

Analysis & design of control for distributed embedded systems under communication constraints

Prateep Kumar Roy

▶ **To cite this version:**

Prateep Kumar Roy. Analysis & design of control for distributed embedded systems under communication constraints. Embedded Systems. Université Paris-Est, 2009. English. NNT : 2009PEST1010 . tel-00534012

HAL Id: tel-00534012
https://pastel.hal.science/tel-00534012

Submitted on 8 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-EST

ÉCOLE DOCTORALE, ICMS

Thèse de doctorat

Spécialité : Sciences de l'Information et de la Communication (CNU 71)

ROY PRATEEP KUMAR

Analysis & Design of Control for Distributed Embedded Systems under Communication Constraints

---

Analyse et Conception de la Commande des Systèmes Embarqués Distribués sous des Contraintes de Communication

Thèse dirigée par Yskandar Hamam

Soutenue le 04 décembre 2009

Jury :

– Marcel STAROSWIECKI	Examinateur
– Konstantin DIMITROV	Examinateur
– Gerard BLANCHET	Rapporteur
– Cosmin IONETE	Rapporteur
– Yskandar HAMAM	Directeur de thèse
– Arben ÇELA	Co-directeur de thèse

Thèse préparée au sein du Département Systèmes Embarqués, ESIEE Paris

Dedicated to My Parents

Résumé

Les Systèmes de Contrôle Embarqués Distribués (SCED) utilisent les réseaux de communication dans les boucles de rétroaction. Étant donné que les systèmes SCED ont une puissance de batterie, une bande passante de communication et une puissance de calcul limitée, les débits des données ou des informations transmises sont bornées et ils peuvent affecter leur stabilité. Ceci nous amène à élargir le spectre de notre étude et y intégrer une étude sur la relation entre la théorie du contrôle d'un coté et celle de l'information de l'autre.

La contrainte de débit de données induit la quantification des signaux tandis que les aspects de calcul temps réel et de communication induit des événements asynchrones qui ne sont plus réguliers ou périodiques. Ces deux phénomènes donnent au SCED une double nature, continue et discrète, et en font des cas d'étude spécifiques.

Dans cette thèse, nous analysons la stabilité et la performance de SCED du point de vue de la théorie de l'information et du contrôle.

Pour les systèmes linéaires, nous montrons l'importance du compromis entre la quantité d'information communiquée et les objectifs de contrôle, telles que la stabilité, la contrôlabilité/observabilité et les performances. Une approche de conception conjointe de contrôle et de communication (en termes de débit d'information au sens de Shannon) des SCED est étudiée.

Les principaux résultats de ces travaux sont les suivants :

- Nous avons prouvé que la réduction d'entropie (ce qui correspond à la réduction d'incertitude) dépend du Grammien de contrôlabilité. Cette réduction est également liée à l'information mutuelle de Shannon.

- Nous avons démontré que le Grammien de contrôlabilité constitue une métrique de l'entropie théorique de l'information en ce qui concerne les bruits induits par la quantification. La réduction de l'influence de ces bruits est équivalente à la réduction de la norme du Grammien de contrôlabilité.

- Nous avons établi une nouvelle relation entre la matrice d'information de Fisher (FIM) et le Grammien de Contrôlabilité (CG) basé sur la théorie de l'estimation et la théorie de l'information.

- Nous proposons un algorithme qui distribue de manière optimale les capacités de communication du réseau entre un nombre "n" d'actionneurs et/ou systèmes concurrents se basant sur la réduction de la norme du Grammien de Contrôlabilité.

Mots-clés: Commande sous les Contraintes de Communication, FIM, Information Réciproque, Entropie, Contrôlabilité Gramian, Théorie de l'Information, Systèmes de Contrôle Embarqués et Distribués (SCED).

Abstract

The Networked Embedded Control System (NECS) uses communication networks in the feedback loops. Since the embedded systems have the limited battery power along with limited bandwidth and computing power, the feedback data rates are limited. The rate of communications can drastically affect system stability. Hence, there is a strong need for understanding and merging the Control Theory with Communication or Information Theory.

The data rate constraint introduces quantization into the feedback loop whereas the communication or computational model induces discrete events which are no more periodic. These two phenomena give the NECS a twofold nature : continuous and discrete, and render them specific.

In this thesis we analyze the stability and performance of NECS from Information-theoretic point of view. For linear systems, we show how fundamental are the tradeoffs between the communication-rate and control goals, such as stability, controllability / observability and performances. An integrated approach of control and communication (in terms of Shannon Information Rate) of NECS or distributed embedded control systems is studied. The main results are as follows :

- We showed that the entropy reduction which is same as uncertainty reduction is dependent on Controllability Gramian only. It is also related to Shannon Mutual-Information.
- We demonstrated that the gramian of controllability constitutes a metric of information theoretic entropy with respect to the noises induced by quantization. Reduction of these noises is equivalent to the design methods proposing a reduction of the controllability gramian norm.
- We established a new relation of Fisher Information Matrix (**FIM**) and Controllability Gramian (**CG**) based on estimation-theoretic and information-theoretic explanations.
- We propose an algorithm which optimally distributes the network capacity between a number "n" of competing actuators. The metric of this distribution is the Controllability Gramian.

Keywords: Control under Communication Constraints, FIM, Mutual Information, Entropy, Controllability Gramian, Information Theory, Networked Embedded Control System (NECS).

Remerciements

J'exprime ma chaleureuse gratitude à l'égard du Professeur Associé Arben Çela dont l'enthousiasme, la disponibilité et la patience m'ont permis de réaliser ce travail Doctoral. Je le remercie également pour les échanges scientifiques fascinants que nous avons partagés et pour son appui continu.

J'exprime également ma gratitude et mes sincères remerciements à l'égard du Professeur Yskandar Hamam pour avoir accepté de diriger cette thèse. Son appui et conseil continus ont été d'une aide et d'une valeur inestimables tout au long de ces trois années de travail.

Je remercie cordialement les membres du jury, particulièrement, Professeur Gérard Blanchet, Professeur Marcel Staroswiecki, Professeur Konstantin Dimitrov et Professeur Cosmin Ionete qui ont gentiment accepté de participer au jury.

Je remercie sincèrement tout le personnel d'ESIEE Paris, pour leur aide apportée au niveau technique, documentation et démarches administratives.

Mes sincères remerciements vont également aux membres du personnel d'ICMS et de l'Université Paris-Est pour leur assistance administrative.

ACKNOWLEDGEMENTS

I express my hearty thanks to Asso. Prof. Arben Çela, who made this Doctoral Work possible and provided me the great enthusiasm for this work. I also thank him for the enthralling scientific discussions which we shared and for his continuous support.

I also express my gratitude and sincere thanks to Prof. Yskandar Hamam to have agreed to direct this thesis. His continuous support and advice was of an invaluable help throughout these three years of work.

I make a point of cordially thanking the members of the Jury, especially, Prof. Gerard Blanchet, Prof. Marcel Staroswiecki, Prof. Konstantin Dimitrov and Prof. Cosmin Ionete who have very pleasantly accepted the task of Reviewers and Examiners as well.

I extend my sincere thanks to all the faculty members and colleagues of Département Systèmes Embarqués, ESIEE Paris. My sincere thanks also go to the supporting Staff Members of this department and the institution ESIEE Paris as a whole for their assistance concerning the administrative, technical and documentation issues.

My sincere thanks also go to the Staff members of ICMS, Université Paris-Est. for their administrative support.

Contents

1 Introduction **1**

1.1 Literature Overview 4

1.2 Overview of Thesis 6

1.3 Outline and Publications 7

1.3.1 Outline 7

1.3.2 Publications 7

2 Basic Concepts Used **9**

2.1 Introduction 9

2.2 Information Theory 10

2.3 Estimation Theory 14

2.4 Control Theory 15

2.5 Conclusion 17

3 Control under Communication Constraints **19**

3.1 Introduction 19

3.2 Control over Networks 20

3.2.1 Control-aware Network 20

3.2.2 Network-aware Control 21

3.3 Quantized Control Systems 22

3.4 Conclusion 24

4 Information-theoretic Analysis of Quantized Control **27**

4.1 Introduction 27

4.2 Information and Entropy 27

4.2.1 Relation between Differential Entropy and Discrete Entropy 28

4.2.2 Entropy and Estimation Error Relation 32
4.3 On the Value of Information 34
4.4 Extracting State Information from Quantized Output 35
4.5 Information-Theoretic Rate Requirement under Asymptotic Quantization 36
4.5.1 Justification for the Equation (4.3) of Mutual Information 37
4.6 Certainty Equivalence Principle 38
4.7 Quantization and Mean Square Error 39
4.8 Conclusion 39

5 Convergence of Information Theory, Communications Theory, Estimation Theory and Control Theory 41
5.1 Introduction 41
5.2 Convergence of Information, Communication, Estimation and Control . . 42
5.3 Achievable Rate of Information Transfer and Bode's Integral Formula . . . 45
5.4 Relation among Entropy, Unstable Poles and Bode-Integral 47
5.4.1 Relation between Entropy and Bode-Integral 47
5.4.2 Relation between Bode-Integral and Unstable Poles 49
5.5 Entropy-Theoretic Control 51
5.6 Mutual Information and Optimal Estimation Error 54
5.7 Conclusion 55

6 Information-Theoretic View of Control 57
6.1 Introduction 57
6.2 Problem Formulation 58
6.3 Preliminaries 58
6.3.1 Mutual Information 59
6.3.2 Shannon Entropy 59
6.3.3 Bode Integral 60
6.4 Information Induced by Controllability Gramian 62
6.5 Conclusion 66

7 Relation of FIM and Controllability Gramian 69
7.1 Introduction 69
7.2 Problem Formulation 70

7.3 Preliminaries . . . 71
7.3.1 Fisher Information Matrix . . . 71
7.3.2 Fisher Information as Metric for Accuracy of a Distribution . . . 71
7.4 Information Induced by FIM and Controllability Gramian . . . 73
7.5 Conclusion . . . 74

8 Degree of Reachability / Observability as a Metric for Optimal Integrated Control and Scheduling of Networked Control Systems 75
8.1 Introduction . . . 75
8.2 Problem Formulation . . . 77
8.2.1 State Representation of Resource-Constrained Systems . . . 79
8.2.2 Lifting of Periodic Systems . . . 80
8.2.3 Reachability / Observability of NCS . . . 82
8.3 Optimal Off-Line Scheduling of Control / Sensor Signals . . . 85
8.3.1 An Example . . . 86
8.3.2 Results Obtained . . . 87
8.4 Optimal On-Line Control and Scheduling of Control Signals . . . 90
8.5 Conclusions and Perspectives . . . 93

9 Conclusion 95
9.1 Summary . . . 95
9.2 Future work . . . 96

Appendices 97

A Analysis of First-Order Information 99

B Receding Horizon Optimal Control with Constraints 107
B.1 Introduction . . . 107
B.2 The Receding Horizon Control Principle . . . 108
B.3 Stability of Receding Horizon Optimal Control . . . 112
B.4 Conditions for Stability . . . 114
B.5 Terminal Conditions for Stability . . . 114
B.6 Conclusion . . . 115

Bibliography **117**

List of Figures

2.1 Feedback Loop 16

3.1 Networked Embedded Control System (NECS) 20
3.2 Network-Aware Control Architecture 21

4.1 Information Gathering 34

5.1 Sensitivity Diagram 49

6.1 (a)Feedback loop and (b)Sensitivity function 61

8.1 Schematic Representation of a Resource-Constrained System 77

B.1 Receding Horizon Control 109
B.2 Receding Horizon Control (RHC) Principle 111
B.3 Receding Horizon Control in Closed Loop 112

List of Tables

8.1 Traces of Inverse Reachability Gramians W_r^{-1} w.r.t. Scheduling Sequences $\mathcal{S}_i$ 87
8.2 Maxima-Minima Moduli of Eigenvalues of Inverse Reachability Gramians W_r^{-1} w.r.t. Scheduling Sequences $\mathcal{S}_i$ 88
8.3 Traces of Inverse Observability Gramians W_0^{-1} w.r.t. Scheduling Sequences $\mathcal{S}_i$ 89
8.4 Maxima-Minima Moduli of Eigenvalues of Inverse Observability Gramians W_o^{-1} w.r.t. Scheduling Sequences $\mathcal{S}_i$ 89

1 Introduction

In recent years, there has been an increased interest for the fundamental limitations in feedback control. A *dynamical system* is a system whose behaviour changes over time, often in response to the external stimulation or excitation. The term *feedback* refers to the situation when two or more dynamical systems are connected together such that each system influences the other and their dynamics are strongly coupled. At its core, control theory is an *Information Science* and includes the use of information in both analog and digital representations. According to researchers, control is an encoding problem and decoding is an estimation problem as well. In the words of Kalman, "*Our ultimate objective is to answer question : What kind and how much information is needed to achieve a desired type of control?*"
Traditional feedback control systems pay less attention to issues related to the flow of information through the feedback loop. Whereas the quality and reliability of information transmitted in a communication network is typically the primary focus, feedback control systems study the performance of manipulating a dynamical system with information that is generally assumed to be communicated through ideal channels (Loss-free, error-free and delay-free). Conventional control engineers generally assume perfect transmission of information within the closed loop and that data processing is done with zero time delay. On the other hand, in communication networks, data packets that carry the information can be dropped, delayed, or even reordered due to the network traffic conditions and limited bandwidth. As new technologies and applications emerge, the fields are coming closer, and we are indeed witnessing that the "next phase of the information technology revolution is the convergence of communication, computing, and control" [Murray et al., 2003].
With this convergence of the fields, feedback loops are now being implemented with information passing through communication networks. New applications utilizing these

types of feedback loops include future battlefield systems, urban search and rescue, internet-based control, "smart homes", sensor networks, unmanned air vehicles, multi-vehicle systems, and many more [Murray et al., 2003]. Introducing communication networks in the feedback loops gives several advantages, including modularity and reconfigurability of system components, simple and fast implementations, powerful system diagnosis tools, etc. Of course, a main advantage of Networked Control Systems (NCS) is that they allow for the use of control systems with spatially distributed components, i.e., actuators, sensors, processing units and plants that do not need to be physically collocated. In the case of these components being embedded systems we can call the system as Networked Embedded Control System (*NECS*). There are, of course, potential issues that arise when closing the loop around imperfect communication links, including data dropout, delays, and quantization effects. Researchers [Antsaklis and Baillieul, 2007] [Hespanha et al., 2007] have addressed many of these issues indicating that the information flow management through the loop becoming equally important as the design of the controller. Quality of information (relevance to the context and time) directly affects the performance of the system and specially when the information carrying channel has limited bandwidth. Hence, we must focus on the informational aspects related to control. We can exploit the information flow management to reduce the uncertainty of the system by increasing the degree of controllability /observability.

It has been known that control theory and information theory share a common background as both theories study signals and dynamical systems in general. One way to describe their difference is that the focal point of information theory is the signals involved in systems while control theory focuses more on systems which represent the relation between the input and output signals. Thus, in a certain sense, we may expect that they have a complementary relation. For this reason, many researchers have studied the interactions of the two theories : Control Theory and Information Theory.

In networked control systems, there are issues related to both control and communication since communication channels with data losses, time delays, and quantization errors are employed between the plants and controllers [Antsaklis and Baillieul, 2007]. To guarantee the overall control performance in such systems, it is important to evaluate the quantity of information that the channels can transfer. Thus, for the analysis of networked control systems, information theoretic approaches are especially useful, and notions/results from this theory can be applied. The results in [Nair and Evans, 2004] and [Tatikonda and Mitter, 2004] show the limitation in the communication rate for the existence of controllers, encoders, and decoders to stabilize discrete-time linear feedback systems.

The focus of information theory is more on the signals and not on their input-output relation. Thus, based on information theoretic approaches, we may expect to extend prior results in control theory. One such result can be found in [Martins et al., 2007], where a sensitivity property is analyzed and Bode's integral formula [Bode, 1945] is extended to a more general class of systems. A fundamental limitation of sensitivity functions is presented in relation to the unstable poles of the plants. In fact the unstable poles contribute to the sensitivity and not the stable ones. Unstable systems carry more information than

the stable ones which normally carry no information with respect to the stabilization objectives.
Recent emerging applications, such as sensor networks, micro-electromechanical systems (MEMS), mobile telephony, distributed networked embedded systems and industrial control networks, have posed a challenge to the validity of the modular approach. In these applications, the aim is to control one or more dynamical systems, using multiple sensors and actuators transmitting and receiving information over a digital communication network. Although the total communication capacity in bits per second may be large, each component is effectively allocated only a small portion. This can introduce large quantization errors affecting the control performance, due to the low resolution (quantization accuracy) of the transmitted data. Quantization errors are not new in control theory, and there exists a significant amount of work in which quantization is modeled as extra additive white noise, thereby allowing the standard solutions of stochastic control to be applied [Curry, 1970]. Though this approach is reasonable if the quantizer resolution is high, it is invalid if the resolution is coarse and the open-loop dynamics are unstable. There exists a critical positive data rate below which there does not exist any quantization and control scheme able to stabilize an unstable plant [Nair et al., 2007].

Quantization converts regions of real numbers (analog signals) into discrete points (digital signals) via a finite set of integers. Some of the information of the signal is lost due to this quantization process, and it affects the closed loop system. Low communication capacity coupled with quantization errors severely affect the attainable control performance. Hence, there is a strong need for analyzing the control and communication aspects jointly rather separately. There are many issues related to the network effects (delay, packet-loss, bit errors, jitter, quantization noises, channel induced noises etc.) of NCS but we focus here mainly on the limitations caused by constrained communication rate. There is a minimum communication rate

$$R \geq \sum_{\lambda(A)} \max\left\{0, \log_2 |\lambda(A)|\right\} \text{ in bits/sec} \tag{1.1}$$

for a linear discrete-time system [Tatikonda and Mitter, 2004] above which the system may be stabilized like Shannon Source Coding Theorem [Haykin, 2001] [Cover and Thomas, 2006] [Shannon, 1948] of communication which states that there is a minimum data rate above which a given random process can be reliably (within small tolerable limit of bit error probability, say, 10^{-6} to 10^{-10}) communicated. Here, $\lambda(A)$ is any eigenvalue of the discrete model plant matrix A in a state-space representation. The basic difference between communication and control is that the information is merely transmitted from sender to receiver in the former but it is also used (apart from just transmitting) in the feedback loop in the latter. Shannon's Rate Distortion Theorem in simple words "one can reduce the communication rate but below the threshold limit which is the minimum communication rate, the probability of the distortion of the message to be communicated is higher and if the rate is still lowered (below the term *Shannon Entropy*) then the distortion in the communicated message is increased". Shannon En-

tropy (uncertainty or randomness) puts the lower limit on such lowering of the communication rate to have an acceptable quality of communication. Applying the same principle in control (which the researchers have already proved) there is a trade-off between communication rate and the control performance which needs to be measured and characterized mathematically. Basic aim of this research work is to analyze the relation between control performance (in terms of stability, controllability, state norms etc.) and information-theoretic parameters like Shannon information rate, Shannon entropy and even Estimation-theoretic parameter Fisher Information Matrix (**FIM**). However, before we proceed for main work in details we would like to provide a brief overview of the related work in the literature.

1.1 Literature Overview

In modern control theory, the data rate theorem refers to the smallest feedback data rate above which an unstable dynamical system can be stabilized. In paper [Minero et al., 2009] a data rate theorem for stabilization of a linear, discrete-time, dynamical system with arbitrarily large disturbances, over a rate-limited, time-varying communication channel was presented. Necessary conditions were derived employing information-theoretic techniques, while a stabilization scheme based on an adaptive successively refinable quantizer was constructed. The survey paper [Nair et al., 2007] gives a historical and technical account of various formulations. In [Delchamps, 1990] stability studies were made when closing the feedback loop with quantized measurement signals. First results on minimum data rates for stabilizability appeared in [Baillieul, 1999], [Wong and Brockett, 1999], where it was shown that a noiseless scalar plant with parameter $|A| \geq 1$ can be kept bounded by memoryless quantized control if and only if the available data rate exceeds $\log_2 A$ bits per sample. These results were the first instances of the Data Rate Theorem. The work in [Nair and Evans, 2000], [Nair and Evans, 2003], [Nair and Evans, 2004] studied the stochastic stability of feedback control systems with limited data rate and show the relation to quantization theory. In addition to the limited data rate effect, they introduced system process and measurement noise into their model. Taking an information-theoretic point of view, the thesis [Sahai, 2001] derived stability conditions based on anytime information which quantifies the "time value" of data bits. Ph.D. thesis [Tatikonda, 2000] analyzed the necessary data rate and coding schemes to stabilize a plant across a noisy channel. A general extension of Bode's integral inequality is provided in [Martins and Daleh, 2008] to assess the performance limitations of feedback control over finite capacity memory-less channels. Researchers also consider issues such as nonlinear systems [Nair et al., 2004], robustness to plant uncertainties [Phat et al., 2004] and disturbance attenuation [Martins and Daleh, 2005]. Due to the nature of transmitting signals across networks, the information passed through the feedback loop can be delayed. In [Branicky et al., 2000], the authors analyzed the influence of the sampling rate and network delay on system stability. In [Nilsson, 1998] the author analyzed delays that are either fixed or random according to a Markov chain.

The LQG optimal control problem was solved for the different delay models. Authors in [Luck and Ray, 1990] compensated for delays by placing observers throughout the loop. In [Liou and Ray, 1991a], [Liou and Ray, 1991b], the authors used a stochastic approach to study time-varying delays. Authors in [Montestruque and Antsaklis, 2004] studied the stability in the presence of time varying delays that could be driven by an underlying Markov chain. In [Montestruque and Antsaklis, 2002a], [Montestruque and Antsaklis, 2002b], [Montestruque and Antsaklis, 2003], the maximum time between samples was determined to ensure stability for periodic sampling.
The improvement from boundedness to asymptotic stability became possible by permitting the quantizer or encoder to possess memory, and follow an adaptive zooming-in/zooming-out strategy as in [Brockett and Liberzon, 2000], [Liberzon, 2003], [Petersen and Savkin, 2001]. This is based on dynamically adjusting the range of the quantizer so that it increases as the plant state approaches the target (zooming-in phase), and decreases if the state diverges from the target (zooming-out phase). The underlying intuition is that, in order to drive the state to the target, the quantizer resolution (Quantization-Accuracy) should be high close to the target but coarse far from it [Nair et al., 2007]. Trade-off between sampling-rate and quantization precision in NCS to optimize the information of the system is given in [Ben Gaid and Çela, 2006]. These techniques can be adapted to design controllers yielding guaranteed rate of state convergence. The concept of *topological entropy* in limited data-rate control has been used in [Savkin, 2006] [Nair et al., 2004]. Topological entropy, in some sense, measures the fastest rate at which the uncertainty about the initial state can be reduced, or equivalently the least rate at which the initial state information can be generated.
All the above-mentioned results concern plants that are deterministic apart from a possibly random initial condition. For stochastic plants, the major contributions are in [Martins and Daleh, 2005], [Borkar and Mitter, 1997], [Matveev and Savkin, 2004], [Nair et al.,], [Nair and Evans, 2004], [Tatikonda et al., 2004], [Tsumura and Maciejowski, 2003]. In [Borkar and Mitter, 1997], data rate-limited control of partially observed linear Gaussian systems is considered under a quadratic cost. It is shown there that if the measurements are passed through a minimum variance filter, and the input to the quantizer is chosen to be the innovations of the filter process, then the design of the coding and control laws can be performed separately. Separation and certainty equivalence for linear Gaussian plants are addressed in a more general way in [Tatikonda et al., 2004], which also presents rate-distortion-theoretic lower bounds on performance over additive white Gaussian noise channels and high rate noiseless digital channels. In the paper [Nair and Evans, 2004], the mean square stabilizability of linear plants with possibly non-Gaussian noise is considered. By exploiting the properties of differential entropy power, a universal lower bound is obtained on the time-asymptotic mean square state norm. In particular, this bound implies that as the data rate approaches the intrinsic entropy rate

$$\dot{H} := \log_2 |\det A|$$

of the plant, the mean square state becomes arbitrarily large, regardless of the coding and control scheme. Here, coding refers to the translation of processes into alphabet of binary strings. [Martins and Daleh, 2005] also shows that the Shannon Channel Capacity C (which is the maximum possible data rate the channel can support with tolerable level of error) of the feedback channel decreases towards H, in other words the ability of the controller to produce a stabilizing signal diminishes.The possibility of getting tight bounds on the data rate needed to stabilize a system is based on the use of dynamic encoders and controllers with unlimited memory. Under memoryless, finite-level quantization, the set of reachable points is discrete or at most dense [Anzai, 1974], [Delchamps, 1989], [Bicchi et al., 2002], and only practical stability can be achieved, namely states in some initial set can be driven to a smaller target set, and not asymptotically to the origin [Baillieul, 2001], [Nair et al., 2007], [Fagnani and Zampieri, 2003], [Picasso et al., 2002], [Wong and Brockett, 1999]. The main difficulty in such case is due to the fact that the performance should be evaluated depending on the indices: (1) the steady-state properties of the closed-loop system, and (2) the quality of the transient [Fagnani and Zampieri, 2004] [Fagnani and Zampieri, 2005]. This prevents the existence of a unique optimal controller, since it would generally depend on the weights associated with the two indices. According to the relative weights assigned to the steady state and to the transient, there are three different optimal strategies: the first based on the uniform quantizer, widely used in applications, the second on the logarithmic quantizer [Elia and Mitter, 2001], similar to the μ- and A-law companders of communications [Haykin, 2001], and the third on the chaotic quantizer as in [Fagnani and Zampieri, 2003] and [Fagnani, 2004]. Memoryless quantization has been discussed [Delvenne, 2006] in a new perspective making the problem much more treatable.

1.2 Overview of Thesis

The aim of the work is to present the information-theoretic explanation of analysis and design of control for distributed embedded systems (DES) under communication constraints. Our main contribution in this research work is to establish a relation between information-theoretic parameters such as Shannon Entropy, Shannon Information and the concept of degree of controllability/observability via Controllability / Observability Gramian. We have established a relation between estimation-theoretic parameter like Fisher Information Matrix (**FIM**) and the concept of degree of controllability/observability via control-theoretic Controllability / Observability Gramian. Controllability is binary in nature as defined by Kalman, that is, by observing the rank of controllability gramian matrix one can conclude that a system is controllable or not. But, degree of controllability is more informative than just concluding a system is controllable or not. We have also presented the Degree of Controllability/Observability as metrics for optimal integrated control and scheduling of NCS and analyzed some interesting results related to scheduling based on the Reachability/Observability Gramians and their Eigenvalues.

1.3 Outline and Publications

1.3.1 Outline

The thesis is having the following remaining chapters as organized below :
In Chapter 2, we discuss the basic concepts of Information Theory, Communications Theory, Estimation Theory and Control Theory which we have used in our work.
In Chapter 3, we analyze the problem of Control under Communication Constraints and discuss quantized control systems.
In Chapter 4, we analyze and discuss the Quantized Control from Information-theoretic viewpoint.
In Chapter 5, we discuss the convergence of fundamental limits of Information Theory, Communications Theory, Estimation Theory and Control Theory.
In Chapter 6, we formulate and analyze the Information-theoretic view of Control.
In Chapter 7, we analyze and establish the relationship between the *FIM* and Controllability Gramian.
In Chapter 8, we analyze and discuss the Degree of Reachability / Observability as a Metric for Optimal Integrated Control and Scheduling of NCS.
In Chapter 9, we provide the conclusion and future work.

1.3.2 Publications

The chapters 3, 5, 6, 7, 8 of this thesis are based on our following publications :

- Prateep Roy and Arben Çela, A Survey on Control under Communication Constraints, *6th International RoEduNet Conference*, University of Craiova, Romania, November 2007. **(Chapters - 3 & 5)**

- Prateep Roy, Arben Çela and Yskandar Hamam, Information-Theoretic View of Control, *6th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2009)*, Milan, Italy, July 2009. **(Chapter - 6)**

- Prateep Roy, Arben Çela and Yskandar Hamam, On the Relation of FIM and Controllability Gramian, *IEEE International Symposium on Industrial Embedded Systems (SIES 2009)*, EPFL, Lausanne, Switzerland, July 2009. **(Chapter - 7)**

- Arben Çela, Abdel Reama, Silviu Niculescu, and Prateep Roy, Degree of Reachability / Observability as a Metric for Optimal Integrated Control and Scheduling of Networked Control Systems, submitted to *IFAC Symposium on System, Structure and Control*, Ancona, Italy, 2010. **(Chapter - 8)**

2
Basic Concepts Used

2.1 Introduction

Information is a central component of most decision processes. Intuitively, information is needed to conduct a decision task in the same way that directions are needed to find our way to a precise location in an unknown place. Each bit of information gathered directly from a dynamical system by a controller can serve to decrease the entropy of the controlled system by at most one bit additional to the reduction of entropy attainable without such information (i.e., in open-loop control). The notion of Maxwell's demons (demons are nothing but the concept of getting the information to reduce the entropy of the system) generalizes to controllers and extends the scope of the second law of thermodynamics to control problems.
Channel capacity is the maximum data transmission rate across a communication channel with the probability of error approaching zero, and the rate distortion function is the minimum information transfer rate needed to describe a source under a distortion constraint. Here, distortion constraint refers to the minimum distortion (in terms of error) because with the reduced information transfer rate there is a high probability that there would be errors. Limit to the maximum information transfer rate is the channel capacity and limit to the minimum information transfer rate (i.e. data compression) is the entropy. The two fundamental limits of data communication and data compression are dual.

Sensors are first used to gather information about the state of the system to be controlled (observation step); this information is processed according to a determined control strategy (decision step), and then fed back to actuators which try to update or redirect the state of the system. Such control systems paired with information-processing

devices are found in many modern control systems, ranging from sophisticated automatic flight guidance systems to very simple servomechanisms such as thermostats. In the former example, the sensors are the various measurement devices (e.g., altimeter, speedometer) which provide the navigation unit with the necessary data to steer a plane, whereas in the latter the sensor is just a thermometer. Information is required in control by demonstrating that it is possible to quantify exactly the amount of information gathered by a controller, and that there exists a direct relationship between the amount of information gathered and the performance of the controller.

According to the paper [Saridis, 1988] the cost (resource utilization in terms of complexity of computation and communication, bandwidth usage, delay incurred etc.) of control with inaccessible states is given by the cost of control with estimated states plus the cost of state estimation minus the equivocation for details) of active transmission of information (Chapter - 5). Here, conditional entropy or equivocation is the residual entropy (or uncertainty) of a random variable given that the other random variable is known.

2.2 Information Theory

Information is measured as

$$I_k = \log_2 \frac{1}{p_k}$$

where I_k and p_k are the information associated with the k-th symbol of the message-word and the probability of occurring the k-th symbol of message-word respectively (Note : A message consists of words, a word consists of symbols and a symbol is represented by bits). Evidently, higher the probability of occurrence, smaller is the information associated with and vice-versa. We code more frequently used symbols with fewer number of bits and vice-versa.

Shannon proposed a measure of uncertainty in a discrete distribution based on the Boltzmann entropy of classical statistical mechanics. He called it the entropy and defined as follows.
We have to take into account the statistics of the alternatives by replacing our original measure of the number of alternatives by the more general expression defining the entropy as follows :

$$H = -\sum_i p_i \log_2 (p_i)$$

where p_i is the probability of the alternative i. This measure of uncertainty has many important properties which agree with our intuitive notion of randomness. We mention three :

1. It is always positive.

2. It vanishes if and only if it is a certain event.

3. Entropy is increased by the addition of an independent component, and decreased by conditioning. Conditioning means *a priori* information as precondition which helps in reducing the entropy or the uncertainty.

The above quantity is known as the *binary entropy* as we use logarithmic base of 2, and was shown by Shannon to correspond to the minimum average number of bits needed to encode a probabilistic source of N states distributed with probability p_i.
Intuitively, H can also be considered as a measure of uncertainty : it is minimum, and is equal to zero, when one of the alternatives appears with probability one, whereas it is maximum and equals to $\log_2 N$ when all the alternatives are equiprobable so that $p_i = \frac{1}{N}$ for all i.
This interpretation of entropy is of foremost importance here. The term entropy is associated with the uncertainty or randomness whereas information is used to reduce this uncertainty. Thus information and entropy (i.e. uncertainty) are complementary aspects of control. At the lowest level of description, controllers can arguably be thought of as devices aimed at reducing the uncertainty associated with a system. Control reduces uncertainty, reduced uncertainty means higher degree of controllability. Uncertainty is nothing but entropy and control uses information to reduce entropy. We cannot control anything unless we measure that. Hence, to measure or to quantify 'information' in terms of mutual information we use entropy (Shannon) and probability associated with the information. How much information is required in control? How much information is required for observing and controlling the system based on observed informations?

Mutual-Information, the term coined by Shannon, is the information carried by one random variable about the other.

Mutual-Information $I(X;Y) = H(X) + H(Y) - H(X,Y)$

where $H(X)$ is the uncertainty that X has about Y, $H(Y)$ is the uncertainty that Y has about X, and $H(X,Y)$ is the uncertainty that X and Y hold in common. Information-theoretically, $H(X)$, $H(Y)$ are called the respective entropies associated with X and Y and $H(X,Y)$ is the joint entropy between X and Y.
Mutual Information $I(X;Y)$, between X as the input variable and Y as the output variable, has the lower and upper bounds given by the following:

Lower Bound of Mutual Information :

$R(D) =$ Rate Distortion $=$ Min $I(X;Y)$

where D is the distortion which happens when information is compressed (i.e. fewer bits are used to represent or code more frequent or redundant informations) and entropy is the limit to this compression i.e. if one compresses the information beyond the

entropy limit there is a high probability that the information will be distorted or erroneous. This is as per Shannon's Source Coding Theorem.

Upper bound of Mutual Information :

$C =$ Communication Channel Capacity $=$ Max $I(X;Y)$

Mutual information is also the difference of entropies (as explained below), where entropy is nothing but the measure of uncertainty. The conditional version of the chain rule [Cover and Thomas, 2006] is written as :

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)\ ;$$

valid for any random variables X and Y.

Mutual information $I(X;Y)$ is the amount of uncertainty in X, minus "the amount of uncertainty in X which remains after Y is known", which is also "the amount of uncertainty in X which is removed by knowing Y". Thus, mutual-information is the difference between two entropies and helps in reducing the uncertainty (entropy). This corroborates the intuitive meaning of mutual information as the amount of information (that is, reduction in uncertainty) that each variable is having about the other.
The *Conditional Entropy* $H(X|Y)$ or read as conditional entropy of X knowing Y or conditioned on Y, is often interpreted in communication theory as representing an information-loss (the so-called *equivocation* of Shannon [Shannon, 1948]), which results from subtracting the maximum noiseless capacity $I(X;X) = H(X)$ of a communication channel with input X and output Y from the actual capacity of that channel as measured by $I(X;Y)$.
Just as entropy [Middleton, 1960] in physical systems tends to increase in the course of time, the reverse is true for information about an information source : as information about the source is processed, it tends to decrease with time, becoming more corrupt or noisy until it is evidently destroyed unless additional information is made available. Here, information refers to the case of desired messages. Information value degrades over time and entropy value increases over time in general. We can use the timely information (information as and when it is generated) through feedback-loop to reduce the uncertainty and thus achieving better controllability because uncertainty in terms of information-theoretic entropy is the obstruction in the path of controllability. Lack of information (of appropriate quality and time-value) is the degradation in control performance in terms of controllability, stability etc.

In Ph.D. thesis [Lloyd, 1988], entropy reduction has been described as equivalent to state reduction in line with Maxwell's demon and thermodynamic concept of microscopic states of gas and thermal entropy. Because entropy is nothing but state uncertainty and reducing entropy is thus reducing the states which are not really relevant or uncertain to a great extent. Entropy can also be interpreted as a measure of our lack of

knowledge of the actual state of the system because of the uncertainty associated with the term entropy and we reduce the entropy (uncertainty) by gaining information.

Hence, $\Delta I = -\Delta H$.

Here, Δ refers to the change and the positive and negative signs indicate gain (increase) and loss (decrease) respectively. Thus, increase in information is accompanied with decrease in entropy and vice-versa. Hence, there is a trade-off between information and entropy.

It has shown via a physical analogy that a law of non-decrease of entropy (i.e. $H \geq 0$) need not apply to such systems in the presence of observations that continue to supply new information. Because new informations gained by observations lead to reduce the uncertainties in the system and hence leads to reduction in entropy.

Since mutual information is not dependent on any underlying reference measure (such as Lebesgue measure) [Mitter and Newton, 2005] i.e. it can be thought of as being the observation-derived information on state X_t stored by the system at time t. It therefore has absolute meaning, unlike quantities such as the signal entropy which only has meaning relative to the reference measure used in its definition. So, for example, if mutual information were zero, it would imply that the observations up to time t , at which the mutual information is measured, were completely useless for estimating the signal value at time t. Mutual information is shown [Mitter and Newton, 2005] to be the difference between the information supply and the information dissipation. Here, the dissipation process represents observation-derived information that was useful for estimating the past of X, but is of no use in estimating its future. We can take the view that entropy is simply unobservable information because entropy or uncertainty is caused by unobservability. However, this is reduced in the presence of the partial observations. Mutual information is increased by the supply of new information, and reduced by the dissipation of historical information.

Mutual Information gain can be used as a control metric and is an *a priori* measure of average information gain following observation.
It is given by :

$$I(x(t); z(t)) = E\left[\log\left\{\frac{P(x(t)|z(t)}{P(x(t)}\right\}\right]$$

which basically measures the compression of information. Choosing the sequence of observations $z(t)$ which maximizes mutual information gain over a horizon. Observations depend on platform state $x(t)$ and state is governed by some control input $u(t)$. Therefore, by choosing $u(t)$ suitably we can maximize the information gain.

2.3 Estimation Theory

When we cannot get the exact information or the information gets distorted / corrupted / lost due to communication induced noises and the channel properties, we need to have the estimation of the information and hence we need the Estimation Theory. Fundamental limit of Estimation Theory is the *Fisher Information Matrix* (**FIM**) and the related parameter called *Fisher Information* is of paramount importance.
Fisher information is the amount of information that an observable random variable Y carries about an unobservable parameter X upon which the likelihood function of X represented by $L(X) = p(Y;X)$ depends. It is a measure of accuracy in estimating a parameter. Fisher Information Matrix is given by

$FIM_x := -E(\frac{\partial \log p(x,y)}{\partial x})^2$

is the (Bayesian) FIM, where x is the input state, y is the observation at the output and E is the expectation operator .

Also, **FIM** $= -E\left[\left\{\frac{\partial \ln \rho(f)}{\partial x}\right\}\left\{\frac{\partial \ln \rho(f)}{\partial x}\right\}^T\right]$

$\rho(f)$ being the Gaussian Probability Density Function of random variable f. Thus Fisher information is the negative of the expectation of the second derivative of the $\log$ of probability p function with respect to x. Information may thus be seen to be a measure of the "sharpness" of the support curve near the maximum likelihood estimate of x. A "blunt" support curve (one with a shallow maximum) would have a low expected second derivative, and thus low information; while a sharp one would have a high expected second derivative and thus high information. Second derivative signifies the curvature which is defined as

Curvature or bending of a trajectory $\kappa = \frac{\ddot{y}}{(1+(\dot{y})^2)^{3/2}}$

where, y, $\ddot{y}$ and $\dot{y}$ denote the function $y = f(x)$, its second derivative and first derivative (slope) respectively. If the slope is small compared with unity, so that the approximation of the curvature results in the second derivative. Mathematically,

$\dot{y} < 1 \Rightarrow (\dot{y})^2 << 1 \Rightarrow \kappa \approx \ddot{y}$.

Cramer-Rao Bound (*CRB*) gives the lower bound of variance σ^2 as

$\sigma^2 \geq \frac{1}{FIM}$

Idea behind *CRB* : There is a "Gain Factor" that describes how much a change in the parameter (to be estimated) changes the probability distribution of the quantities being

observed. Quality of any estimate is limited both by variance of the noise and by the magnitude of the change in the probability distribution caused by a change in quantity being estimated.
Fisher Information (*FI*) as defined in [Stam, 1959]

$$FI = -\int p(x_1, x_2, \ldots, x_n, \theta) \frac{\partial^2}{\partial \theta^2} \ln p(x_1, x_2, \ldots, x_n, \theta) dx_1, dx_2, \ldots, dx_n$$

where, $x_1, x_2, \ldots, x_n$ being the observable variables and θ being the parameter getting estimated.

2.4 Control Theory

Kalman in his pioneering paper [Kalman, 1960] has given a nice definition of the state of a dynamical system. The state space is the smallest collection of state variables which must be specified at time $t = t_0$ in order to be able to predict the behaviour of the system for any time $t \geq t_0$. In other words, the state is the minimal "record" of the past history needed to predict the future behaviour.
Physically a stable system needs no information on its internal state or the environment to assure its stability. So, if we consider a well designed stable feedback control system with disturbances or/and noises as inputs and performance signals as outputs then it is not needed to have extra feedback loop to assure its stability. We may say the same thing for systems which are open-loop stable. For example, a pendulum with non-zero friction coefficient subject to a perturbation will return back to the equilibrium position after a transient period without any need of extra information. For unstable systems the mutual information between the initial state and the output of the system is related to its unstable poles.
Bode Sensitivity is the fundamental limitation of control theory which bears a correlation with information-theoretic entropy. It is established by researchers [Wu and Jonckheere, 1992], [Mehta et al., 2006], [Iglesias, 2001], [Iglesias, 2002], [Okano et al., 2008] that unstable poles do contribute to Bode sensitivity integral and stable poles do not. The simplest (and perhaps the best known) result is that, for an open-loop stable plant, the integral of the logarithm of the sensitivity is zero; i.e.

$$\int_0^\infty \ln |S_0(j\omega)| \, d\omega = 0$$

Where, S_0 and ω being the sensitivity function and frequency respectively.
Now, we know that the logarithm function has the property that it is negative if $|S_0| < 1$ and it is positive if $|S_0| > 1$. The above result implies that set of frequencies over which sensitivity reduction occurs (i.e. where $|S_0| < 1$) must be matched by a set of frequencies over which sensitivity magnification occurs (i.e. where $|S_0| > 1$). Referring to Fig. (2.1) for *SISO* system with a stable rational transfer function $L(j\omega) = P(j\omega) C(j\omega)$, sensitivity is defined as

$S(j\omega) = \frac{1}{1+L(j\omega)}$.

This has been given a nice interpretation as thinking of sensitivity as a pile of dirt. If we remove dirt from one set of frequencies, then it piles up at other frequencies. Hence, if one designs a controller to have low sensitivity in a particular frequency range, then the sensitivity will necessarily increase at other frequencies – a consequence of the weighted integral always being a constant; this phenomenon has also been called the Water-Bed Effect (pushing down on the water bed in one area, raises it somewhere else).
For linear systems Bode Integral is the difference in the entropy rates (entropy per unit time) between the input and output of the systems which is an information-theoretic interpretation. Entropy of the signals in the feedback loop help provide another inter-

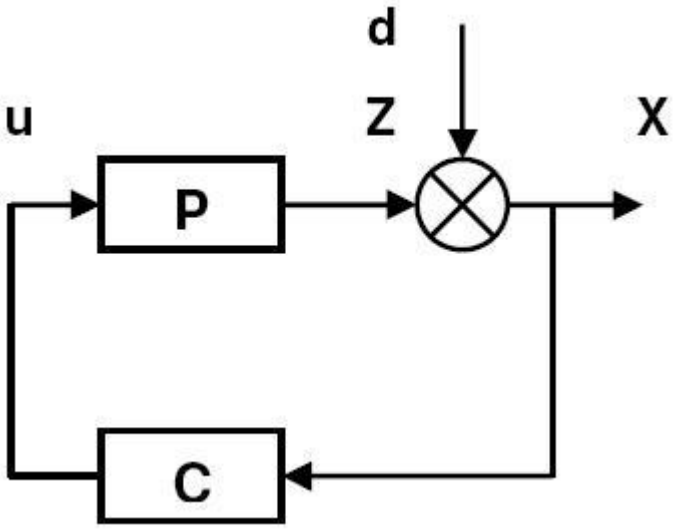


Figure 2.1: Feedback Loop

pretation of the Bode integral formula [Zang and Iglesias, 2003] [Mehta et al., 2006] as follows. Shannon Entropy - Bode Integral Relation (for derivation refer Chapter 5 of this thesis) can be rewritten as : For an LTI discrete-time system —-

$$H_c(x) - H_c(d) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left|S\left(e^{j\omega}\right)\right| d\omega = \sum_k \log(p_k) \tag{2.1}$$

Where $S(e^{j\omega})$ is the transfer function of the feedback loop from the disturbance d to output x and p_{ui} 's are unstable poles ($|p_k| > 1$) of the open-loop plant; S is referred to as the sensitivity function for an open-loop plant gain P and a stabilizing feedback controller gain C, S is given by $S = \frac{1}{1+PC}$. Sensitivity shows how much sensitive is the observable output state to input Gaussian disturbance. Here, $H_c(x)$ and $H_c(d)$ denote the conditional entropy of the random processes associated with the output x and disturbance d respectively as per Fig. (2.1) [Mehta et al., 2006].

The problem of associating the degree of controllability or the degree of observability was considered in an early study by Kalman et al. in paper [Kalman et al., 1963] where the *trace* of $\mathbf{W}_c^{-1}$ and the *determinant* of $\mathbf{W}_c^{-1}$ were proposed as two possible options for a quantitative measure of complete controllability of the linear dynamical systems. Here, $\mathbf{W}_c$ being the controllability gramian matrix. These are measures of information which we are discussing in later chapters of this thesis. Controllability, observability, controllability gramian and observability gramian are given in Annexure - A of the thesis.

2.5 Conclusion

This chapter introduced the basic concepts of Information / Communication Theory, Estimation Theory and Control Theory used in our research work.
The Internet, wireless networks, and the like are making it possible to have many remote plants linked together via communication channels. Thus a substantive theory of distributed control is increasingly becoming important in today's control systems. An important aspect of distributed control is the role of communication between the different components in the system. Since control problems have strict delay constraints therefore it is important to find the smallest possible rate or find the minimum rate required on the channel to achieve the desired control objectives or find the largest tolerable state estimation error such that the control objectives can still be met.

3 Control under Communication Constraints

3.1 Introduction

In this chapter, we present our abstract view of a distributed embedded control system operating under communication constraints. This abstract view is described by the class of *computer-controlled systems*, which was introduced in [Hristu, 1999]. This class allows to model, in a finely-grained and abstract way, the impact of the resource limitations on the behavior of the controlled system. In Distributed Embedded Systems (*DES*) the plant (**P**) or the process being controlled, the sensor (**S**), the actuator (**A**), the controller (**C**) — each of these may be one or many in number, are distributed in space and use shared communication network(s) — wired or wireless with limited use of power supply (Battery!), even computational/processing power and communication bandwidth. Limited power and bandwidth are the communication constraints which cause packet-drop or information loss, delay (*latency*), delay-variation (*jitter*), quantization noise etc. More communication means more frequent transfer of information which implies more consumption of limited power supply (e.g. battery power), this is in addition to the computational overhead in the context of the power supply. Refer to the Fig. (3.1) for Networked Embedded Control System (NECS). [1]

A unified view of control and communication clarifies many of the conceptual issues underlying the distributed control problem. We discussed in [Roy and Çela, 2007]

[1] Actuators(A),Sensors(S), Controller(C) are distributed in space, connected by Communication Network for Information sharing, powered on their own and using Embedded System Technology (having controller, processor, memory and power source embedded in **SoC**(System-on-Chip) form)

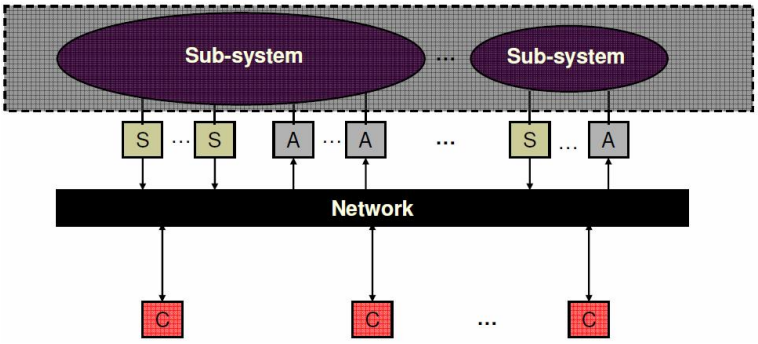


Figure 3.1: Networked Embedded Control System (NECS)

our survey result on control under communication constraints in the light of the interaction between information and control. Our main goal is to understand the fundamental limits of control performance in distributed systems when there are communication constraints. Main emphasis is on the practically achievable control performance issues rather than just stability issues wherein many work have already been done. We are eager to find the satisfactory controllability and observability of the system with achievable optimality of information exchange among the various subsystems with the emphasis on performance of the system under communication constraints.

3.2 Control over Networks

There are two complimentary approaches to control over networks (whether wired or wireless) :

1. Control-aware Network
2. Network-aware Control

3.2.1 Control-aware Network

By Control-aware networking we mean layered communication models and control of communication resources capable of communicating / networking taking into account the hard / soft real-time constraints of control. We need to modify the network protocols and communication links for better real-time performance, for example, using CAN (Controller Area Network) from Bosch and IWLAN (Industrial Wireless Local Area Network) from Siemens. We have to investigate the suitability of the type of communication network protocols needed at various layers suitable for our control applications.

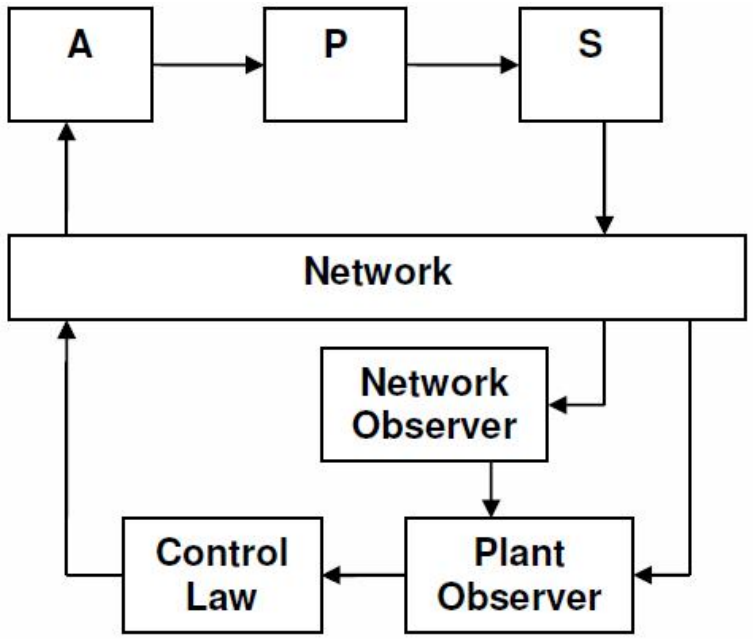


Figure 3.2: Network-Aware Control Architecture

3.2.2 Network-aware Control

By Network-aware control we mean that the control system will take measures to optimize the utilization of resources for control information exchange among sensors, controllers and actuators via shared communication network which has the constraints like Delay, Jitter, Packet Loss, Out-of-Sequence packet etc. Modify control algorithms to cope with communication imperfections — e.g., control with time-stamped sensor data. Communication imperfections in single control loop include — Delay and jitter, Bandwidth limitations, Data loss and bit errors, Outages and disconnection. To realize a network-aware controller we need to modify the conventional controller to cope with communication imperfections in the following ways :

1. Control under varying network delay
2. Control under data loss
3. Control under bandwidth limitation
4. Control under topology constraints

Network-aware control architecture is needed to be designed and analyzed which will look like Fig.(3.2) which is basically an architecture for control over communication networks. Here, we have to estimate network state in presence of network delay, data loss probability and bandwidth.

3.3 Quantized Control Systems

The literal definition of quantization is the division of a quantity into a discrete number of small parts, often assumed to be integral multiples of a common quantity. The oldest example of quantization is rounding off. Any real number x can be rounded off to the nearest integer $q(x)$, say, with a resulting quantization error $e = q(x) - x$ so that $q(x) = x + e$. More generally, we can define a quantizer as consisting of a set of intervals or *cells* $\mathcal{S} = \{S_i; i \in \mathcal{I}\}$, where the index set $\mathcal{I}$ is ordinarily a collection of consecutive integers beginning with 0 or 1, together with a set of reproduction values or points or levels $\mathcal{C} = \{y_i; i \in \mathcal{I}\}$, so that the overall quantizer is defined by $q(x) = y_i$ for $x \in S_i$, which can be expressed [Gray and Neuhoff, 1998] concisely as

$$q(x) = \sum_i y_i l_{S_i}(x).$$

where the indicator function $l_S(x)$ is 1 if $x \in S$ and zero otherwise.
Due to quantization technique used in the communication channel used for networked control system there exists quantization errors which lead to control performance degradation. A quantized measurement $q(x)$ of a real number x as input can be treated like a partial observation of x or as an entity containing a limited quantity of information about x [Delchamps, 1990]. Thus this explanation seems to be more qualitative than quantitative. Due to finite bandwidth we have to restrict the number of bits assigned to the quantization which has got a direct bearing on the quantization accuracy.
Information-theoretic source coding technique reduces the number of bits to represent x based on entropy. Entropy puts the lowest limit on the number of bits assigned to represent x so as to successfully reconstruct the original value of x within specified tolerable error probability or bit-error rate which is called the *fidelity*. The quality of a quantizer can be measured by the *goodness* (*fidelity*) of the resulting reproduction in comparison to the original. One way of accomplishing this is to define a distortion measure $d(x, \hat{x})$ that quantifies cost (performance cost) or distortion resulting from reproducing x as $\hat{x}$ and to consider the average distortion as a measure of the quality of a system, with smaller average distortion meaning higher quality. The most common distortion measure is the squared error $d(x, \hat{x}) = \|x - \hat{x}\|^2$. In practice, the average will be a sample average when the quantizer is applied to a sequence of real data, but the theory views the data as sharing a common probability density function (pdf) $f(x)$ corresponding to a generic random variable X and the average distortion $D(q)$ [Gray and Neuhoff, 1998] becomes an expectation as follows:

$$D(q) = E\left[d\left\{X, q(X)\right\}\right] = \sum_i \int_{S_i} d(x, y_i) f(x) dx.$$

If the distortion is measured by squared error, $D(q)$ becomes the mean squared error (MSE).
It is desirable to have the average distortion as small as possible, and, in fact, negligible average distortion is achievable by letting the cells become numerous and tiny (i.e.

denser quantization). There is a cost in terms of the number of bits required to describe the quantizer output to a decoder, however, and arbitrarily reliable reproduction will not be possible for digital storage and communication media with finite capacity. A simple method for quantifying the cost for communications or storage is to assume that the quantizer "codes" an input x into a binary representation or channel codeword of the quantizer index i specifying which reproduction level should be used in the reconstruction. If there are N possible levels and all of the binary representations or binary codewords have equal length (i.e. *uniform quantization*), the binary vectors will need $\log N$ (or the next larger integer, $\lceil \log N \rceil$, if $\log N$ is not an integer) bits. Thus one definition of the rate of the code in bits per input sample is $R(q) = \log N$. Here, the base of the logarithm is 2 as we are concerned here with bits or binary representations.
A quantizer with fixed-length binary codewords is said to have *fixed rate* because all quantizer levels are assumed to have binary codewords of equal length.

In summary, the goal of quantization is to encode the data from a source, characterized by its probability density function, into as few bits as possible (i.e., with low rate) in such a way that a reproduction may be recovered from the bits with as high quality as possible (i.e., with small average distortion). Clearly, there is a trade-off between the two primary performance measures: average distortion (or simply distortion, as we will often abbreviate) and rate. This trade-off may be quantified as the operational *Distortion-Rate function* $\delta(R)$, which is defined to be the least distortion of any scalar quantizer with rate R or less. Mathematically,

$$\delta(R) \equiv \inf_{q:R(q)\leq R} D(q)$$

Alternatively, one can define the operational *Rate-Distortion function* $R(D)$ as the least rate of any fixed-rate scalar quantizer with distortion D or less, which is the inverse of $\delta(R)$.

The measurements of the plant and/or control inputs have to go through finite-capacity digitalized communication channels and hence are quantized to finite precision; in addition, in many situations, these signals are transmitted and processed only intermittently (namely, they are transmitted and processed only if a change occurs). Therefore, quantized controller usually generates piecewise constant control inputs and are event-driven systems. In [Brockett, 1997] the author proposed the minimum attention control which uses piecewise constant control and takes into account of the "attention cost" measuring the overhead for information transmitting and processing.
Coarser quantization implies that less information flows between the controller and the plant. Therefore, the minimum quantization density that stabilizes an unstable plant is of interest. It can be used to measure the minimum information needed for stabilization, and it codifies how difficult a system can be controlled. In [Elia and Mitter, 2001] the authors devised the quantizer with minimum density for stabilizing a discrete-time linear time-invariant (LTI) single-input plant that is open-loop unstable. The quantizer design was shown to be an Minimum Energy Control (MEC) problem, and the minimum density of the quantizer ρ is given by

$\rho = \frac{DI-1}{DI+1}$,

depending only on the *Degree of Instability*

$$DI := \prod_{i}^{m} |\lambda_{u,i}|$$

where $\lambda_{u,i}$ are the unstable poles of the system. This implies that if the plant is more unstable, then more information is needed for accomplishing the stabilization task. Later, the minimum density quantizer design was obtained for multi-input plants [Elia, 2002], for nonlinear plants [Liu and Elia, 2004], and for control with performance requirements [Elia, 2000].
In [Tatikonda, 2000] and [Nair and Evans, 2000] is provided the minimum bit-rate for stabilizing an unstable linear plant. The minimum bit-rate R depends only on the degree of instability, namely,

$$R = \log DI$$

Similar results hold for state estimation. Since the bit-rate is directly linked to mutual information and entropy, one may attempt to link this result to the (generalized Kolmogorov) entropy generated by the plant. Researchers showed that such an unstable plant generates entropy at a rate equal to $\log DI$, and hence apparently a channel which can sustain communication rate of at least $\log DI$ is needed [Nair et al., 2004].

Information quantization makes a dynamical system hybrid in many cases. Such a hybrid nature may cause technical difficulties (such as the discontinuity in the vector fields for continuous-time dynamical systems, refer [Brockett and Liberzon, 2000]). Despite this, sometimes people intensionally introduce information quantization to a system to reduce the communication/computation costs and to address the design problems of hybrid systems or hierarchical systems [Brockett, 1997], [Elia, 2000], [Elia and Mitter, 2001], [Egerstedt and Hu, 2002].

Trade-off between quantization accuracy and sampling frequency are illustrated with examples and simulation results in papers [Ben Gaid and Çela, 2006] and [Zhang and Liu, 2008].

3.4 Conclusion

In this chapter, we have briefly discussed the control system (e.g. NECS) operating under communication constraints. We have discussed the two major approaches (Network-aware Control and Control-Aware Networking) of solving the constrained control problems. We have reviewed the theoretical results established in the literatures in the context of quantized control under communication constraints. To summarize, quantized control systems become an interesting device that integrates dynamical systems and

control, information theory and communication and hybrid systems.

4

Information-theoretic Analysis of Quantized Control

4.1 Introduction

In this chapter, we present an information-theoretic view and analysis of *Quantized Control* frequently encountered in a distributed embedded control system especially under communication constraints.
Because of the feedback nature of control and the importance of the information pattern it is important to distinguish sharply between two aspects of the word "control".
The first aspect is the control law which maps available data into values of control inputs to the system. The control law is a function which is selected from a set of functions (infinite dimensional!), and there is nothing stochastic or uncertain about this function. The problem is to make an optimal selection. This is to be done by the designer before the system is put together.
The second aspect is the realization of the control variables actually applied to the system in operation.

4.2 Information and Entropy

Ideally, what we would like is for the message to have low entropy by having prior information so that there is less information for the decoder to try and extract from the signal. The intuition involved is that low entropy implies better predictability. Better predictability means that our prior knowledge is quite strong.

However, to get the message across intact, we would like the messenger to have high-energy so that the signal-to-noise ratio is favourable (high mutual information = informative likelihoods). The intuition for the case of signaling is that we want to reduce the effect of the noise. We do this by having a large mutual information between the input and output of the channel. Unfortunately, when we restrict ourselves to affine controllers for this problem, these two objectives are in direct opposition. An affine controller implies Gaussian state and for a Gaussian random variable, high energy implies high entropy and low entropy implies low energy.
There is one important difference between the continuous and discrete entropies. In the discrete case, the entropy measures in an absolute way the randomness of the random variable. In the continuous case, the measurement is relative to the coordinate system. If we change coordinates the entropy will in general change. Since the mutual information is the difference between the two entropies, it is co-ordinate independent due to common terms getting canceled by subtraction; whereas the entropy is co-ordinate dependent.
In spite of this dependence on the coordinate system the entropy concept is as important in the continuous case as the discrete case. This is due to the fact that the derived concepts of information rate and communication channel capacity depend on the difference of two entropies and this difference does not depend on the coordinate frame, each of the two terms being changed by the same amount. Hence, we would like to use mutual information related to entropy in defining and correlating some important control parameters of the system.

4.2.1 Relation between Differential Entropy and Discrete Entropy

Due to quantization the quantized output are discrete in nature (because quantization results in discretization) and hence we need to investigate how these quantized values are related to entropy. Here, we analyze the quantization in the context of information-theoretic entropy.
Based on [Cover and Thomas, 2006] the relation is established as follows :
Consider a random variable X with density $f(x)$. Suppose that we divide the range of X into bins of length Δ. Here, bin is the interval between two consecutive values of $f(x)$ corresponding to the two consecutive values of x.
Now, for multidimensional continuous case, entropy (precisely *differential entropy*) of a continuous random variable X with probability density function $f(x)$
(if $\int_{-\infty}^{\infty} f(x)dx = 1$) is defined as

Differential Entropy $h(X) = -\int_S f(x) \ln f(x)dx$;

where the set S for which $f(x) > 0$ is called the *support set* of X.

As in discrete case, the differential entropy depends only on the probability density

of the random variable and therefore the differential entropy is sometimes written as $h(f)$ rather than $h(X)$.
Let us assume that the density is continuous within the bins. Then by mean value theorem, there exists a value x_i within each bin such that

$f(x_i)\Delta = \int_{i\Delta}^{(i+1)\Delta} f(x)dx.$

Now, quantized random variable X^Δ is defined as

$X^\Delta = x_i$, if $i\Delta \leq X < (i+1)\Delta$.

Then the probability that $X^\Delta = x_i$ is given by

$p_i = \int_{i\Delta}^{(i+1)\Delta} f(x)dx = f(x_i)\Delta.$

Therefore, the Entropy of the Quantized Random Variable is

$H(X^\Delta) = -\sum_{-\infty}^{\infty} p_i \log p_i$

$= -\sum_{-\infty}^{\infty} f(x_i)\Delta \log(f(x_i)\Delta)$

$= -\sum \Delta f(x_i) \log f(x_i) - \sum f(x_i)\Delta \log \Delta$

$= -\sum \Delta f(x_i) \log f(x_i) - \log \Delta$. (since, $\sum f(x_i)\Delta = \int f(x)dx = 1$.)

$\therefore H(X^\Delta) = -\sum \Delta f(x_i) \log f(x_i) - \log \Delta.$

If $f(x)\log f(x)$ is Riemann integrable (to ensure that the limit is well defined), the first term in the above equation approaches the integral of $-f(x)\log f(x)$ as $\Delta \to 0$ by definition of Riemann Integrability. Using these facts and arguments as mentioned following theorem can be written :

Theorem 4.1. *[Cover and Thomas, 2006] If the density $f(x)$ of the random variable X is Riemann integrable, then*

$H(X^\Delta) + \log \Delta \to h(f) = h(X)$, *as* $\Delta \to 0$.

Thus, the entropy of an n-bit quantization of a continuous random variable X is approximately $h(X) + n$.

Generally, $h(X) + n$ is the number of bits required on an average to describe X to n-bit accuracy.

We note the following points [Cover and Thomas, 2006]:

- Given n bits to describe a random variable X with differential entropy $h(X)$, the error can have differential entropy no less than $h(X) - n$.
- Given a random variable X with differential entropy $h(X)$, the lowest possible variance of X is $\frac{1}{2\pi e}2^{2h(X)}$.
- The *Entropy-Power Inequality* : Given two independent random variables X and Y with differential entropy $h(X)$ and $h(Y)$ respectively,

  $2^{2h(X+Y)} \geq 2^{2h(X)} + 2^{2h(Y)}$.
- Entropy of a random variable X is no less than the entropy of X given additional information about another random variable Y. i.e. $h(X) \geq h(X|Y)$.

Proof: **Variance-Entropy Relation**

Let X be the random variable with density function $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-x^2}{2\sigma^2}}$ with normal distribution. Then, differential entropy can be calculated as follows :

$h(f) = -\int f \ln f$

$= -\int f(x)\left[\frac{-x^2}{2\sigma^2} - \ln\sqrt{2\pi\sigma^2}\right]$

$= \frac{EX^2}{2\sigma^2} + \frac{1}{2}\ln\{2\pi\sigma^2\}$

$= \frac{1}{2} + \frac{1}{2}\ln\{2\pi\sigma^2\}$

$= \frac{1}{2}\ln e + \frac{1}{2}\ln\{2\pi\sigma^2\}$

$= \frac{1}{2}\ln\{2\pi e\sigma^2\}$

Therefore, $h(f) = \frac{1}{2}\ln\{2\pi e\sigma^2\}$

$\Rightarrow$ The Variance $\sigma^2 = \frac{1}{2\pi e}e^{2h(X)}nats$

Changing the base of the logarithm from natural to binary we have

The Variance $\sigma^2 = \frac{1}{2\pi e}2^{2h(X)}bits$

Hence, the variance-entropy relation is established.

Using *Asymptotic Equipartition Property (AEP)* [Cover and Thomas, 2006] of Information Theory we can proceed as follows :

Let $X_1, X_2, \ldots, X_n$ be a sequence of random variables drawn i.i.d according to the density $f(x)$. Then

$-\frac{1}{n}\log f(X_1, X_2, \ldots, X_n) \to E\left[-\log f(x)\right] = h(X)$ in probability.

Definition 4.1. For $\epsilon > 0$ and any n, we define the *typical set* $A_\epsilon^{(n)}$ w.r.t. $f(x)$ as follows :

$A_\epsilon^{(n)} = \left\{(x_1, x_2, \ldots, x_n) \in S^n : \left|-\frac{1}{n}\log f(x_1, x_2, \ldots, x_n) - h(X)\right| \leq \epsilon\right\}$

where $f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n f(x_i)$.

The analog of cardinality of the typical set for the discrete case is the volume of the typical set for continuous random variables.

Definition 4.2. The *volume* $Vol(A)$ of a set $A \subset \Re^n$ is defined as $Vol(A) = \int_A dx_1 dx_2 \cdots dx_n$.

The typical set $A_\epsilon^{(n)}$ has the following property :

$Vol\left(A_\epsilon^{(n)}\right) \leq 2^{-n(h(X)+\epsilon)}$ for all n.

Proof. Now, $1 = \int_{S^n} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \cdots dx_n$.

$\geq \int_{A_\epsilon^{(n)}} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \cdots dx_n$.

$\geq \int_{A_\epsilon^{(n)}} 2^{-n(h(X)+\epsilon)} dx_1 dx_2 \cdots dx_n$.

$= 2^{-n(h(X)+\epsilon)} \int_{A_\epsilon^{(n)}} dx_1 dx_2 \cdots dx_n$.

$= 2^{-n(h(X)+\epsilon)} Vol\left(A_\epsilon^{(n)}\right)$.

Therefore, $2^{-n(h(X)+\epsilon)} Vol\left(A_\epsilon^{(n)}\right) = 1$.

$\Rightarrow Vol\left(A_\epsilon^{(n)}\right) = \frac{1}{2^{-n(h(X)+\epsilon)}}$

Hence, Volume of the smallest set with $\epsilon \to 0$ would be given by

$Vol\left(A_\epsilon^{(n)}\right) = \frac{1}{2^{-n(h(X)+\epsilon)}} = 2^{nh(X)}$. □

Volume of the smallest set that contains most of the probability is approximately $2^{nh(X)}$, where n refers to the n-dimensional volume.

Hence, **Side length** $l = \left(2^{nh(X)}\right)^{\frac{1}{n}} = 2^{h(X)}$

$\Rightarrow h(X) = \log_2 l$

Therefore, the differential entropy $h(X)$ is the logarithmic value of the equivalent side length of the smallest set that contains most of the probability. Hence, low entropy implies that the random variable (r.v.) is confined to a small effective volume and high entropy indicates that the r.v. is widely dispersed. Changing the size of the quantization blocks, one can extract more information about the behaviour of the system.

4.2.2 Entropy and Estimation Error Relation

Referring to the paper [Gupta et al., 2006], for an LTI Discrete system represented by

$x_{k+1} = ax_k + u_k + w_k$, $u_k = f(\hat{x}_k)$

where u_k is the control input, x_k being the state, $\hat{x}_k$ being the estimated state, w_k being the white noise.
At time step $k = 0$, the entropy is simply $h(x_0)$, thus the entropy of Δ_0 is at least $h(x_0) - n$.
At time step $k = 1$, we have

$h(x_1) \geq h(x_1|\hat{x}_0) = h(ax_0 + w_0|\hat{x}_0)$.

Now x_0 and w_0 are independent (even when $\hat{x}_0$ is given). Denoting the entropy of the noise by $h(w)$ we can write

$2^{2h(x_1)} \geq 2^{2h(ax_0|\hat{x}_0)} + 2^{2h(w|\hat{x}_0)}$

$= 2^{2log(a)+2h(x_0|\hat{x}_0)} + 2^{2h(w)}$

$\geq 2^{2log(a)}2^{2h(x_0)-2n} + 2^{2h(w)}$.

Putting $\alpha = a^2 2^{-2n}$ we get

$h(x_1) \geq \frac{1}{2}\log\left[\alpha 2^{2h(x_0)} + 2^{2h(w)}\right]$

$h(\Delta_1) \geq \frac{1}{2}\log\left[\alpha 2^{2h(x_0)} + 2^{2h(w)}\right] - n$.

Similarly we can get

$$h(\Delta_k) \geq \frac{1}{2}\log\left[\alpha^k 2^{2h(x_0)} + \sum_{j=0}^{k-1}\alpha^j 2^{2h(w)}\right] - n$$

Finally the error variance at time step k is : $E\left[\Delta_k^2\right] \geq \frac{1}{2\pi e} 2^{-2n} \left[\alpha^k 2^{2h(x_0)} + \sum_{j=0}^{k-1} \alpha^j 2^{2h(w)}\right]$

Thus, we can conclude that the Minimum Mean Square Error (MMSE) or error variance is lower bounded by two terms — (1) the term containing Initial State, (2) the term containing the Noise alone. Here, we correlate the estimation-theoretic parameter MMSE with information-theoretic entropy-powers of control-theoretic initial state and scalar a.

Given the quantized output Y_q as side information (side information means the information indirectly used for finding the state) and estimator $\hat{X}(Y_q)$, the Mean-square-error is given by

$$E\left(X - \hat{X}(Y_q)\right)^2 \geq \frac{1}{2\pi e} e^{2h(X|Y_q)}.$$

Shannon's lossless source coding theorem implies that given a source and a quantizer partition, one can always find an assignment of binary codewords (indeed, a prefix set) with average length not more than $H(q(X)) + 1$, and that no uniquely decodable set of binary codewords can have average length less than $H(q(X))$, where

$$H(q(X)) = -\sum_i p_i \log_2 p_i$$

is the Shannon Entropy of the quantizer output, p_i being the probability that the source sample $X \in S_i$ lies in the i-th cell S_i.
[Renyi, 1959] shows that a uniform scalar quantizer with infinite number of levels and small cell width Δ has output entropy given approximately by

$$H(q(X)) \approx h(X) - \log_2 \Delta$$

where $h(X)$ is the differential entropy of the source variable X as defined earlier.

$$\Rightarrow \log_2 \Delta = h(X) - H(q(X))$$

Therefore,

$$\Delta = 2^{h(X) - H(q(X))} \tag{4.1}$$

We can define an operational distortion-rate function as the minimum average distortion over all variable-rate quantizers with output entropy $H(q(X)) \leq R$. Since the quantizer output entropy is a lower bound to actual rate, this operational distortion-rate function may be optimistic; i.e., it falls below $\delta(R)$ defined using average length as rate. A quantizer designed to provide the smallest average distortion subject to an entropy constraint is called an entropy-constrained scalar quantizer. Based on [Zador, 1966] operational distortion-rate function (i.e., the least distortion of such codes with a given rate) is

$$\delta(R) \approx \frac{\Delta^2}{12} \tag{4.2}$$

Using (4.1) in (4.2) results in $\delta(R) \approx \frac{1}{12}2^{2h(X)-2H(q(X))} \approx \frac{1}{12}2^{2h(X)-2R}$

Therefore, $\delta(R) \approx \frac{1}{12}2^{2\{h(X)-R\}}$

Hence, we conclude that higher the code rate in terms of R in bits/sample, lower the distortion i.e. lower the quantization noise which implies better quantization precision or accuracy. Also, more uncertainty leads to more distortion.

4.3 On the Value of Information

Referring to paper [Witsenhausen, 1971] : "the minimum (or infimum) expected cost achievable in a problem depends upon the prevailing information pattern. Changes in information produce changes in optimal cost. This suggests the idea of measuring information by its effect upon the optimal cost, as has been proposed many times".
Such a measure of information is entirely dependent upon the problem at hand and is clearly not additive. The only general property that it is known to possess is that additional information, if available free of charge, can never do harm though it may well be useless. This simple monotonicity property is in sharp contrast with the elaborate results of information transmission theory. The latter deals with an essentially simpler problem, because the transmission of the information is considered independently of its use, long periods of use of a transmission channel are assumed, and delays are ignored.
Refer to the Fig. (4.1) as follows :

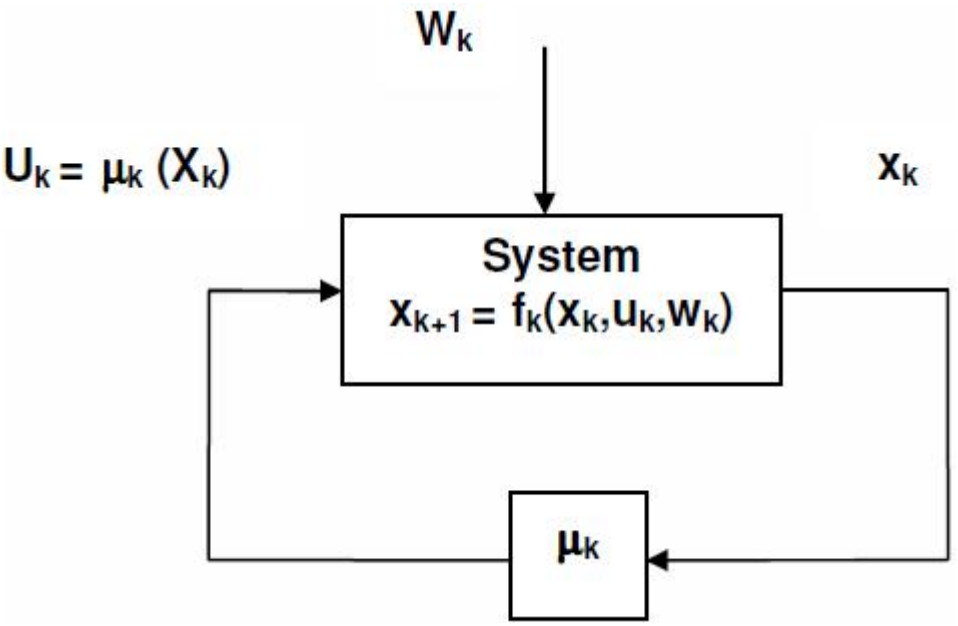


Figure 4.1: Information Gathering

Consider a discrete-time dynamic system :

$x_{k+1} = f_k(x_k, u_k, w_k)$ for $k = 0, 1, \ldots, N-1$;

where k is the time-instant, the state $x_k \in X$, the control $u_k \in U$ and the random disturbance $w_k \in W_k$.
The distinction between open-loop minimization of the cost, where we select all controls $u_0, u_1, \ldots, u_{N-1}$ at once at time 0, and closed-loop minimization, where we select a policy $\{\mu_0, \mu_1, \ldots, \mu_{N-1}\}$ that applies the control $\mu_k(x_k)$ at time k with the knowledge of the current state x_k. With closed-loop policies, it is possible to achieve lower cost, essentially by taking advantage of the extra information (the value of the current state). The reduction in cost may be called the *value of the information* and can be significant indeed. If the information is not available, the controller cannot adapt appropriately to unexpected values of the state, and as a result the cost can be adversely affected. As suggested by Bertsekas [Bertsekas, 2000] no open-loop strategy can give a chess player a greater than 50-50 chance of winning the match, whereas with closed-loop strategy the player has more than 50% chance of winning. It is further emphasized [Bertsekas, 2000] that the value of information (the outcome of an event) is the difference of the optimal closed-loop and open-loop probability (of favourable outcome) values.

4.4 Extracting State Information from Quantized Output

Quantized measurements [Delchamps, 1989] of real numbers are in fact partial informations rather than mere approximations (as perceived by many researchers). Here, partial informations mean not the full information we are getting due to quantization as some informations are lost in the process of quantization. Our aim should be to extract valuable or useful information from these partially observed information using the concepts of entropy from information theory. Conventionally, the quantized informations are treated as approximations when the quantization losses are ignored but these quantization losses are used here as useful information in extracting state informations. Quantized measurement $q(x)$ is not just an approximation but an entity that provides us a limited amount of information about state x. Considering a system :

$$\begin{aligned} x(k+1) &= ax(k) + bu(k), \qquad k \geq 0\,; \\ y(k) &= q(x(k)), \end{aligned}$$

where a and b are real numbers and q is a mapping from $\Re$ into a finite set $\mathbf{Y} = \{y_i, i \in \mathbf{Z}\}$. An admissible feedback control strategy may be therefore modeled as a family of mappings $\left\{f^{(k)} : \mathbf{Y}^k \rightarrow \Re\right\}$ which satisfies

$u(k) = f^{(k)}\left\{y(0), \ldots, y(k)\right\}, k \geq 0.$

Let us now define the uniform quantizer q with sensitivity Δ as follows (with $\Delta > 0$) :

$$q_\Delta(x) = \begin{cases} i & \text{if } x \geq 0, \\ & \text{and } x \in \left[\left(i - \frac{1}{2}\right)\Delta, \left(i + \frac{1}{2}\right)\Delta\right], \\ -q_\Delta(-x) & \text{when } x < 0\ . \end{cases}$$

We model the uncertainty about $x(k)$ given the observation set $\{y(m) : 0 \leq m \leq k\}$ as the differential entropy of $x(k)$ given $\{h(m) : 0 \leq m \leq k\}$. We denote this quantity (uncertainty) by $h(x(k)|y_0^k)$ and defined as follows :

$$\begin{aligned} h(x(k)|y_0^k) &= -\sum_{(y_0,\ldots,y_k)\in \mathbf{Y}^k} Probab\left\{y(0) = y_0, \ldots, y(k) = y_k\right\} \\ &= \left[\int f(x(k)|y_0^k)\log(f(x(k)|y_0^k))dx(k)\right] \end{aligned}$$

where, $f(x(k)|y_0^k)$ is the conditional density of $x(k)$ given $y(0) = y_0, \ldots, y(k) = y_k$. The lower the value of $h(x(k)|y_0^k)$, the more information about $x(k)$ is contained on an average in the output set till time k. Observing

$h(x(0)|y(0)) = \log \Delta$;

In general, $h(x(k)|y(k)) = \log \Delta$.

Therefore, $\Delta = 2^{h(x(k)|y(k))}$.

4.5 Information-Theoretic Rate Requirement under Asymptotic Quantization

Consider an autonomous scalar linear system
$x_{t+1} = ax_t$,
where x_t is the state at time $t = 0, 1, \ldots, T$ and the pdf of initial state has finite support. Here we assume the distribution to be Gaussian as it gives the maximum entropy (worst-case) with unbiasness. Then the following hold [Yüksel and Başar, 2006]:

1. The bit-rate required for bounding the state estimator error variance as the terminal time $T \to \infty$, is $\max\{0, \log_2 |a|\}$ per time-stage.
2. This rate is achievable by quantization. Also, the uniform quantization is optimal.

Proof:

As per the theorem [Cover and Thomas, 2006] :

$H(aX) = H(X) + \log|a|$

$\therefore H(ax_0) = H(x_0) + \log_2|a|$

Similarly, $H(a^2x_0) = H(x_0) + \log_2|a^2| = H(x_0) + 2\log_2|a|$
$\vdots$
$\therefore H(a^tx_0) = H(x_0) + t\log_2|a|.$

Hence, we note the following :
1. The entropy at time t is

$$H(x_t) = H(a^tx_0) = H(x_0) + t\log_2|a|.$$

and mutual information between the transmitted state x_t and the received estimated state $\hat{x}_t$ is

$$I(x_t;\hat{x}_t) = H(x_t) - H(x_t|\hat{x}_t) = H(x_t) - H(x_t - \hat{x}_t|\hat{x}_t) \tag{4.3}$$

4.5.1 Justification for the Equation (4.3) of Mutual Information

Intuitively, it says that the uncertainty in the state does not change as a result of a deterministic shift in the random variable, since $\hat{x}$ is a deterministic function of itself.

$P(x|\hat{x})$ is a distribution, with some mean.

$P(x - \hat{x}|\hat{x})$ is another probability distribution, with a shifted mean.

They both have the same entropy, since in the definition of differential entropy (assuming there exists a density which is Gaussian, without any loss here) there is no change with translation :

$$H(x|\hat{x}) = -\int P(x = z|\hat{x})log(P(x = z|\hat{x}))dz$$

$$= -\int P(x - \hat{x} = z - \hat{x}|\hat{x})log(P(x - \hat{x} = z - \hat{x}|\hat{x}))dz$$

$$= -\int P(x - \hat{x} = z'|\hat{x})log(P(x - \hat{x} = z'|\hat{x}))dz'$$

where in the last equality, we shift the integral boundaries.

$\therefore H(x|\hat{x}) = H(x - \hat{x}|\hat{x})$

i.e. $H(x_t|\hat{x}_t) = H(x_t - \hat{x}_t|\hat{x}_t)$

Hence, the equation (4.3) is justified.

Let, $\mathcal{P}_\mathcal{D}$ be the set of all probabilistic maps that achieve a given distortion level D. Then, for corresponding rate R we have $R \geq \inf_{\mathcal{P}_\mathcal{D}} I(x_t; \hat{x}_t)$. Thus,

$R \geq \inf_{\mathcal{P}_\mathcal{D}} \{H(x_t) - H(x_t - \hat{x}_t|\hat{x}_t)\}$

$= \inf_{\mathcal{P}_\mathcal{D}} \{H(x_0) + t \log_2 |a| - H(x_t - \hat{x}_t|\hat{x}_t)\}$

$\geq \inf_{\mathcal{P}_\mathcal{D}} \{H(x_0) + t \log_2 |a| - H(x_t - \hat{x}_t)\}$; because conditioning reduces entropy.

$$R \geq H(x_0) + t \log_2 |a| - \frac{1}{2} \log(2\pi e D) \tag{4.4}$$

because maximum differential entropy for Gaussian distribution being $\frac{1}{2} \log(2\pi e D)$ [Cover and Thomas, 2006]. Here, the random variable is having finite variance D.

It may be noted that we are interested in quantizers, the quantization error conditioned on the quantized values will, in fact, have a compact support set for its distribution. Gaussian density entropy provides here the lower bound on the bit-rate.
On dividing both sides of (4.4) by time t and taking limits as $t \to \infty$ we get the necessary condition on the time-average rate $R_{av} = \frac{R}{t}$ which is lower bounded by $\log_2 |a|$.

2. Although the rate per stage of time is finite, as the terminal time goes to ∞, we can regard this problem as the quantization where the number of quantization levels goes to ∞. Paper [Gish and Pierce, 1968] showed that the uniform quantizer followed by entropy coder is at most 0.255 bit worse than the optimal quantizer. Since, $\frac{0.255}{T} \to 0$, as $T \to \infty$ where T being the terminal time, we can conclude that uniform quantizer is optimal.

4.6 Certainty Equivalence Principle

The basic idea of this is to define a decision policy by putting the uncertainty equal to its expected value given the observations. If the optimal policy is unaffected when the disturbances are replaced by their means, we say that *certainty equivalence principle* holds. The *certainty equivalence controller* [Bertsekas, 2000] is a suboptimal control scheme that is inspired by linear-quadratic control theory. It applies at each stage of the control that would be optimal if some or all of the uncertain quantities were fixed at some typical values; i.e. it acts as if a form of the certainty equivalence principle were holding. According to [Åström and Wittenmark, 1997] optimal control strategy can be separated

into two parts

1. State Estimator : gives the best estimate of the states from the observed outputs.
2. Linear Feedback Law from the estimated states : linear controller used is same as the one used if there are no disturbances acting on the system.

Definition : *Certainty Equivalence Control Law* [Water and Willems, 1981]:
Assume that we face the decision of choosing a control $u \in U$ so as to minimize the cost function $J(u, w)$, where w, u, y being the uncertainty, control input and observed output respectively. In the case of perfect information, i.e., if $w \in \Omega$ (where Ω is the Uncertainty Space) is known, this leads to the decision $u^*(w)$ chosen, such that $J(u^*(w), w) \leq J(u, w)$ for all $u \in U$. However, if we only know w only through the observations $y(w)$ and we cannot solve the problem of finding the corresponding optimal feedback control law $u^*(y)$, it is not totally without merit to use the control law given by
$\hat{u}(y) := E\{u^*(w)|y(w) = y\}$.
Here we have assumed that this conditional expectation is well defined. This idea is called certainty equivalence and can be formalized in a dynamic framework as well.

4.7 Quantization and Mean Square Error

Let us consider a uniform scalar quantizer with step size Δ. An N-level uniform quantizer q whose levels are $y_1, \ldots, y_N$ with $y_i = y_{i-1} + \Delta$. If this quantizer is applied to a r.v. X with pdf $f(x)$ when Δ is small and overload distortion (Overload distortion is a result of too small step size of quantization, to avoid it one needs to increase the step size) is ignored, the Mean Square Error (MSE) distortion can be approximated as follows :

$$D(q) = E\left[(X - q(X))^2\right] \cong \frac{\Delta^2}{12}$$

4.8 Conclusion

In this chapter, we have presented the informational aspect of quantized control which is pertinent in a distributed embedded control system operating under communication constraints. We have analyzed the various theoretical results in the context of quantized control in the light of information theory. Relations between entropy and estimation error, entropy and variance are discussed. Also, the extraction of state information from quantized output is presented in this chapter.

5

Convergence of Information Theory, Communications Theory, Estimation Theory and Control Theory

5.1 Introduction

This chapter is based on our survey work vide our paper [Roy and Çela, 2007].
It has been suggested [Mitter, 2000] many times that a unified view of Control and Communication is badly needed if progress has to be made towards a science of distributed systems, where subsystems are linked via Communication channel. Control engineers must master computer and software technologies to be able to build the systems of the future, and software engineers need to use control concepts to master the ever-increasing complexity of real-time computing systems. The results surveyed address the real-time information exchange limitations in terms of packet rates, sampling, network delay and packet dropouts along with the problems faced by control community and also the possible solutions in the light of convergence of Control Theory, Information Theory and Communications Theory along with Estimation Theory.

Unlike regular control systems, in networked control systems the synchronization between different sensors, actuators and control units is not guaranteed. Furthermore, there is no guarantee for zero delay or even constant delay in sending information from sensors to the control units and control signals from the control units to the actuators. The paper [Årzén et al., 2006], highlights how the complexity of the closed-loop system is increased, due to additional disturbances introduced by the communication system : additional delays, jitter, data rate limitations, packet losses etc.

Communication is an important component of distributed and networked control systems. Hence, there is a need to understand the interactions between the control components and the communication components of the distributed system. A control problem is formulated with a communication channel connecting the sensor to the controller and designed the channel encoder and channel decoder along with the controller to achieve different control objectives. Specifically it is shown by researchers that a necessary condition on the rate for asymptotic observability and asymptotic stabilizability in a linear, discrete-time system is (recalling Eq. (1.1) of Chapter - 1)

$$R \geq \sum_{\lambda(\mathbf{A})} \max\{0, \log|\lambda(\mathbf{A})|\}$$

where $\lambda(\mathbf{A})$ denotes the eigenvalue of the system matrix $\mathbf{A}$ and R denotes the bit rate without error.
This result relates the speed of the dynamics of the plant to the information rate of the channel. Thus, control theoretic parameter are connected to communication parameters. Paper [Liu and Elia, 2006] showed integration of Information, Estimation and Control, and unified fundamental limitations of information, estimation and control. It is shown that the achievable information rate in the feedback communication problem can be alternatively given by the decay rate of the *Cramer-Rao Bound (CRB)* in the associated estimation problem, or by the Bode sensitivity integral in the associated control problem. Here, we may recall the CRB as the lower limit of variance $\sigma^2 \geq \frac{1}{FIM}$ from Chapter - 2 of the thesis.

Therefore, it is concluded that the three fundamental limitations namingly : *Achievable Information Rate* in information transmission (Communication), *CRB* in information processing (Estimation), and *Bode Sensitivity Integral* in information utilization (Control), seemingly different and usually separately treated, are in fact three sides of the same entity.

There is a general equivalence among a feedback communication system over a Gaussian channel, an estimation system over the same channel, and a control system over the same channel. Roughly speaking, in steady state, the communication system achieves reliable communication if and only if the estimation system has bounded estimation error and if and only if the control system is stabilized in closed-loop.

5.2 Convergence of Information, Communication, Estimation and Control

Referring to [Liu and Elia, 2006] we find the convergence as follows :
For a SISO AWGN Channel with *ISI* (*Inter-Symbol Interference*, a common communication problem caused by improper sampling interval and bandwidth resulting in signal distortion and hence errors generated in reception), denoted as $\mathcal{F}$ and described in state-

space as

$$\mathcal{F}: \begin{cases} s_{t+1} = Fs_t + Gu_t, s_0 = 0, \\ y_t = Hs_t + u_t + N_t; \end{cases}$$

Where $F \in \Re^m$, m being the order of $\mathcal{F}$, u_t is the channel input, s_t is the channel state, N_t is the AWGN and y_t is the channel output. We assume that (F, G) is controllable and (F, H) is observable. Then using the results of [Liu and Elia, 2006] and CRB concept we may conclude that

$$\text{Mutual Information} I(u;y) = \frac{1}{2}\log(\frac{1}{MMSE}) = -\frac{1}{2}\log(CRB) = \frac{1}{2}\log(FIM)$$

The Feedback Communication Limitation is the Achievable Rate; Estimation Limitation is the CRB Decay-Rate and the Control System Fundamental Limitation is the Bode Sensitivity Integral. It is pertinent to find that

Achievable Rate in Feedback Communication System

$= \frac{1}{2}$ *Decay-rate of CRB in Estimation System*

$=$ *Bode Sensitivity Integral in Control System*

There exists a critical value for the signaling rate, above which reliable communication is not achievable and below which reliable communication is achievable. In the (recursive) estimation system, the fundamental limitation is the decay rate of Cramer-Rao bounds (CRB): whatever estimator one may design, the decay rate of mean-squared error (MSE) cannot be made faster than the decay rate of CRB. In the control system, the fundamental limitation is the Bode sensitivity integral. No matter how one designs the controller, the sensitivity integral cannot be made smaller than a constant determined by how unstable the plant is. Hence, the fundamental limitations in the three systems agree. Thus, it is said that the three problems : the optimal estimation problem, the optimal feedback communication problem and the minimum-energy control problem, are equivalent and hence if any one of the problems is solved, then the other two are solved.

In paper [Guo et al., 2005] a new formula connected the input-output mutual information and the minimum mean-square error (MMSE) achievable by optimal estimation of the input given the output. That is, the derivative of the mutual information (in nats units) with respect to the signal-to-noise ratio (SNR) is equal to half the MMSE, regardless of the input statistics

$$\frac{dI\left(SNR\right)}{dSNR} = \frac{1}{2}MMSE\left(SNR\right).$$

Where, mutual information is $I(x, y)$ between the input x and the output y of a channel, and $MMSE$ is the minimum mean-square error in estimating the input with a given

output.
Thus, an estimation-theoretic parameter MMSE is correlated with information-theoretic parameter of mutual information or Shannon information. CRB gives the lower bound of covariance which is nothing but the reciprocal of Fisher Information Matrix (**FIM**).

Researchers in [Fernando and Nicholson, 1981] have defined measures to quantify the *Degree of Controllability due to Individual Inputs* based on an information theoretic measure called *Mahalanobis Distance*. Mahalanobis Distance is defined as follows :

$$d_i{}^2(x) = (x-\mu)^T \Sigma^{-1} (x-\mu)$$

where μ is the mean value of a random vector x, T represents the Transpose of the matrix, Σ is the covariance matrix and this distance is measured from the mean. This is also a measure of *oscillatory energy* of the vector x^i. where $x^i \in X^i, x \in X, X^i \subseteq X$ for individual input $i = 1, \ldots, m$.

$\Sigma = E\left[(x-\mu)(x-\mu)^T\right]$ where E denotes the expectation operator.

Mahalanobis Distance can also be defined as

$$D_M^2(x^i, \bar{x^i}) = \left\|x^i - \bar{x^i}\right\|^2 \Sigma^{-1}$$

According to [Fernando and Nicholson, 1981] "The loci of points of constant energy defines a hyper-ellipsoid with the principal axes in the directions of the eigenvectors of the covariance matrix Σ. The lengths of semiaxes are given by the square roots of the eigenvalues. Mahalanobis distance, which is an information theoretic metric measure, can be used as an index to investigate the control effectiveness of individual inputs in multi-variable control systems. However, the basic weakness of the existing approaches is that the controllability gramian $\mathbf{W}_c$ - matrix is not invariant under similarity transformations of the system (A, B). Mahalanobis distance is scale-invariant and hence does not depend on the units or scales of measurement".

Assuming that linear system (A, B) is asymptotically stable and fully controllable. The controllability gramian matrix W_c for deterministic unit input impulses is given by

$$W_c(\tau) = \int_0^\tau x(t)x^T(t)dt.$$

For stochastic inputs of the form

$$E[u(t)] = 0 \text{ , } E[u(t)u^T(\tau)] = I\delta(t-\tau)$$

W_c matrix can be taken as the covariance

$$W_c = \lim_{t\to\infty} W_c(t) = \lim_{t\to\infty} E\left[x(t)x^T(t)\right]$$

Without loss of generality we can assume that with $\tau \to \infty$ the results for deterministic and stochastic approaches coincide. If there is a white noise at the *i*-th input with all other inputs being held at zero, the response of the system is given by

$x^i(t) = e^{At}b^i, \;\; x^i(t), \;\; b^i \in M_{n,1}.$

where b^i being the *i*-th column of the matrix B.

Now we can compare the energy in the controllability subspace due to the i th input only using the Mahalanobis distance. This is a measure of the *Degree of Controllability* of the state space, and this index is defined by

$d_i{}^2(x) = \frac{1}{n} trace W_c^{-1} W_c^i$

Where, controllability gramian matrix due to individual i th input, represented by

$W_c^i = \lim_{t\to\infty} E\left[x^i(t)\left\{x^i(t)\right\}^T\right] = \int_0^\infty e^{At}b^i\left(b^i\right)^T e^{A^T t}dt.$

5.3 Achievable Rate of Information Transfer and Bode's Integral Formula

Referring to [Elia, 2004] we consider a linear time-invariant (LTI), single-inputsingle-output (SISO), discrete-time, unstable system with the following state-space model:

$$\begin{aligned} K: \begin{bmatrix} x_s(k+1) \\ x_u(k+1) \end{bmatrix} &= \begin{bmatrix} A_s & 0 \\ 0 & A_u \end{bmatrix}\begin{bmatrix} x_s(k) \\ x_u(k) \end{bmatrix} + \begin{bmatrix} B_s \\ B_u \end{bmatrix} y(k) \\ u(k) &= \begin{bmatrix} C_s & C_u \end{bmatrix} + \begin{bmatrix} x_s(k) \\ x_u(k) \end{bmatrix} \end{aligned} \tag{5.1}$$

The eigenvalues of A_u are outside the unit disc (strictly unstable) while the eigenvalues of A_s are all strictly inside the unit disc (strictly stable). Without loss of generality we assume that A_s and A_u are in Jordan form. Assuming that A_u has m eigenvalues denoted by $\lambda_i(A_u)$ with $i = 1, \ldots, m$.

$DI(K) = \prod_1^m |\lambda_i(A_u)|$ is called the degree of instability *(DI)* of K.

We also consider an LTI, SISO, discrete-time, stable channel with an additive Gaussian white noise $w(k)$ at its output

$$\begin{aligned} F: x_c(k+1) &= A_f x_c(k) + B_f u(k) \\ y(k) &= C_f x_c(k) + D_f u(k) + w(k) \end{aligned} \tag{5.2}$$

In compact notation, the channel can be represented as follows :

$$y = Fu + w$$

where w is an additive white Gaussian noise with zero mean and unit variance without loss of generality. If S be the Sensitivity Transfer Function from Additive White Gaussian Noise (AWGN) w to output y and the average directed information $\vec{I}\left(U^T \to Y^T\right)$, then they are equal as illustrated below and thus, they are two different ways of looking at the same fundamental limitation. Let h be the differential entropy defined as follows: Now, for multidimensional continuous case, entropy (precisely *differential entropy*) of a continuous random variable X with probability density function $f(x)$ (if $\int_{-\infty}^{\infty} f(x)dx = 1$) is defined [Cover and Thomas, 2006] as

Differential Entropy $h(X) = -\int_{\mathcal{S}} f(x) \ln f(x)dx$;

where the set $\mathcal{S}$ for which $f(x) > 0$ is called the *support set* of X.
As in discrete case, the differential entropy depends only on the probability density of the random variable and therefore the differential entropy is sometimes written as $h(f)$ rather than $h(X)$.

Theorem 5.1. *[Elia, 2004]:*
$\lim_{T\to\infty} \frac{1}{T}\vec{I}\left(U^T \to Y^T\right) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln\left|S(e^{j2\pi\theta})\right| d\theta = \sum_{i=1}^{m} \ln|\lambda_i(A_u)|$. *where* $\lambda_i(A_u)$ *are the eigen values corresponding to unstable poles (lying outside the unit disk).*

Definition 5.1. *Directed information* [Massey, 1990] is defined as :

If $X^N = \{X_1, X_2, \ldots, X_N\}$ and $Y^N = \{Y_1, Y_2, \ldots, Y_N\}$ are the input and output sequences respectively of a channel, then directed information is given by

$$I(X^N \to Y^N) \triangleq \sum_{n=1}^{N} I(X^n; Y_n|Y^{n-1}).$$

Also, $I(X^N \to Y^N) \leq I(X^N; Y^N)$

with equality if the channel is used without feedback. Here, $I(X^N; Y^N)$ being mutual information.

Proof :

$$\vec{I}\left(U^T \to Y^T\right)$$

$$= \sum_{t=0}^{T} I\left(U^t; Y_t|Y^{t-1}\right)$$

$$= \sum_{t=0}^{T} h\left(Y_t|Y^{t-1}\right) - h\left(Y_t|Y^{t-1}, U^t\right)$$

$= \sum_{t=0}^{T} h\left(Y_t|Y^{t-1}\right) - h\left(W_t + \left(FU^t\right)_t|Y^{t-1}, U^t\right)$

$= \sum_{t=0}^{T} h\left(Y_t|Y^{t-1}\right) - h\left(W_t|W^{t-1}, U^t\right)$

$= \sum_{t=0}^{T} h\left(Y_t|Y^{t-1}\right) - h\left(W_t|W^{t-1}\right)$

$= h\left(Y^T\right) - h\left(W^T\right) = \frac{1}{2}\ln\frac{\det\left[\Sigma_{Y^T}\right]}{\det\left[\Sigma_{W^T}\right]}$

where Σ and h represent the Error Covariance Matrix and Entropy respectively.

$Now, \lim_{T\to\infty}\frac{1}{T}\vec{I}\left(U^T \to Y^T\right) = \lim_{T\to\infty}\frac{1}{2T}\ln\frac{\det\left[\Sigma_{Y^T}\right]}{\det\left[\Sigma_{W^T}\right]}$

$= \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln\frac{\left|Y(e^{j2\pi\theta})\right|^2}{\left|W(e^{j2\pi\theta})\right|^2} d\theta = \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln\left|S(e^{j2\pi\theta})\right| d\theta = \sum_{i=1}^{m} \ln\left|\lambda_i\left(A_u\right)\right|.$

The last equality is the Poisson-Jensen's formula, and holds because the open-loop system is strictly proper and the sensitivity function is analytic and stable, thus the integral only depends on the zeros of S outside the unit disc. However, such zeros are the eigenvalues of A_u, since the channel is assumed to be stable. Stabilizing an unstable system through a communication channel is equivalent to communicating information at a bit rate equal to the $\log$ of the degree of instability of the system. In loose terms, while the instability of the dynamical system can be controlled by channel state in the feedback communication system, this is often not possible in the feedback control problem, where the plant instability is independent from the channel state.

5.4 Relation among Entropy, Unstable Poles and Bode-Integral

5.4.1 Relation between Entropy and Bode-Integral

It is established by researchers [Zhang and Sun, 2003] that the difference between the entropy rates of the two signals equals Bode Integral (which measures the performance of disturbance rejection). Recalling, (2.1) we can express mathematically,

$$H_c\left(x\right) - H_c\left(d\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left|S\left(e^{j\omega}\right)\right| d\omega = \sum_k \log\left(p_k\right) \tag{5.3}$$

We are presenting now the proof of the relation as given above.

Let, $X_1^n = \{x_1, x_2, \cdots, x_n\}$ and $Y_1^n = \{y_1, y_2, \cdots, y_n\}$ be the sets of observations of stochastic processes X and Y respectively. Then, entropy rate of a discrete-time stochastic process X is

$\bar{H}(X) = \lim_{n\to\infty}\frac{1}{n}H(X_1^n)$ = Per unit time uncertainty of X

Mutual Information Rate between two stochastic processes X and Y is

$\bar{I}(X;Y) = \lim_{n\to\infty} \frac{1}{n} I(X_1^n; Y_1^n)$

which measures the average information transmitted between processes X and Y, where $I(X_1^n; Y_1^n)$ is the mutual information between X_1^n and Y_1^n.
Referring to [Iglesias et al., 1990]:

Let, any transfer function $G(z)$ with $z = e^{j\omega}$ satisfies

$\|G(z)\|_\infty = \sup_\omega \bar{\sigma}\left[G(e^{j\omega})\right] < \lambda;$

where, $\bar{\sigma}$ is the maximum singular value and $G^*(e^{j\omega}) = G^T(e^{-j\omega})$. Then the entropy of $G(z)$ in H_∞ control defined in [Mustafa and Glover, 1990] [Qiu and Chen, 1999] as

$E(G,\lambda) = -\frac{\lambda^2}{4\pi}\int_{-\pi}^{\pi} \ln\left[I - \lambda^{-2} G^*(e^{j\omega}) G(e^{j\omega})\right] d\omega$

$E(G,\lambda)$ entropy being different from Shannon entropy. It is known that this H_∞ entropy denoted by $E(G,\lambda)$ is in fact a performance index measuring the trade-off between the H_∞ optimality and H_2 optimality [Zhou, 1998].
Now let, $G(z) \in \Re H_\infty$ be the $m \times m$ transfer function matrix of a discrete-time MIMO LTI system where $\Re H_\infty$ denotes the set of all stable and proper rational transfer function matrices and stationary stochastic input $x(k) \in \Re^m$ with $k = 0, 1, 2, \cdots$ has a spectral density $S_x(\omega)$. Then, the entropy rate of system output $y(k) \in \Re^m$ is

$\bar{H}(y) = \bar{H}(x) + \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left|\det G(e^{j\omega})\right| d\omega$

$\bar{H}(x)$ being the entropy rate of the input x given by

$\bar{H}(x) = \frac{1}{2}\ln\left\{(2\pi e)^m\right\} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\det S_x(\omega) d\omega.$

Proof. x being Gaussian process , $x^n = \{x_1, \ldots, x_n\}$

Toeplitz Matrix $T_{x^n} = [Q(j-k)]_{j,k=-n,\ldots,-1,0,1,\ldots,n}$;

where $[Q(j-k)] \in \Re^{m\times m}$ is the covariance matrix of x ;

Then, entropy of x^n is $H(x^n) = \frac{1}{2}\ln(2\pi e)^{nm} + \frac{1}{2}\ln|T_{x^n}|$

From Szego Theorem and the properties of Toeplitz Matrix

$\lim_{n\to\infty}\left[\det T_{x^n}\right]^{\frac{1}{n}} = e^{\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\det S_x(\omega) d\omega\right]}.$

Hence, $\bar{H}(x) = \frac{1}{2}\ln(2\pi e)^n + \frac{m}{2} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\det S_x(\omega)d\omega$

Similarly, $\bar{H}(y) = \frac{1}{2}\ln(2\pi e)^n + \frac{m}{2} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\det S_y(\omega)d\omega$

Since, $S_y(\omega) = G(e^{j\omega})S_x(\omega)G^*(e^{j\omega})$

Therefore,$\bar{H}(y) = \frac{1}{2}\ln(2\pi e)^n + \frac{m}{2} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\det S_x(\omega)d\omega +$

$\frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\det\left[G(e^{j\omega})G^*(e^{j\omega})\right]d\omega$

$= \bar{H}(x) + \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\det\left[G(e^{j\omega})\right]d\omega$

Hence, $\bar{H}(y) = \bar{H}(x) + \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\det\left[G(e^{j\omega})\right]d\omega$.

$$\Rightarrow \bar{H}(y) - \bar{H}(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\det\left[G(e^{j\omega})\right]d\omega \tag{5.4}$$

□

i.e. the Entropy Rate Difference between the output and input is the Bode Integral. Thus, information-theoretic entropies are related to control-theoretic Bode Integral. The reduction in entropy can also be interpreted as the gain in information.

5.4.2 Relation between Bode-Integral and Unstable Poles

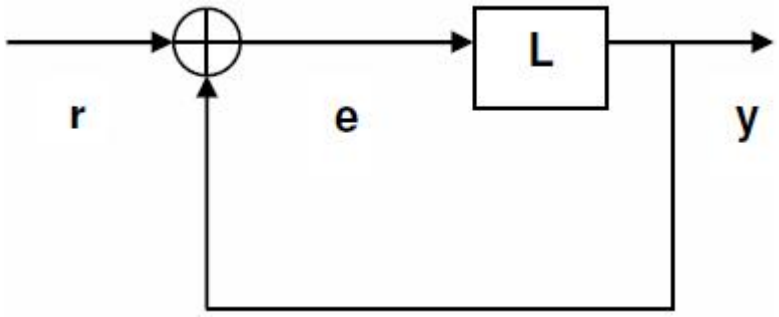


Figure 5.1: Sensitivity Diagram

Based on the paper [Wu and Jonckheere, 1992] and referring to the Fig. (5.1) we can establish the relation as follows :

Consider a SISO LTI discrete-time system with Sensitivity $S(z)$ and open-loop transfer function $L(z)$.

$L(z) \triangleq \frac{K\prod_{i=1}^{m}(z-z_i)}{\prod_{i=1}^{n}(z-p_i)}$

where $K \neq 0$ and is chosen to stabilize the closed-loop system. p_i are open-loop poles, some of them are allowed outside the open-unit disk. z_i are zeros. The sensitivity function of the system is defined as

$S(z) = \frac{1}{1+L(z)}$.
Then,

$$\int_{-\pi}^{\pi} \ln\left|S(e^{j\omega})\right| d\omega = 2\pi\left(\sum_i \ln|p_{ui}| - \ln|\gamma+1|\right). \tag{5.5}$$

where p_{ui} are unstable open-loop poles (outside the closed-unit disk) and $\gamma \triangleq \lim_{z\to\infty} L(z)$. Let the *relative degree* $\nu \triangleq n-m$.

1. **Case - 1** : If $\nu \geq 1$, then $\gamma = 0$ and

   $S(z) = \frac{\prod_{i=1}^{n}(z-p_i)}{\prod_{i=1}^{n}(z-r_i)}$

   where p_i and r_i are open-loop and closed-loop poles respectively. Note that the degrees of denominator and numerator of $S(z)$ are the same. since, the closed-loop system is asymptotically stable, $|r_i| < 1$ for $i = 1, 2, \ldots, n$; and since the open-loop poles, which are inside the closed-unit disk do not contribute to this integration in view of the following Lemma (For proof of this lemma refer to [Wu and Jonckheere, 1992]).

   Lemma 5.1. $\int_{-\pi}^{\pi} \ln\left|e^{j\omega}-r\right|^2 d\omega = \begin{cases} 0 & \text{if } |r|<1 \\ 2\pi\ln|r|^2 & \text{if } |r|>1 \end{cases}$, (Here $\theta = arg(r)$ the usual notation in polar form)

   Hence, we have

   $\int_{-\pi}^{\pi} \ln\left|S(e^{j\omega})\right|^2 d\omega = 4\pi\sum_i \ln|p_{ui}|$.

   Since $\gamma = 0$ in this case, we can rewrite the above as

   $\int_{-\pi}^{\pi} \ln\left|S(e^{j\omega})\right| d\omega = \frac{1}{2}\int_{-\pi}^{\pi} \ln\left|S(e^{j\omega})\right|^2 d\omega = 2\pi\left(\sum_i \ln|p_{ui}| - \ln|\gamma+1|\right)$

   For *relative degree* $\nu \geq 1$, say $\nu = n-m \geq 2$, then

   $$\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left|S(e^{j\omega})\right| d\omega = \sum_i \ln|p_{ui}| \tag{5.6}$$

Hence, we can conclude that Bode Integral equals the summation of logarithms of unstable poles.

2. **Case - 2** : If *relative degree* $\nu = 0$, i.e. $n = m$ then

$$L(z) \triangleq \frac{\gamma \prod_{i=1}^{n} (z - z_i)}{\prod_{i=1}^{n} (z - p_i)}$$

If $\gamma \neq -1$, then

$$S(z) = \frac{\prod_{i=1}^{n} (z - p_i)}{(\gamma+1) \prod_{i=1}^{n} (z - r_i)}$$

There is one more term, $\gamma + 1$, here. Then, following the same approach as **Case - 1**, we get

$$\int_{-\pi}^{\pi} \ln |S(e^{j\omega})| \, d\omega = 2\pi \left(\sum_i \ln |p_{ui}| - \ln |\gamma + 1| \right).$$

If $\gamma = -1$, then atleast one of the closed-loop poles is at infinity, which is outside the unit circle, so that the closed-loop system is not causal. Therefore, $\gamma = -1$ is not allowed when $\nu = 0$.

From the discrete system Bode Integral and unstable pole relation as illustrated above implies that, if the open-loop system is stable and $\nu \geq 1$, i.e. $\gamma = 0$, then

$$\int_{-\pi}^{\pi} \ln |S(e^{j\omega})| \, d\omega = 0.$$

This is Bode's Theorem for discrete-time systems.

Combining the Entropy - Bode Integral (5.4) and Bode Integral - Unstable Pole relations (5.6) we get the final relation as follows :

$$\bar{H}(y) - \bar{H}(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \det \left[G(e^{j\omega}) \right] d\omega = \sum_i \ln |p_{ui}|$$

5.5 Entropy-Theoretic Control

At the execution level, the design of the desirable control may be expressed by the uncertainty of selecting the optimal control that minimizes a given performance index. By choosing a density function over the set of admissible controls to minimize the differential control entropy , one may show that the optimal control problem is equivalent to the problem of minimization of the assigned entropy function with respect to the associated control. It is shown [Saridis, 1988] that the optimal entropy is decomposed into

three terms: the optimal control term with given parameters, the parameter identification term, and the equivocation term which accounts for the active transition of dual control.
In [Conant, 1976] was created a mathematical model for an intelligent control system using entropy as the common measure. According to this formulation, the solution of the intelligent control problem can be formulated as the sequence of decisions in a hierarchical system structured according to the principle of increasing precision with decreasing intelligence, that minimized the total entropy system. It is natural for the higher levels of the system to deal with knowledge-based concepts that use information theoretic techniques to assign probability distributions that generate entropy measures. It is not so easy to do the same with the execution level that requires the application of system theoretic methods to control the system's hardware as postulated earlier. This requires the expression of control system theoretic measures in terms of information theoretic measures, e.g., entropy.

A reformulation of the optimal control problem is based on the idea of expressing the design of the desirable control by the uncertainty of selecting a control law that minimized a given performance index. This is manifested by choosing a density function law over the set of admissible control for the particular problem and obtaining its differential entropy. If the density function is selected to satisfy Jaynes' principle of maximum entropy, it is shown that minimization with respect to control entropy $H(u)$ is equivalent to the minimization of the given performance index.

The purpose here is to unify the theory of optimal, approximately optimal, and adaptive control design, by using entropy measures, which represents a different approach of looking at the problem. The problem of control system design will be posed probabilistically by assigning a distribution function representing the uncertainty of selection of the appropriate control law over the space of admissible controls. By selecting the distribution to satisfy Jaynes' maximum entropy criterion [Jaynes, 1957] the control's performance criterion is associated with the entropy of selecting the appropriate control. Minimization of the differential entropy, which is equivalent to the average performance of the system, yields the optimal control.
Consider the system

$$\begin{aligned} \dot{x}(t) &= f(x,u,w,t); \\ z(t) &= g(x,\nu,t); \\ x(t_0) &= x_0; \text{ with probability density } p(x_0); \\ p(w(t)) &= \text{ probability density associated with } w(t); \\ p(\nu(t)) &= \text{ probability density associated with } \nu(t); \end{aligned}$$

where, state x, control u, Process noise w, time t, z being measured value, ν being measurement noise.

Since only the measurements $z(t)$ are available for feedback, a state estimation procedure $\hat{x}(t)$ should be implemented to yield the optimal solution. The cost of the additional state estimation process may be computed as follows :

Let, Control entropy $H(u)$ corresponding to a cost $J(u(x,t),t)$ for some control $u(x,t) \in \Omega_u$. Using the marginal probabilities over the n-dimensional space defined by the measurements $\Omega_{\hat{x}} \subset \Omega_x$, for specifying the optimal estimation and control processes :

$$H(u) = -\int_{\Omega_u} p(u)\ln p(u)dx = -\int_{\Omega_u}\int_{\Omega_{\hat{x}}} p(u,\hat{x})\ln p(u)dxd\hat{x}$$

$$= -\int_{\Omega_u}\int_{\Omega_{\hat{x}}} p(u|\hat{x})p(\hat{x})\left[\ln p(u|\hat{x})p(\hat{x})/p(\hat{x}|u)\right]d\hat{x}dx$$

$$= -\int_{\Omega_{\hat{x}}}\left[\int_{\Omega_u} p(u|\hat{x})\ln p(u|\hat{x})dx\right]p(\hat{x})d\hat{x} - \int_{\Omega_{\hat{x}}}\left[\int_{\Omega_u} p(u|\hat{x})dx\right]\ln p(\hat{x})p(\hat{x})d\hat{x}$$

$$+\int_{\Omega_u}\left[\int_{\Omega_{\hat{x}}} p(\hat{x}|u)\ln p(\hat{x}|u)d\hat{x}\right]p(u)dx.$$

$$= H(u|\hat{x}) + H(\hat{x}) - H(\hat{x}|u).$$

Therefore, $H(u) = H(u|\hat{x}) + H(\hat{x}) - H(\hat{x}|u)$

i.e. The cost of control with inaccessible states is given by the cost of control with estimated states plus the cost of state estimation minus the equivocation of active transmission of information
$\Rightarrow$ Control Cost (with inaccessible states) $=$ Control Cost with estimated states $+$ Estimation Cost $-$ Equivocation of Active Transmission of Information.
The optimal control cost is given by

$$H(u^*) = \min_u H(u|\hat{x}) + \min_{\hat{x}} H(\hat{x}) - \max_{\hat{x}} H(\hat{x}|u).$$

which is equal to the optimal deterministic control with states replaced by their optimal estimates plus the minimum cost of state estimation minus the maximum active transmission of information from the controller to the estimator.
Discussion :
The result above has given the credit of reducing the uncertainty by providing *a priori* information or the conditional version of uncertainty but the result derived in [Tatikonda, 2000] shows only the splitting of optimal cost into two – control cost (full state cost) and communication cost (dependent only on state estimation error covariance) with no credit to the information which can be utilized to reduce the cost in terms of uncertainty. In paper [Zhang and Liu, 2008] *LQG* cost is divided into the control cost, the communication cost, and the sampling cost. The controller prefers to have more frequent communication with the plant, even if the information is crude instead of to obtain a piece of more precise information with a long delay. But, in a control problem under constrained communication this scenario does not fit well simply because passing on information as and when generated leads to severe load on the already communication constrained channel of limited bandwidth. This scenario would severely aggravate the problem rather than solving it. Optimal *LQG* cost comprises of control cost under full

observation of the state, the communication cost and the sampling cost. All costs depend on the sampling rate but only the communication cost depends on the channel capacity and in communication-constrained control problem the communication cost would severely increase by communicating information at a faster rate causing possible violation of channel capacity. Hence, the need is to find the solution for control under communication through noisy channel which is realistic and makes sense.
Therefore, the result discussed (derived) in details in this section has the potential of reducing the control cost by the way of providing information so as to reduce the uncertainty rather than to increase the load on communication channel indiscriminately. The better way of reducing the communication cost by working on the protocol aspects of communication and co-designing control and communication. One possible way is the scheduling part of control as suggested by us in Chapter - 8 of this thesis, [Ben Gaid, 2006] and the associated references cited therein.

5.6 Mutual Information and Optimal Estimation Error

Based on [Freudenberg et al., 2007] the relation between optimal prediction estimation error and mutual information can be established. The (differential) entropy rate of a stationary, continuous-valued, discrete-time scalar random process a is given by [Cover and Thomas, 2006]

$h_\infty(a) = \lim_{k\to\infty} h_\infty(a_k|a^{k-1})$.

The entropy rate of a stationary Gaussian random process a may be computed from its power spectral density $S_a(\omega)$ [Cover and Thomas, 2006] by the formula

$h_\infty(a) = \frac{1}{2}\ln 2\pi e + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln S_a(\omega)d\omega$.

The entropy rate of a random sequence a is related to the problem of estimating from previous values a^{k-1}.
Denoting such an estimate by $\hat{a}_{k|k-1}$, and the resulting estimation error by $\tilde{a}_{k|k-1} = a_k - \hat{a}_{k|k-1}$;
Then the variance of the minimal mean square estimation error satisfies (refer page-255 of [Cover and Thomas, 2006])

$$E\left[\tilde{a}^2_{k|k-1}\right] = \frac{1}{2\pi e}e^{2h_\infty(a)}. \tag{5.7}$$

Applying this relation to the feedback control system, the power spectrum of the channel (having s and r as the input and output respectively) output may be written as

$S_r(\omega) = \left|S(e^{j\omega})\right|^2 \sigma_n^2 \left(1 + \left|G_d(e^{j\omega})\right|^2 \frac{\sigma_d^2}{\sigma_n^2}\right)$. G_d being the transfer function with disturbance input

Since disturbance d and channel noise n are assumed Gaussian, the channel output is also Gaussian, with entropy rate given by

$$h_\infty(r) = \frac{1}{2}\ln 2\pi e\sigma_n^2 + \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\left|S(e^{j\omega})\right|d\omega + \frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\left(1+\left|G_d(e^{j\omega})\right|^2\frac{\sigma_d^2}{\sigma_n^2}\right)d\omega. \quad (5.8)$$

Assuming the plant with control input u and the transfer function G_u is strictly proper, having m unstable poles $\phi_i > 1$, and no non-minimum phase zeros. Then it is possible to stabilize the system using a controller with no unstable poles, and the sensitivity function must satisfy the Bode integral [Seron et al., 1997]

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\left|S(e^{j\omega})\right|d\omega = \sum_{i=1}^{m}\ln|\phi_i|\,. \quad (5.9)$$

The mutual information rate [Martins and Daleh, 2008] between the disturbance and channel output satisfies

$$I_\infty(r;d) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\left(1+\left|G_d(e^{j\omega})\right|^2\frac{\sigma_d^2}{\sigma_n^2}\right)d\omega. \quad (5.10)$$

Substituting (5.9) and (5.10) into (5.8) and applying the formula (5.7) shows that the minimum mean square error in estimating r_k given r^{k-1} is given by

$$E\left[\tilde{r}_{k|k-1}^2\right] = \sigma_n^2\prod_{i=1}^{m}|\phi_i|^2\,e^{2I_\infty(r;d)}. \quad (5.11)$$

Hence, the minimal channel capacity required for stabilization is thus

$$C > \sum_{i=1}^{m}\ln|\phi_i| + I_\infty(r;d)\text{nats/transmission} \quad (5.12)$$

Interpretation of the two terms that contribute to channel capacity in (5.12) is as follows:

1. the Channel Capacity required for Stabilization alone is given by $\sum_{i=1}^{m}\ln|\phi_i|$
2. the Additional Capacity required to Stabilize in the presence of a *disturbance* depends on the mutual information between the disturbance and the channel output, once the plant has been stabilized.

5.7 Conclusion

In this chapter, we have analyzed and discussed the convergence of the control, communication, information and estimation theory. We have presented an entropy theoretic explanation of optimum control and formulation of the relation concerning the information-theoretic entropy, control-theoretic Bode integral and unstable poles. Relation between Shannon Mutual-Information and Estimation Error has also been presented.

6

Information-Theoretic View of Control

6.1 Introduction

In recent years, there has been an increased interest for the fundamental limitations in feedback control. Bode's sensitivity integral (Bode Integral, in short) is a well-known formula that quantifies some of the limitations in feedback control for linear time-invariant systems. In [Sandberg and Bernhardsson, 2005], it is shown that there is a similar formula for linear time-periodic systems.

In this chapter, we are presenting the information-theoretic explanation of Bode Sensitivity Integral, a fundamental limitation of control theory, controllability gramian and the issues of control under communication constraints based on our paper [Roy et al., 2009a]. As resource-economic use of information is of prime concern in communication-constrained control problems, we need to emphasize more on informational aspect which has got direct relation with uncertainties in terms of Shannon Entropy and Mutual Information. Bode Integral which is directly related to the disturbances can be correlated with the difference of entropies between the disturbance-input and measurable output of the system (refer Eq. 2.1). These disturbances due to communication channel-induced noises and limited bandwidth are causing the information packet-loss and delays resulting in degradation of control performances. Shannon Entropy of information theory is a stronger metric for uncertainty which hinders control of a system.

It has been known that control theory and information theory share a common background as both theories study signals and dynamical systems in general. In distributed embedded systems (*DES*) or networked control systems (*NCS*), there are issues related to both control and communication since communication channels with data losses, time delays, and quantization errors are employed between the plants and con-

trollers [Antsaklis and Baillieul, 2007]. To guarantee the overall control performance in such systems, it is important to evaluate the quantity of information that the channels can transfer. Thus, for the analysis of DES or NCS, information theoretic approaches are especially useful, and notions and results from this theory can be applied. The results in [Nair and Evans, 2004] and [Tatikonda and Mitter, 2004] show the limitation in the communication rate for the existence of controllers, encoders, and decoders to stabilize discrete-time linear feedback systems.
The focus of information theory is more on the signals and not on their input-output relation. Thus, based on information theoretic approaches, we may expect to extend prior results in control theory. One such result can be found in [Martins et al., 2007], where a sensitivity property is analyzed and Bode's integral formula [Bode, 1945] is extended to a more general class of systems. A fundamental limitation of sensitivity functions is presented in relation to the unstable poles of the plants.

6.2 Problem Formulation

Networked control systems suffer from the drawbacks of packet losses, delays and quantization in particular. These cause degradation of control performances and under some conditions instability. Uncertainties due to packet losses, delays, quantization, communication channel induced noises etc. have a great influence on the system's performance. If we consider only the uncertainties induced by channel noise and quantization we may write:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + w(t); \\ y(t) &= Cx(t) + Du(t) + v(t); \end{aligned} \tag{6.1}$$

where $A \in R^{n\times n}$ is the system or plant matrix and $B \in R^{n\times q}$ is the control or input matrix. Also, $x(t)$ is the state, $u(t)$ is the control input, $y(t)$ is the output, C is the output or measurement matrix, D is the Direct Feed matrix, $w(t)$ and $v(t)$ are the external disturbances and noises of Gaussian nature respectively.
Our aim is to achieve better control performance of system by tackling these uncertainties using Shannon's Mutual-Information, Information-Theoretic Entropy and Bode Sensitivity. We present the information-theoretic model of such uncertainties and their possible reduction using information measures.

6.3 Preliminaries

By means of the connection between Bode integral and the entropy cost function, paper [Iglesias, 2001] provided a time-domain characterization of Bode integral. The traditional frequency domain interpretation is that, if the sensitivity of a closed-loop system

is decreased over a particular frequency range – typically the low frequencies the designer "pays" for this in increased sensitivity outside this frequency range. This interpretation is also valid for the time-domain characterization presented in [Iglesias, 2001] provided one deals with time horizons rather than frequency ranges. Time-domain characterization of Bode's integral shows how the frequency domain trade-offs translate into the time-domain. Under the usual connection between the time and frequency domains: low (high) frequency signals are associated with long (short) time horizons. In Bode's result, it is important to differentiate between the stable poles, which do not contribute to the Bode sensitivity integral and the unstable poles, which do. Time-varying systems which can be decomposed into stable and unstable components are said to possess an exponential dichotomy. What the exponential dichotomy says is that the norm of the projection onto the stable subspace of any orbit in the system decays exponentially as $t \to \infty$ and the norm of the projection onto the unstable subspace of any orbit decays exponentially as $t \to -\infty$, and furthermore that the stable and unstable subspaces are conjugate. The existence of an exponential dichotomy allows us to define a stability preserving state space transformation (a Lyapunov transformation) that separates the stable and unstable parts of the system.

6.3.1 Mutual Information

Recalling the section 2.2, Mutual Information $I(X;Y)$, between X as the input variable and Y as the output variable, has the lower and upper bounds given by the following:

$R(D) =$ Rate Distortion $= MinI(X;Y)$
$C =$ Communication Channel Capacity $= MaxI(X;Y)$

where D is the distortion which happens when information is compressed (i.e. fewer bits are used to represent or code more frequent or redundant informations) and entropy is the limit to this compression i.e. if one compresses the information beyond the entropy limit there is a high probability that the information will be distorted or erroneous. This is as per Shannon's Source Coding Theorem. We code more frequently used symbols with fewer number of bits and vice-versa.

6.3.2 Shannon Entropy

Recalling the definition of Shannon entropy discussed in section 2.2 :

$$H = -\sum_{i} p_i \log_2 (p_i) \tag{6.2}$$

where p_i is the probability of the alternative i. The above quantity is known as the binary entropy in *bits* as we use logarithmic base of 2 (with logarithmic base e the entropy

is in *nats*), and was shown by Shannon to correspond to the minimum average number of bits needed to encode a probabilistic source of N states distributed with probability p_i. The term entropy is associated with the uncertainty or randomness whereas information is used to reduce this uncertainty. Uncertainty is the main hindrance to control and if we can reduce the uncertainty by getting the relevant information and utilizing the information properly so as to achieve the desired control performance of the system. Many researchers have posed the same question: "*How much information is required for controlling the system based on observed informations in the case where these informations are passed through communication channels in a networked based system?*"
Intuitive meaning of mutual information is the amount of information (that is, reduction in uncertainty) that each variable is having about the other.

6.3.3 Bode Integral

As pointed out in Chapter - 2 of this thesis a stable system needs no information on its internal state or the environment to assure its stability. For unstable systems the mutual information between the initial state and the output of the system is related to its unstable poles.
The simplest (and perhaps the best known) result (recalling section 2.4) is that, for an open loop stable plant, the integral of the logarithm of the sensitivity is zero; i.e.

$$\int_0^\infty \ln |S_0(j\omega)| \, d\omega = 0$$

Where, S_0 and ω being the sensitivity function and frequency respectively.
For linear systems Bode Integral is the difference in the entropy rates between the input and output of the systems which is an information-theoretic interpretation. For nonlinear system (if the open loop system is globally exponentially stable and has fading memory) this difference is zero. Fading Memory Requirement is used to limit the contributions of the past values of the input on the output. Fading is the commonly encountered phenomenon in wireless networks wherein the signals fade away (attenuates and varies with time and space). Entropy of the signals in the feedback loop help provide another interpretation of the Bode integral formula [Zang and Iglesias, 2003], [Mehta et al., 2006] as follows. Recalling chapter - 2 , Shannon Entropy - Bode Integral Relation can be rewritten as :

$$H_c(x) - H_c(d) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \left| S\left(e^{j\omega}\right) \right| d\omega = \sum_k \log(p_k) \tag{6.3}$$

Consider a random variable $x \in \Re^m$ of continuous type with entropy associated with this is given by

$H(x) := -\int_{\Re^m} p(x) \ln p(x) dx$;

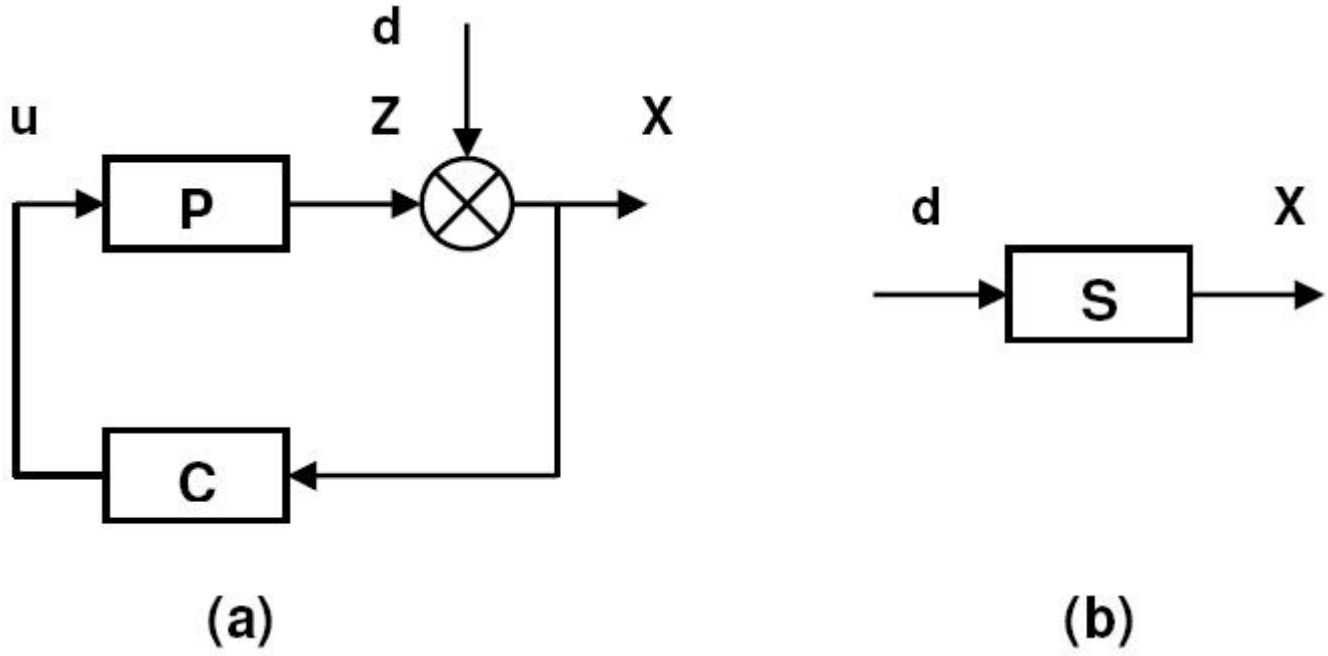


Figure 6.1: (a)Feedback loop and (b)Sensitivity function

where $p(x)$ being the probability density function of x and the conditional entropy of order n is defined as
$H(x_k|x_{k-1},\ldots,x_{k-n}) := -\int_{\Re^m} p(.)\ln p(.)dx$

where $p(.) = (x_k|x_{k-1},\ldots,x_{k-n})$.

This conditional entropy is a measure of the uncertainty about the value of x at time k under the assumption that its n most recent values have been observed. By letting n going to infinity, the conditional entropy of x_k is defined as

$H_c(x_k) := \lim_{n\to\infty} H(x_k|x_{k-1},\ldots,x_{k-n})$ assuming the limit exists.

Thus the conditional entropy is a measure of the uncertainty about the value of x at time k under the assumption that its entire past is observed. Difference of conditional entropies between the output and input is nothing but the Bode sensitivity integral which equals the summation of logarithms of unstable poles.
For a stationary Markov process, conditional entropy [Cover and Thomas, 2006] is given by

$H(x_k|x_{k-1},\ldots,x_{k-n}) = H(x_k|x_{k-1})$.

It is well known that the sensitivity and complementary sensitivity functions represent basic properties of feedback systems such as disturbance attenuation, sensor-noise reduction, and robustness against uncertainties in the plant model. Researchers have worked earlier on the issues of relating the entropy and the Bode Integral and complementary sensitivity. For details refer to the references cited in paper [Roy et al., 2009a]. In [Iglesias, 2002] the sensitivity integral is interpreted as an entropy integral in the time domain, i.e., no frequency-domain representation is used.
One has to gather relevant information, transmit the information to the relevant agent, process the information, if needed, and then use the information to control the system. The fundamental limitation in information transmission is the achievable information rate (i.e. a fundamental parameter of Information Theory), the fundamental limitation in information processing is the Cramer-Rao Bound (*CRB*) which deals with Fisher Information Matrix (**FIM**) in Estimation Theory, and the fundamental limitation in information utilization is the *Bode Integral* (i.e. a fundamental parameter of Control Theory), seemingly different and usually separately treated, are in fact three sides of the same entity as per [Liu and Elia, 2006]. Hence, there is a need of dealing the control problem from information-theoretic point of view.

Even Kalman et al. in their paper [Kalman et al., 1963] have stated that Controllability Gramian (refer Appendix - A) Matrix $\mathbf{W}_c$ - matrix is analogous to **FIM** and the determinant $\det \mathbf{W}_c$ is analogous to Shannon Information.
These research work motivated us to investigate some important correlations amongst mutual information, entropy and design control parameters (like controllability gramian) of practical importance rather than just concentrating on stability issues.

6.4 Information Induced by Controllability Gramian

In general, from the viewpoint of the open-loop system, when the system is unstable, the system amplifies the initial state at a level depending on the size of the unstable poles [Okano et al., 2008]. Hence, we can say that in systems having more unstable dynamics, the signals contain more information about the initial state. Using this extra information (in terms of mutual information between the control input and the initial state) we can reduce the entropy (uncertainty) and thus rendering the observation of initial state easier.
Suppose that we have a feedback control system in which control signal is sent through a network with limited bandwidth. We will consider the case where the state of the system is measurable and the controller can send the state of the system without error. Under these conditions we may write:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu^*(t); \\ u^*(t) &= -K_c x(t) + u^e(t); \end{aligned} \tag{6.4}$$

where K_c, $u^*(t)$ and $u^e(t)$, represent, respectively, the feedback controller gain, the applied control input and control error due to quantization noise of limited bandwidth

network. In the sequel we are supposing that the control signal errors are caused by Gaussian White Noise which may be given by $u_i^e(t) = \sqrt{D_i}\delta(t)$. So we may write :

$$\begin{aligned} \dot{x}(t) &= (A - BK_c)x(t) + Bu^e(t); \\ u^*(t) &= -K_c x(t) + u^e(t); \end{aligned} \tag{6.5}$$

or more compactly:

$$\begin{aligned} \dot{x}(t) &= A_c x(t) + Bu^e(t); \\ u^*(t) &= -K_c x(t) + u^e(t); \end{aligned} \tag{6.6}$$

where, $A_c = A - BK_c$.

The feedback system (6.6) is a stable one which is perturbed by quantization errors or noises due to the bandwidth limitation.

Lemma : The controllability gramian matrix $\mathbf{W}_c$ of system (6.6) can related with the information-theoretic entropy H as follows [Mitra, 1969]:

$$H(x,t) = \frac{1}{2}\ln\left\{\det \mathbf{W}_c(\mathbf{D},t)\right\} + \frac{n}{2}(1 + \ln 2\pi) \tag{6.7}$$

= Average *a priori* uncertainty of the state x at time t for an order n of the system.

Where $\mathbf{D}$ being the Diagonal Matrix (positive definite symmetric matrix) with D_i being the ith diagonal element. Here unit impulse inputs are considered.
and

$$\mathbf{W}_c(\mathbf{D},\tau) = \int_0^\tau e^{A_c t} B\mathbf{D}B^T e^{A_c{}^T t} dt$$

for a system modeled as (6.6).

Proof of Eqn.(6.7) :

Referring to [Cover and Thomas, 2006] we are providing the proof. The input of (6.6) being Gaussian White Noise, the state of the system is with probability density having mean-value $\bar{x}(t) = e^{A_c t}\bar{x}(0)$ and Covariance Matrix Σ at time t is given by

$$\Sigma = E\left\{(x - \bar{x})(x - \bar{x})^T\right\} = \mathbf{W}_c(\mathbf{D},t).$$

In a more detailed form :

$$\begin{aligned} x(t) &= e^{A_c t}x(0) + \int_0^t e^{A_c t}Bu^e(t)dt \\ E\{x(t)\} &= E\left\{e^{A_c t}x(0) + \int_0^t e^{A_c t}Bu^e(t)dt\right\} \\ \Rightarrow \bar{x}(t) &= e^{A_c t}\bar{x}(0) \end{aligned}$$

where $\bar{x}(t)$ denotes the mean value of $x(t)$ and Covariance Matrix

$$\Sigma = E\left\{(x-\bar{x})(x-\bar{x})^T\right\}$$

$$\Rightarrow \Sigma = E\left[\left\{e^{A_c t}x(0) + \int_0^\tau e^{A_c t}Bu^e(t)dt - e^{A_c t}x(0)\right\}\left\{e^{A_c t}x(0) + \int_0^\tau e^{A_c t}Bu^e(t)dt - e^{A_c t}x(0)\right\}^T\right]$$

Therefore, $\Sigma = \int_0^\tau e^{A_c t}B\mathbf{D}B^T e^{{A_c}^T t}dt = \mathbf{W}_c(\mathbf{D},\tau)$.

where, $u_i^e(t) = \sqrt{D_i}\delta(t)$ i.e. weighted impulses and $\mathbf{D}$ being the Diagonal Matrix (positive definite symmetric matrix) with D_i being the ith diagonal element. Here, unit impulse inputs are considered.

$$p(x,t) = \frac{1}{(2\pi)^{n/2}\left\{\det \mathbf{W}_c(\mathbf{D},t)\right\}^{1/2}} e^{\left[-1/2\left\{(x-\bar{x})^T\mathbf{W}_c^{-1}(\mathbf{D},t)(x-\bar{x})\right\}\right]} \tag{6.8}$$

Now, for multidimensional continuous case, entropy (precisely *differential entropy*) of a continuous random variable X with probability density function $f(x)$ (if $\int_{-\infty}^{\infty} f(x)dx = 1$) is defined [Cover and Thomas, 2006] as :
Differential Entropy $h(X) = -\int_S f(x)\ln f(x)dx$;
where the set S for which $f(x) > 0$ is called the *support set* of X.
As in discrete case, the differential entropy depends only on the probability density of the random variable and therefore the differential entropy is sometimes written as $h(f)$ rather than $h(X)$. Here, we call differential entropy as $H(x,t)$ and $f(x)$ as $p(x,t)$ which are correlated as

$$H(x,t) = -\int p(x,t)\ln p(x,t)dx \tag{6.9}$$

Using equation (6.8) in equation (6.9) we get

$$H(x,t) = -\int p(x,t)\left[-\tfrac{1}{2}(x-\bar{x})^T\mathbf{W}_c^{-1}(\mathbf{D},t)(x-\bar{x}) - \ln\left\{(2\pi)^{n/2}\left\{\det \mathbf{W}_c(\mathbf{D},t)\right\}^{1/2}\right\}\right]dx$$

$$H(x,t) = \tfrac{1}{2}E\left[\sum_{i,j}\left\{(X_i-\bar{X}_i)(\mathbf{W}_c^{-1}(\mathbf{D},t))_{ij}(X_j-\bar{X}_j)\right\}\right] + \tfrac{1}{2}\ln\left[(2\pi)^n\left\{\det \mathbf{W}_c(\mathbf{D},t)\right\}\right]$$

$$= \tfrac{1}{2}E\left[\sum_{i,j}\left\{(X_i-\bar{X}_i)(X_j-\bar{X}_j)(\mathbf{W}_c^{-1}(\mathbf{D},t))_{ij}\right\}\right] + \tfrac{1}{2}\ln\left[(2\pi)^n\left\{\det \mathbf{W}_c(\mathbf{D},t)\right\}\right]$$

$$= \tfrac{1}{2}\sum_{i,j}\left[E\left\{(X_j-\bar{X}_j)(X_i-\bar{X}_i)\right\}(\mathbf{W}_c^{-1}(\mathbf{D},t))_{ij}\right] + \tfrac{1}{2}\ln\left[(2\pi)^n\left\{\det \mathbf{W}_c(\mathbf{D},t)\right\}\right]$$

$$= \tfrac{1}{2}\sum_j\sum_i(\mathbf{W}_c(\mathbf{D},t))_{ji}(\mathbf{W}_c^{-1}(\mathbf{D},t))_{ij} + \tfrac{1}{2}\ln\left[(2\pi)^n\left\{\det\mathbf{W}_c(\mathbf{D},t)\right\}\right]$$

$$= \tfrac{1}{2}\sum_j\left\{(\mathbf{W}_c(\mathbf{D},t))(\mathbf{W}_c^{-1}(\mathbf{D},t))\right\}_{jj} + \tfrac{1}{2}\ln\left[\left\{(2\pi)^n\left\{\det\mathbf{W}_c(\mathbf{D},t)\right\}\right]\right.$$

$$= \tfrac{1}{2}\sum_j\mathbf{I}_{jj} + \tfrac{1}{2}\ln\left[\left\{(2\pi)^n\right\}\left\{\det\mathbf{W}_c(\mathbf{D},t)\right\}\right]$$

(Where $\mathbf{I}_{jj}$ is the Identity Matrix)

$$= \tfrac{n}{2} + \tfrac{1}{2}\ln\left[\left\{(2\pi)^n\right\}\left\{\det\mathbf{W}_c(\mathbf{D},t)\right\}\right] = \tfrac{n}{2} + \tfrac{1}{2}\ln\left\{(2\pi)^n\right\} + \tfrac{1}{2}\ln\left\{\det\mathbf{W}_c(\mathbf{D},t)\right\}$$

$$= \tfrac{n}{2} + \tfrac{n}{2}\ln\left\{(2\pi)\right\} + \tfrac{1}{2}\ln\left\{\det\mathbf{W}_c(\mathbf{D},t)\right\}$$

$$\therefore H(x,t) = \tfrac{1}{2}\ln\left\{\det\mathbf{W}_c(\mathbf{D},t)\right\} + \tfrac{n}{2}(1+\ln 2\pi)$$

Based on the equation (6.7) we can write the entropy reduction as

$$\Delta H(x,t) = \tfrac{1}{2}\Delta[\ln\left\{\det\mathbf{W}_c(\mathbf{D},t)\right\}]$$

This shows that the entropy reduction which is same as uncertainty reduction is dependent on Controllability Gramian only. Other term being constant for constant n, gets canceled.

Therefore, $\Delta H(x,t) = H(x(t_1),t_1) - H(x(t_2),t_2)$

$$= \tfrac{1}{2}\ln\left\{\det\mathbf{W}_{c1}(\mathbf{D}_1,t_1)\right\} - \tfrac{1}{2}\ln\left\{\det\mathbf{W}_{c2}(\mathbf{D}_2,t_2)\right\}$$

$$\Rightarrow \Delta H(x,t) = \frac{1}{2}\ln\left\{\frac{\det\mathbf{W}_{c1}(\mathbf{D}_1,t_1)}{\det\mathbf{W}_{c2}(\mathbf{D}_2,t_2)}\right\} \tag{6.10}$$

For simplicity we denote $\Delta H(x,t)$ by ΔH, $\mathbf{W}_{c1}(\mathbf{D}_1,t_1)$ by $\mathbf{W}_{c1}$ and $\mathbf{W}_{c2}(\mathbf{D}_2,t_2)$ by $\mathbf{W}_{c2}$.

Therefore, $\Delta H = \frac{1}{2} \ln \left\{ \det(\mathbf{W}_{c1} \mathbf{W}_{c2}^{-1}) \right\}$

$$\Rightarrow \det(\mathbf{W}_{c1} \mathbf{W}_{c2}^{-1}) = e^{2(\Delta H)}$$

Using the above expression along with the concept of mutual information being the difference of the entropy and the residual conditional entropy i.e. $I(X;U) = H(X) - H(X|U)$ (gain in information is reduction in entropy), we can conclude that Mutual Information $I(X;U)$ between the state X and control input U denoted simply by Shannon Information I_{sh} is given by this ΔH which can be expressed further as

Finally,

$$\det(\mathbf{W}_{c1} \mathbf{W}_{c2}^{-1}) = e^{2(\Delta H)} = e^{2 I_{sh}} \tag{6.11}$$

We may conclude that the uncertainty reduction which is directly related to the $\Delta H(x,t)$ will reduce the variance of the state with respect to the steady-state if $\Delta H(x,t)$ converges to zero. The only influence we have on the control signal is related to that of feedback gain ans scheduling of control signals, to be chosen such that the norm of gramians, represented by $\det(\mathbf{W}_c(D_i,t))$ converge rapidly to their norm to infinity $\det(\mathbf{W}_c(D_\infty,\infty))$. We can also conclude from (6.11) that gramian is directly related to the entropy and in turn with Shannon mutual information.

6.5 Conclusion

This chapter has addressed some new ideas (based on our paper [Roy et al., 2009a]) concerning the relation between control design and information theory. Since the distributed embedded system or networked control system has communication constraints due to limited bandwidth or quantization noises, we must have to adopt a policy of resource allocation which enhances the information transmitted. This may be done possible if we know the characteristics of the networks, the bandwidth constraints and that of the dynamical system under study.

As demonstrated in this chapter, the gramian of controllability constitutes a metric of information theoretic entropy with respect to the noises induced by quantization. Reduction of these noises is equivalent to the design methods proposing a reduction of the controllability gramian norm.

Since zoom - in (near the target) is equivalent to reducing the entropy (contracting the uncertainty domain) and zoom - out (away from the target) is equivalent to increasing the entropy (expanding the uncertainty domain), work in this direction to propose an information-theoretic analysis of the zooming algorithm proposed in [Ben Gaid and Çela, 2006] would be dealt in future.

Controllability gramian is, in some sense, a measure of energy. Higher the controllability gramian ($\mathbf{W}_c$), lower is the energy associated with it and hence lower the uncertainty

(randomness, like the Brownian motion or the physical notion of energy and entropy). Hence, entropy reduction is associated with controllability gramian reduction and entropy reduction is at the cost of gain in Shannon mutual information which is established in this chapter vide equation (6.11).

7 Relation of FIM and Controllability Gramian

7.1 Introduction

In this chapter our work [Roy et al., 2009b] on the correlation of Fisher Information Matrix(**FIM**) and Controllability Gramian(**CG**) based on estimation-theoretic and information-theoretic explanations is being presented. Resource-economic use of information is of primary concern in issues of control under communication constraints. Disturbances due to communication channel-induced noises, quantization noises and limited bandwidth are causing the information packet-loss and delays resulting in degradation of control performance. Controllability Gramian is a measure of control performance and Fisher Information is a measure of accuracy of estimation (which needs to be done in uncertainties caused by packet-loss and delays of the network). Hence, in order to have better controllability, we need to explore the relation between the estimation theoretic (because information losses and delays force us to resort to estimation) / information theoretic parameters and the control theoretic parameters of practical importance.
Embedded systems are prevalent in most of the spheres of life today. These systems have the limited power / battery life, memory, processing power, communication bandwidth and computational capabilities. In order to optimally use the power, memory and bandwidth we need to emphasize on the information transfer aspect as more information transfer-rate means more consumption of battery power and hence drainage of the battery power apart from the load on the communication bandwidth. Our main objective is to use the information optimally using the concepts of information theory, estimation theory and control theory.

Most of the Industrial Embedded Systems (*IES*) have network based control systems. Networked systems are complex dynamical systems composed of a large number of simple systems interacting through a communication medium. These systems arise as natural models in many areas of engineering and sciences, such as sensor networks, autonomous unmanned vehicles, biological networks, and animal cooperative aggregation and flocking. There are two features that are common to all these systems: first, they deal with complex dynamics; second, designing globally optimal behavior for these systems requires the solution of large-scale optimization problems, which typically necessitate a prohibitive amount of computational effort. Desirable features of the operation of these systems include robustness to uncertainties and disturbances, and adaptability to environmental changes.
Since the information in case of uncertainties caused by delays and packet losses lead us to do estimation, we need to focus on estimation theoretic information and its relation with controllability. Researches have been done correlating Estimation Theory with Information Theory and hence we can use these results to achieve our goal of finding the relation between **FIM** and **CG**.

In this chapter, we focus on Controllability Gramian (**CG**) of control theory and Fisher Information Matrix (**FIM**) of Estimation Theory because the latter is a stronger metric for estimation required to be done under uncertainty which hinders control of a system.
In NECS, there are issues related to both control and communication since communication channels with data losses, time delays, and quantization errors are employed between the plants and controllers [Antsaklis and Baillieul, 2007]. To guarantee the overall control performance in such systems, it is important to evaluate the information rate that the channels can transfer. Thus, for the analysis of DES/NCS, information theoretic approaches are especially useful, and notions and results from this theory can be applied. Researchers in [Nair and Evans, 2004] and [Tatikonda and Mitter, 2004] showed the limitation in the communication rate for the existence of controllers, encoders, and decoders to stabilize discrete-time linear feedback systems.
Information theory is focused on the signals and not on the input-output relation. Thus, based on information theoretic approaches, we can expect to extend prior results in control theory.

7.2 Problem Formulation

Referring to Chapter - 6 if we consider only the uncertainties induced by channel noise and quantization we can write the system dynamics as given in (6.1).
In all communication models Gaussian distribution [Rao, 1965] of noise is considered as it possesses the property of having maximum entropy for a given mean and variance and the range$(-\infty, +\infty)$ for the variable.
Our aim is to achieve better control performance (in terms of controllability gramian) of system by tackling these uncertainties using Shannon's Information, Entropy and Fisher

Information. We present the information-theoretic model of such uncertainties and their possible reduction using information measures exploiting the estimation theoretic measure of information such as **FIM**.

7.3 Preliminaries

Here, we discuss first some of the basics on **FIM** and for Controllability Gramian the readers may refer to Appendix - A.

7.3.1 Fisher Information Matrix

Recalling the basics of **FIM** from chapter - 2 we can say that for Gaussian distribution it is related to the variance by *Cramer-Rao Bound* (*CRB*). CRB is a general uncertainty expressing reciprocity between the mean-square-error (e) in an estimate and the Fisher Information FI present in the observed data. Mathematically,

$e^2 \geq \frac{1}{FI}$ or $FI \geq \frac{1}{e^2}$

It is a measure of accuracy in estimating a parameter. In other words,

$FI \geq \frac{1}{\sigma^2}$; σ^2 being the variance

Lower the variance, higher the Fisher Information and vice-versa.

Mathematically, curvature or sharpness has got a definite relation with second derivative as we have explained earlier in this thesis. To a great extent Fisher Information defines what can and what cannot be known about a given physical scenario. **FIM** allows detection of some non-stationary behaviour in situations where Shannon Entropy shows limited dynamics. B. Roy Frieden has characterized **FIM** as a versatile tool to describe the evolution laws of physical systems.

7.3.2 Fisher Information as Metric for Accuracy of a Distribution

By information on an unknown parameter θ contained in a random variable or in its distribution, we mean the extent to which uncertainty regarding the unknown value of θ is reduced as a consequence of an observed value of the random variable. If there is a unique observation with probability 1 corresponding to each value of the parameter, we have a situation where the random variable has the maximum information. On the other hand if the random variable has the same distribution for all values of the parameter, there is no basis for making statements about θ based on an observed value of the random variable. The sensitiveness of a random variable with respect to a parameter may then be judged by the extent to which its distribution is altered by a change in the value of the parameter.

If $P(X;\theta)$ and $P(X;\theta')$ denote the probability densities of X (w.r.t. σ-finite measure v) for two values θ and θ' of the parameter, the difference in the distributions could be measured by some distance function, such as, Hellinger Distance [Rao, 1965], given by

$$\cos^{-1} \int \sqrt{P(X;\theta)P(X;\theta')}dv; \tag{7.1}$$

If we take $\theta' = \theta + \delta\theta$ and expand $P(X;\theta')$ by Taylor series, equation (7.1) reduces to

$$\cos^{-1} \int P(\theta) \left[1 - \frac{1}{8}\left\{\frac{P'(\theta)}{P(\theta)}\right\}^2 (\delta\theta)^2\right] dv = \cos^{-1}\left[1 - \frac{1}{8}J(\theta)(\delta\theta)^2\right] \tag{7.2}$$

where, $J(\theta)$ is Fisher Information, neglecting higher powers of $(\delta\theta)$.

Since $J(\theta)$ is positive, the distance (7.2) increases with increase in the value of $J(\theta)$ and is thus a measure of sensitivity of the random variable w.r.t. an infinitesimal change in the value of the parameter.
Hence, Fisher Information is an index of sensitivity for small changes in the value of the parameter.
In the multi-parameter case we get quadratic differential metric

$$\sum\sum J_{ij}\delta\theta_i\delta\theta_j$$

in place of $J(\theta)(\delta\theta)^2$ in (7.2), as the distance between the distributions of the random variable for the parameters $(\theta_1, \theta_2, \ldots)$ and $(\theta_1 + \delta\theta_1, \theta_2 + \delta\theta_2, \ldots)$.

The sensitivity of the random variable w.r.t. changes in the parameter may then be judged by examining the Fisher Information Matrix (**FIM**) as a whole.

Information value degrades over time and entropy value increases over time in general [Middleton, 1960].

For Gaussian Probability Distribution, Shannon Entropy is given by

$$H = \ln\sqrt{2\pi e}\sigma,$$

Where, Fisher Information is $FI = \frac{1}{\sigma^2}$

$\Rightarrow$ Fisher Information varies faster than Shannon Entropy.

Shannon's measure of information is valid for Hamiltonian Systems [Plastino et al., 1998] whereas Fisher information is valid for arbitrary dynamical systems. Fisher information is more general character than Shannon's or Kullback's. Fisher information is a local measure as it is sensitive to the local behaviour of the probability density. Fisher information decreases under coarse graining (i.e. less precision or quantization accuracy!) and it also decreases with time.

7.4 Information Induced by FIM and Controllability Gramian

As the system model and the initial mathematical treatment are identical with Chapter - 6, readers can refer to Chapter - 6 where we have already established the relation between controllability gramian, Shannon entropy and Shannon Information.

Referring to the result (6.11) of Chapter - 6 of this thesis we can say that

$$\det(\mathbf{W}_{c1}\mathbf{W}_{c2}^{-1}) = e^{2(\Delta H)} = e^{2I_{sh}} \tag{7.3}$$

where $\mathbf{W}_{c1}, \mathbf{W}_{c2}$ denote controllability gramians and H and I_{sh} denote the Shannon Entropy and Shannon Information respectively. Fisher Information is also a measure of the intrinsic accuracy of a distribution and in [Liu and Elia, 2006] it is shown that Mutual Information I_{sh} is related to the Fisher Information $\det(\mathbf{FIM})$ as follows :

$I_{sh} = \frac{1}{2}\ln\{\det(\mathbf{FIM})\}$

Therefore,

$$\det(\mathbf{FIM}) = e^{2I_{sh}} \tag{7.4}$$

Hence, using equations (7.4) and (7.3) we can conclude that

$$\det(\mathbf{FIM}) = \det(\mathbf{W}_{c1}\mathbf{W}_{c2}^{-1}) \tag{7.5}$$

Thus with the reduction of controllability gramian norm from $\mathbf{W}_{c1}$ to $\mathbf{W}_{c2}$ there is an increase in **FIM** and higher the reduction of controllability gramian, more is the Fisher information associated with. Since Cramer-Rao Bound implies that Fisher information is the reciprocal of the variance as its lower limit and controllability gramian is same as variance for a white Gaussian noise as input, we can conclude that our result is in agreement with the already established theories.
From control theory the control energy expression is given by

$$\epsilon_c^2 \equiv \mathbf{x}^T\mathbf{W}_c^{-1}\mathbf{x}$$

By controllability one means that with minimum energy the state of a system can be steered towards the target in a finite time. Minimizing the energy means minimizing $\mathbf{W}_c^{-1}$ i.e. maximizing $\mathbf{W}_c$ in the sense of a given matrix norm. The use of enriched Fisher information which has got a direct relation with variance and controllability gramian helps us to achieve the target state with greater ease and prudent transfer of information which is of great interest in communication and power constrained embedded systems. An important observation in the above energy expression (control-theoretic) is that the locus is an ellipsoid with axis direction given by eigenvector of $\mathbf{W}_c^{-1}$ and length proportional to eigenvalue. Smaller the ellipsoid is, more controllable the system is. Similar analogy can be drawn in information theory [Middleton, 1960] as well :

$$\Delta I(Y;X) = \frac{1}{2}(\Delta X)^T\mathbf{FIM}(\Delta X)$$

where superscript T denoting Transpose, I, X and Y being the mutual information, input and output processes respectively with Δ being the change in respective parameters.
It is shown in [Guo et al., 2005] that the derivative of the mutual information (nats) with respect to the signal-to-noise ratio (*snr*) is equal to half the Minimum Mean Square Error (MMSE), regardless of the input statistics i.e.

$$\frac{dI}{d\mathit{snr}}(\mathit{snr}) = \frac{1}{2}MMSE(\mathit{snr}) \tag{7.6}$$

This relationship holds for both scalar and vector signals, as well as for discrete-time and continuous-time non-causal MMSE estimation. MMSE reduces with *snr* and Shannon Information increases with *snr* for Gaussian and Binary Channel as well. Hence, by reducing quantization noises (in terms of Quantization Precision / Accuracy) in [Ben Gaid and Çela, 2006] at the time when the system approaches steady-state we can increase the *snr* and with increase in *snr* we can have increased Mutual information and thus less uncertainty. Here, we provide the theoretical justification of the simulation work done before in [Ben Gaid and Çela, 2006].

7.5 Conclusion

This chapter has addressed some new ideas (based on our work [Roy et al., 2009b]) concerning the relation between Control Theoretic parameter (Controllability Gramian) and **FIM** of Estimation Theory which, in turn, linked to Information Theory. Since the DES or Networked Embedded Control System (NECS) has communication constraints due to limited bandwidth or noises, we must have to adopt a policy of resource allocation (refer our work in Chapter - 8) which enhances the information transmitted.
It is demonstrated that the controllability gramian norm constitutes a metric of information-theoretic entropy with respect to the noises induced by quantization. In fact, physical interpretation of controllability Gramian is : "If the input to the system is white Gaussian noise, then controllability gramian is the covariance of the state". Reduction of these noises is equivalent to the design methods proposing a reduction of the controllability gramian norm which is related to the Fisher information matrix as well. Motivated by this idea we have chosen the Degree of Controllability/Observability as metric for optimal integrated Scheduling and Control for NCS which is illustrated in our work in Chapter - 8.

8

Degree of Reachability / Observability as a Metric for Optimal Integrated Control and Scheduling of Networked Control Systems

8.1 Introduction

This Chapter addresses the problem of the on-line or dynamic optimal control and scheduling of Networked Control System (NCS) over limited bandwidth deterministic network based on our paper [Çela et al., 2010]. The solution is obtained by decomposing the overall problem in a twofold one as follows :

1. The first level solution aims at obtaining the off-line or static scheduling function of control signals based on system dynamic characteristics, induced periodicity of scheduling signals, communication constraints and the maximization of the *Degree of Reachability / Observability* [Antoulas, 2005]. A static feedback control law design based on the periodicity of scheduling function ensures the stability and steady state performances.

2. The second level solution is based on the periodicity of the static scheduling function obtained from the first level problem solution and the minimization of control energy transfer of the current state to the final or desired state of the system. We propose a control law which enhances the performances of the system in real-time (reducing the time of response to perturbations), by modification of the periodic

scheduling and the value of control signals given by the first phase based on the minimization of the energy needed to transfer the state of the system to the desired one.

Advantages of the proposed method are twofold :

1. It reduces the complexity of the optimal static scheduling of control and sensor signal.

2. The real-time control signal update is done not only based on the scheduling function but also on the feedforward control supplementary signal which minimizes the transfer energy.

These results are proved mathematically and a design methodology is given in the case of NCS composed of interconnected LTI systems.

The use of a communication network in NCS considerably reduces the wiring, especially when the control application contains a large number of nodes. Consequently, the maintenance and the deployment become easier. The automotive industry represents an important application field of NCS. The design process of automotive applications is subject to conflicting criteria including cost, safety, reliability and performance. Consequently, these networked embedded control applications are developed on target platforms with limited CPU speed, memory and communication bandwidth. When computation or communication resources are limited, they have to be exploited as efficiently as possible. In distributed information processing systems, the scheduling function is responsible for the allocation of computation or communication resources. Consequently, using efficiently the computation or communication resources amounts to designing appropriate scheduling algorithms. Recently, it has been shown that a more efficient use of the computation or communication resources could be obtained if the problems of control and resource allocations are considered jointly [Årzén et al., 2000, Cervin et al., 2002, Martí et al., 2004, Lee et al., 2005].

In this chapter, the problem of the distributed control over deterministic real-time networks is addressed. We have to design the control laws and at the same time decide on the control / sensor signals scheduling. The solution is obtained by decomposing this problem in two distinct sub-problems. The first sub-problem formulation aims at obtaining a static or off-line scheduling of control / sensor signals based on the structural properties of the system and the communication constrains. The corresponding optimal control and sensor signals scheduling problem is formulated. The solution obtained consists of a periodic scheduling of a subset of control / sensor signals to send from controller / sensor node to the actuator / controller node in each sampling period. The second sub-problem modifies the periodic or static control scheduling function based on the information given by the state of the system. In this manner, not only the feedback control signals but also the scheduling function is modified with respect to the system state.

We model the NCS as a family of periodic systems induced by the periodic scheduling function and its *linear arrangement* [2]. To obtain the off-line or static scheduling function we use the *Degree of Reachability / Observability* as a metric. Two optimization problems are formulated and resolved based on these concepts and operating over the related gramians (refer Appendix - A) and intrinsic characteristics of the induced periodic systems. As runtime decision making metric for control signals update, we use the energy needed to transfer the system from a given initial state to a target state. It measures the capacity of scheduled control signals to modify the state of the system. Less is the transfer-energy needed, better is the degree of reachability associated with the scheduled control signals. It is clear that the interest of the control signals update is related to the influence of perturbation on the systems during their life cycles and the necessity to consider them via the system state information as well as scheduling of control signals or communication resource allocation. In the absence of perturbations, the on-line scheduling function is identical to the periodic static one.

8.2 Problem Formulation

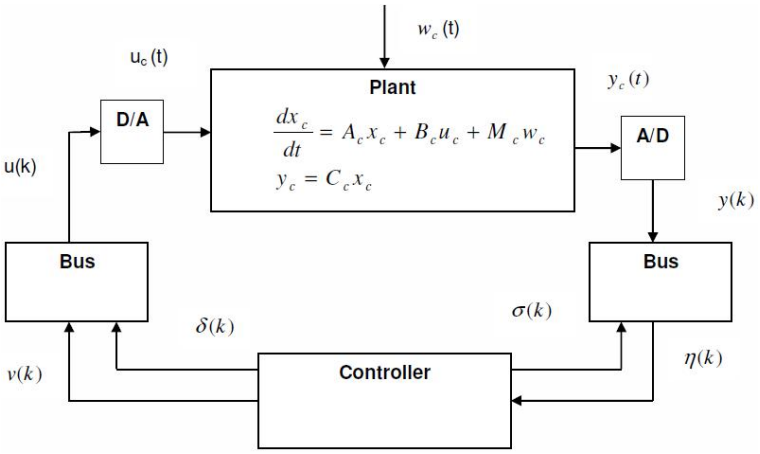


Figure 8.1: Schematic Representation of a Resource-Constrained System

Consider the continuous-time LTI plant described by :

$$\begin{aligned} \dot{x}_c(t) &= A_c x_c(t) + B_c u_c(t) \\ y_c &= C_c x_c(t) \end{aligned} \tag{8.1}$$

where $x_c(t) \in \mathbb{R}^n$, $u_c(t) \in \mathbb{R}^m$ and $y_c(t) \in \mathbb{R}^p$. The transpose of matrix A will be denoted by A'. The plant contains m distinct actuators and p distinct sensors which are spatially distributed. The sensors and actuators are connected to the main controller through a

[2] All linear arrangements of a family of three objects $\{A_1A_2A_3\}$ are $\{A_1A_2A_3\}, \{A_2A_3A_1\}, \{A_3A_1A_2\}$

limited bandwidth communication network. The network bandwidth is limited in the sense that it can carry at most b_r measures and b_w controls.

In order to derive a digital control law, a discrete-time representation of system (8.1) in the sampling period T_s is considered :

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \tag{8.2}$$

where $x(k) = x_c(kT_s)$, $u(k) = u_c(kT_s)$ and $y(k) = y_c(kT_s)$.

In [Hristu, 1999, Ben Gaid et al., 2006a], the description of the constraints affecting the transmission of the control commands to the actuators was done using the notion of *Scheduling Function*. Communication constraints may be formally described by introducing two vectors of Booleans $\sigma(k) \in \{0,1\}^{b_r}$ and $\delta(k) \in \{0,1\}^{b_w}$, defined for each sampling instant k.

Definition 8.1. The vector $\sigma(k)$ defined by

$$\begin{cases} \sigma_i(k) = 1 & \text{if } y_i(k) \text{ is read by the controller at instant } k, \\ \sigma_i(k) = 0 & \text{otherwise.} \end{cases}$$

is called sensors-to-controller scheduling vector at instant k.

Definition 8.2. The vector $\delta(k)$ defined by

$$\begin{cases} \delta_i(k) = 1 & \text{is } u_i(k) \text{ updated at instant } k, \\ \delta_i(k) = 0 & \text{otherwise.} \end{cases}$$

is called controller-to-actuators scheduling vector at instant k.

The vector $\sigma(k)$ indicates the sensor signals that the controller may read at instant k. In a similar way, the $\delta(k)$ indicates the control inputs to the plant that the controller may update at instant k.

The plant (8.1), the analog-to-digital (A/D) and digital-to-analog (D/A) converters, the communication bus and the controller are schematically depicted in Fig. (8.1). In this figure, $\eta(k) \in \mathbb{R}^{b_r}$ represents the vector of partial measurements that the controller receives (through the communication bus) at the sampling period k. In a similar way, vector $v(k) \in \mathbb{R}^{b_w}$ represents the vector of partial control commands that the controller may send to the actuators (through the limited bandwidth communication bus) at the sampling period k. Blocks D/A and A/D respectively represent the digital-to-analog and analog-to-digital converters. The controller may also assign the values of the sensors-to-controller scheduling vector ($\sigma(k)$) as well as the controller-to-actuators scheduling vector ($\delta(k)$). Let $[D_\delta(k)]_{1\leq i\leq m, 1\leq j\leq b_w}$ be the matrix defined by

$$\begin{cases} [D_\delta(k)]_{ij} = 1 & \text{if } \delta_i(k) = 1 \text{ and } \sum_{l=1}^{i} \delta_l(k) = j, \\ [D_\delta(k)]_{ij} = 0 & \text{otherwise,} \end{cases}$$

and

$$E_\delta(k) = \begin{bmatrix} 1-\delta_1(k) & & \\ & \ddots & \\ & & 1-\delta_m(k) \end{bmatrix},$$

then

$$u(k) = D_\delta(k)v(k) + E_\delta(k)u(k-1). \tag{8.3}$$

Let $[M_\sigma(k)]_{1\le i\le b_r, 1\le j\le p}$ the matrix defined by

$$\begin{cases} [M_\sigma(k)]_{ij} = 1 & \text{if } \sigma_j(k) = 1 \text{ and } \sum_{l=1}^{j} \sigma_l(k) = i, \\ [M_\sigma(k)]_{ij} = 0 & \text{otherwise.} \end{cases}$$

then we have the relation

$$\eta(k) = M_\sigma(k)y(k). \tag{8.4}$$

8.2.1 State Representation of Resource-Constrained Systems

For predefined scheduling vectors $\sigma(k)$ and $\delta(k)$, a resource constrained system $\mathcal{S}$ (as defined below) may be viewed, between its input $v(k)$ and its output $\eta(k)$, as a linear time-varying system. Based on the previous definitions, and noting

$$\chi(k) = u(k-1)$$

and

$$\tilde{x}(k) = \begin{bmatrix} x(k) \\ \chi(k) \end{bmatrix},$$

the linear sampled-data and time-varying model of system $\mathcal{S}$ is given by

$$\begin{aligned} \tilde{x}(k+1) &= \tilde{A}(k)\tilde{x}(k) + \tilde{B}(k)v(k) \\ \eta(k) &= \tilde{C}(k)\tilde{x}(k), \end{aligned} \tag{8.5}$$

where

$$\tilde{A}(k) = \begin{bmatrix} A & BE_\delta(k) \\ 0_{m,n} & E_\delta(k) \end{bmatrix},$$

$$\tilde{B}(k) = \begin{bmatrix} BD_\delta(k) \\ D_\delta(k) \end{bmatrix},$$

and

$$\tilde{C}(k) = M_\sigma(k) \begin{bmatrix} C & 0_{p,m} \end{bmatrix}.$$

Remark 1: If we consider the model of actuators signals updating given in ([Zhang and Hristu-Varsakelis, 2006, Hristu, 2007]), the model (8.5) is variable on $\tilde{B}(k)$ and $\tilde{C}(k)$ and constant in $\tilde{A}$. This makes the analysis and design easier but not necessarily implies that better performances would be obtained.

Remark 2: The $\tilde{A}$ matrix is dependent only on the scheduling function of control signals $\delta(k)$.

8.2.2 Lifting of Periodic Systems

Consider the optimal T_δ-periodic off-line scheduling function or communication sequence $\delta^{T_\delta - 1}$ of control signals ([3]) defined by :

$$\delta^{T_\delta - 1} = \{\delta(0), \cdots, \delta(T_\delta - 1)\}$$

and verifying $\delta(k + T_\delta) = \delta(k)$. Without loss of generality we may assume furthermore that $\delta \in \mathcal{S}^c$, where $\mathcal{S}^c$ is the set of periodic communication sequences that guarantees the reachability of system (8.5)([Zhang and Hristu-Varsakelis, 2006, Ionete and Çela, 2006, Hristu, 2007]). The periodicity of the communication sequence induces the periodicity of the resource-constrained system $\mathcal{S}$ (8.5). As a result, matrices $\tilde{A}(k)$ and $\tilde{B}(k)$ satisfy $\tilde{A}(k + T_\delta) = \tilde{A}(k)$, $\tilde{B}(k + T_\delta) = \tilde{B}(k)$.

Definition 8.3. The ι^{th} *linear arrangement* of an ordered set of T_δ elements is obtained by its ι ($\iota \in \{0, 1, \ldots T_\delta - 1\}$) circular right/left shifting. The number of right/left shifting ι is called period index value[4]

As illustrated in [Bittanti et al., 1991, Ben Gaid et al., 2006a], a time invariant model of discrete periodic system (8.5) may be obtained using the lifting technique. The time invariant modeling may be seen as a down sampled representation of system (8.5) with periodicity T_δ called **hyperperiod**. In the following, we will give some basics of the lifting technique. Please refer to [Bittanti et al., 1991] for more details.

Let Φ be the transition matrix associated with the state matrix $\tilde{A}$. Φ is defined by

$$\begin{cases} \Phi(l, s) = \tilde{A}(l-1)\tilde{A}(l-2)\cdots\tilde{A}(s) & \text{if } l > s \\ \Phi(l, l) = I_{n+m}. \end{cases}$$

Let Γ the matrix defined for $s < l < s + T_\delta$ by

$$\Gamma(l, s) = [\Phi(l, s+1)\tilde{B}(s) \;\; \Phi(l, s+2)\tilde{B}(s+1) \;\; \cdots \;\; \cdots \;\; \Phi(l, l)\tilde{B}(l-1) \;\; \underbrace{0_{n+m,b} \;\; \cdots \;\; 0_{n+m,b}}_{T_\delta - l - s}]$$

[3] Control/Sensor scheduling function will be obtained from the solution of **Problem 8.1** and **Problem 8.2** defined in section (8.3)

[4] All linear arrangements of $\delta^2 = \{\delta(0), \delta(1), \delta(2)\}$ is $\{\{\delta(0), \delta(1), \delta(2)\}, \{\delta(2), \delta(0), \delta(1)\}, \{\delta(1), \delta(2), \delta(0)\}\}$

and for $s = l$ by

$$\Gamma(s,s) = [0_{n+m,b} \ \cdots \ 0_{n+m,b}] .$$

Let

$$\bar{x}_\iota(q) = \tilde{x}(\iota + qT_\delta)$$

and

$$\bar{v}_\iota(q) = \begin{bmatrix} v(\iota + qT_\delta) \\ \vdots \\ v(\iota + (q+1)T_\delta - 1) \end{bmatrix}$$

then for $0 \leq i \leq T_\delta$:

$$\tilde{x}(\iota + qT_\delta + i) = \Phi(\iota + i, \iota)\bar{x}_\iota(q) + \Gamma(\iota + i, \iota)\bar{v}_\iota(q).$$

Particularly, let

$$\bar{A}_\iota = \Phi(\iota + T_\delta, \iota)$$

and

$$\bar{B}_\iota = \Gamma(\iota + T_\delta, \iota)$$

So, we may write the following LTI state-space representation of the periodic system $\mathcal{S}$:

$$\bar{x}_\iota(q+1) = \bar{A}_\iota\bar{x}_\iota(q) + \bar{B}_\iota\bar{v}_\iota(q). \tag{8.6}$$

Remark 3: In the sequel we will assume that the scheduling functions of control and sensor signals, $\delta^{(T_\delta - 1)}$ and $\sigma^{(T_\delta - 1)}$ respectively, **have the same length** T_δ ($T_\delta = T_\sigma$). So we may construct the output equation $\bar{y}_\iota(q) = \bar{C}_\iota\bar{x}_\iota(q)$ and thus obtaining the periodic state-space equation :

$$\begin{aligned} \bar{x}_\iota(q+1) &= \bar{A}_\iota\bar{x}_\iota(q) + \bar{B}_\iota\bar{v}_\iota(q) \\ \bar{y}_\iota(q) &= \bar{C}_\iota\bar{x}_\iota(q) \end{aligned} \tag{8.7}$$

It is necessary to see more closely the importance of the scheduling function on the structure of these equations. Recall that

$$\bar{A}_\iota = \Phi(T_\delta + \iota, \iota) = \tilde{A}(T_\delta + \iota - 1)\tilde{A}(T_\delta + \iota - 2)\cdots\tilde{A}(\iota) \tag{8.8}$$

and

$$\bar{B}_\iota = \Gamma(\iota + T_\delta, \iota) \tag{8.9}$$

where $\iota \in \{0, \ldots, T_\delta - 1\}$.

It is clear that $\bar{A}_\iota$ and $\bar{B}_\iota$ are dependent[5] on the value of parameter ι which is the current period index value. So we have T_δ realization of the LTI state-space equations (8.7). We may see the set of T_δ LTI realizations as competing systems to obtain the communication resources based on the performances they may provide with respect to the current state. Each of these realizations is a periodic one with period T_δ dependent on the scheduling of control/sensor signals. Real-time optimal solutions are NP-hard problem so our control design problem may be split into two levels. The first level or the off-line design has to consider only the dynamic model characteristics of the systems and the network bandwidth constraints. It has to give us a periodic scheduling of control/sensor signals and the corresponding feedback control gains based on induced periodic system (8.7). The second level problem has to modify the real-time control signals to handle the perturbation affecting our systems based on the energy criteria consistent with the degree of controllability.
In the following subsection we will give the solution of the first level problem or off-line optimal scheduling design. For clarity of our presentation we will introduce at first the concept of degree of reachability and observability.

8.2.3 Reachability / Observability of NCS

As aforementioned control design of an NCS with bandwidth resource constraints needs to specify the structure of feedback control law and associated scheduling function $\delta(k), \sigma(k), k \in \{0, 1, \ldots, T_\delta - 1\}$ of the control and sensor signals. It is natural that we have to apply (send through the network) the control signals with greatest impact on the system relative to the current state and the target state. This physical fact is closely related to the *degree of reachability* ([Moore, 1981], [Antoulas, 2005]) or the *minimal energy control* needed to transfer the current state of the system x_t to the target state X_d. The dual concept is the *degree of observability* or the *observation energy* produced by the observed state x_0. Before giving the expression of these energies and control law design we recall some definitions and results concerning the reachability and observability and related gramians [Moore, 1981], [Antoulas, 2005], [Georges, 1995].

Definition 8.4 ([Rugh, 1996]). A linear discrete-time system is called l-step reachable (resp. l-step observable) if l is a positive integer such that the system is reachable (resp. observable) on $[i, i + l]$, for any $i \in N$.

Theorem 8.1 ([Zhang and Hristu-Varsakelis, 2006], [Ionete and Çela, 2006], [Hristu, 2007]). *If A is invertible and the pair (A, B) is reachable, then for any integer b_w such that $1 \leq b_w \leq m$, there exists integers $l, T > 0$ and a maximal T-periodic communication sequence of width m such that system (8.5) is l-step reachable.*

[5] In the model adopted in [Zhang and Hristu-Varsakelis, 2006, Hristu, 2007] only $\bar{B}_\iota$ is dependent on the value of index value ι

A similar result is obtained for the observability.

Theorem 8.2 ([Zhang and Hristu-Varsakelis, 2006], [Ionete and Çela, 2006]). *If A is invertible and the pair (A, C) is observable, then for any b_r such that $1 \leq b_r \leq p$, there exist integers $l, T > 0$ and a maximal T-periodic communication sequence of width p such that system (8.5) is l-step observable.*

Proposition 8.1 ([Zhang and Hristu-Varsakelis, 2006, Ionete and Çela, 2006]). *The time T_f needed for the system (8.2) to reach the target state x_f is bounded by $T_f = (n * \lceil \frac{n}{b} \rceil)$*[6]

Proposition 8.2 ([Brockett, 1970], [Antoulas, 2005]). *The reachability gramians $W_r(k)$ (where R_k being the controllability matrix) have the following properties :*

1. $W_r(k) = W_r^{'}(k) \geq 0$
2. *their columns span the reachability subspace, i.e.,*

$$Im(W_r(k)) = Im(R_k(A, B))$$

Corollary 1 : The system (8.2) is reachable if and only if $W_r(k)$ is positive definite for $k > n$ where n is the system order.

Proposition 8.3 ([Brockett, 1970], [Antoulas, 2005]). *The observability gramians $(W_o(k))$ (where $\mathcal{O}_k$ being the observability matrix) have the following properties :*

1. $W_o(k) = W_o^{'}(k) \geq 0$
2. *their columns span the observability subspace, i.e.,*

$$Im(W_o(k)) = Im(\mathcal{O}_k(A, C))$$

Corollary 2 : The system (8.2) is observable if and only if $W_o(k)$ is positive definite for $k > n$.

Definition 8.5 ([Brockett, 1970], [Antoulas, 2005], [Sontag, 1990]). The transient reachability function of the discrete-time system (8.2) is defined respectively as :

$$\begin{aligned} F_r &= \min_{(u(k))} \sum_{k=0}^{T_f} u^{'}(k)u(k) \\ x(k+1) &= Ax(k) + Bu(k) \\ \Rightarrow x(0) &= x_0, x(T_f) = x_d \end{aligned} \tag{8.10}$$

It is clear that the transient reachability function gives the minimal energy needed to transfer the system from the initial state x_0 to the desired one x_d in T_f sampling periods.

[6] In [Hristu, 2007] it is proved that under some conditions the order n of the system is an upper bound for the scheduling periodicity and the reachability horizon.

Lemma 8.1 ([Sontag, 1990], [Brockett, 1970], [Antoulas, 2005]). *The minimal energy required to transfer the initial state x_0 to the desired state x_d in T_f sampling periods or the transient reachability function is given by:*

$$F_r = (x_d - (A^{'})^{T_f} x_0)^{'} W_r^{-1}(T_f)(x_d - (A^{'})^{T_f} x_0) \tag{8.11}$$

and its optimal control solution is :

$$\overline{u}(k) = B^{'}(A^{'})^{(T_f-k)} W_r^{-1}(T_f)(x_d - (A^{'})^{T_f} x_0) \tag{8.12}$$

From this Lemma we have, simultaneously, the function which measures the *degree of controllability/reachability*[7] of our system and the open-loop optimal control signals realizing it. Naturally, for T_δ different state-space realizations (8.6), the *degree of reachability* (control energy needed to transfer the system from zero initial condition, x_0, to the desired one x_d in T_f sampling periods) and the associated control signals will be different. This will be the central point in the real-time update of scheduling function.

Less is this energy required, more important is actuator action on system state. Here, minimization of control energy required is the criterion. It will form our switching function which will modify at runtime the static periodic scheduling function by updating the control signals with the greatest *degree of controllability/reachability* of the system state. It is similar to the problem of optimal location of actuators and sensors as treated in [Georges, 1995] and operating on the augmented LTI systems (8.7). This optimal location problem is operated not only in space (b_w control signals update of m possible ones) but also in each sampling period.

Definition 8.6 ([Brockett, 1970], [Antoulas, 2005]). The transient observability function of the discrete-time system (8.2) is defined respectively as :

$$\begin{aligned} F_o &= \max_{(y(k))} \sum_{k=0}^{k=T_f} y(k)^T y(k) \\ x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \\ \Rightarrow x(0) &= x_0, u(k) \equiv 0, k \in \{0, 1, \ldots, (T_f - 1)\} \end{aligned} \tag{8.13}$$

The transient observability function represents the energy of the output signals relative to the initial state x_0 and gives the degree of its observability from the output signals. Here, maximization of observation energy is the criterion.

Lemma 8.2 ([Sontag, 1990], [Brockett, 1970], [Antoulas, 2005]). *The transient observability function is given by :*

$$F_o = x_0^T W_o(T_f) x_0 \tag{8.14}$$

[7] Supposing that the state matrix is invertible, the controllability and the reachability are equivalent properties of the system

Let note by $N = \mathcal{H}T_\delta$ the horizon of periodic system $\bar{A}_\iota, \bar{B}_\iota, \iota \in \{0, \ldots, (T_\delta - 1)\}$, with $\mathcal{H}$ a positive integer and $N \geq T_f$. So $\mathcal{H}$[8] measures the horizon of our study in number of hyperperiods (T_δ).

8.3 Optimal Off-Line Scheduling of Control / Sensor Signals

From Lemma 8.1 and Lemma 8.2 we can find clearly that the inverse of the reachability gramian constitutes the weighted matrix of optimal transfer energy criterion. It is dependent on the systems dynamics, the periodic scheduling of control signals and the initial states. Less is transfer energy required, greater the controllability we have on the system state via the scheduled control signals. If the initial state and the systems dynamics are given, the scheduling of control signals can change the value of this transfer energy. The same analysis may be done for the observability gramian and its dependence on the scheduling of measurement or sensed signals. Increasing the degree of reachability/controllability means that the quadratic form (A.19) has to be minimized over all periodic scheduling $\delta(k), (k = 1, \ldots, N, \ldots)$. Degree of Observability is optimized by choosing the scheduling which maximizes the quadratic form over all periodic scheduling $\sigma(l), l = 1(\ldots, N, \ldots)$.

Recall that $\delta^{T_\delta - 1} = (\delta(0), \cdots, \delta(T_\delta - 1))$ and $\sigma^{T_\sigma - 1} = (\sigma(0), \cdots, \sigma(T_\sigma - 1))$ represent the periodic scheduling functions of control and sensor signals respectively. As assumed previously these two periodic scheduling functions have the same length. Denote by $\mathcal{S}^c$ and $\mathcal{S}^s$ the sets of all periodic scheduling functions of length $T_\delta - 1$ and $T_\sigma - 1$ respectively (c for control and s for sensor/measurement). So, we can arrive at the optimal scheduling of control and measurement signals by the solution of the two following optimization problems :

Problem 8.1. Optimal Degree of Reachability

$$Max_{\{\delta^{T_\delta - 1} \in \mathcal{S}^c\}} \quad \{\Lambda_{min}(W_r(T_f, \delta^{T_\delta - 1}))\} \tag{8.15}$$

$$\bar{x}_\iota(q+1) = \bar{A}_\iota(\delta^{T_\delta - 1})\bar{x}_\iota(q) + \bar{B}_\iota(\delta^{T_\delta - 1})\bar{v}_{\iota(q)}$$

$$\bar{v}_\iota(q) = -\tilde{K}(\iota)\bar{x}_\iota(q) + \bar{v}_{s\iota}(q)$$

Since reachability and observability are dual concepts we treat the *Problem - 8.2* like the one in *Problem - 8.1* but in Dual system. Substituting $\overline{A} = \overline{A}', \overline{B} = \overline{C}', \overline{C} = \overline{B}'$ in *Problem - 8.1* we can formulate as follows :

[8] the worst case scenario, $b = 1$, induces that $\mathcal{H} = n$, where n is the order of our system. **In the sequel we will consider** $\mathcal{H} = n$

Problem 8.2. Optimal Degree of Observability

$$Max_{\{\sigma^{T_\sigma-1}\in\mathcal{S}^o\}} \quad \{\Lambda_{min}(W_o(T_f,\sigma^{T_\sigma-1}))\} \tag{8.16}$$

$$\bar{x}'_\iota(q+1) = \bar{A}'_\iota(\sigma^{T_\sigma-1})\bar{x}'_\iota(q) + \bar{C}'_\iota(\sigma^{T_\sigma-1})\bar{v}'_\iota(q)$$

$$\bar{y}'_\iota(q) = \bar{B}'_\iota(\sigma^{T_\sigma-1})\bar{x}'_\iota(q)$$

where $\Lambda_{min}(W_r)$ and $\Lambda_{min}(W_o)$ are the respective eigenvalues of matrices W_r and W_o with minimal moduli.

These two problems are integer programming problems. The solution complexity depends on the system order and the bandwidth constraints. An algorithm based on incremental values of $\delta^{T_\delta-1}$, $\sigma^{T_\sigma-1}$ and detection of its induced periodicity on a given maximal length give satisfying results and reduce significantly the complexity with respect to the method proposed in [Ben Gaid et al., 2006a].

Proposition 8.4. *The periodic scheduling function obtained from the solution of* Problem - 8.1 *preserves the reachability and the observability of the original discrete system (8.2).*

To prove this proposition, first we have to check the existence of periodic scheduling function preserving these properties and secondly if this scheduling function is in the set of the feasible solutions of *Problem - 8.1* and *Problem - 8.2*. The first part is proved by the theorems (8.1) and (8.2). The second is clear (proved) from the formulation of the *Problem - 8.1* and *Problem - 8.2* which finds the optimal periodic scheduling function minimizing the energy transfer or observation criteria[9] which is directly related to the degree of reachability of our system [Moore, 1981] [Ben Gaid et al., 2006b].

8.3.1 An Example

System $S_1 = \{A_1, B_1, C_1, D_1\}$

System $S_2 = \{A_2, B_2, C_2, D_2\}$

System $S_3 = \{A_3, B_3, C_3, D_3\}$

Where :

$$A_1 = \begin{bmatrix} 0.5 & 0 \\ 1 & 0 \end{bmatrix}; A_2 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}; A_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix};$$

[9] As stated in [Hristu, 2007], [Ben Gaid, 2006] the length of T_δ obtained by the solution of *Problem - 8.1* is much shorter then that obtained in [Zhang and Hristu-Varsakelis, 2006], [Ionete and Çela, 2006], [Hristu, 2007]

$\mathrm{B}_1 = B_2 = B_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; C_1 = C_2 = C_3 = \begin{bmatrix} 0 & 1 \end{bmatrix}; D_1 = D_2 = D_3 = \begin{bmatrix} 0 \end{bmatrix};$

The open-loop eigenvalues of the systems are respectively : $\{0.5 \;\; 0\}, \{-1 \;\; 0\}$ and $\{1 \;\; 0\}$ which means that the second system is the only open-loop stable system. So, naturally, the order of importance of the resource allocation is given by the ordered set $\{S_3, S_1, S_2\}$. This order is consistent with the information rate definition given in [Tatikonda and Mitter, 2004, Zang and Iglesias, 2003, Mehta et al., 2006, Wu and Jonckheere, 1992, Iglesias, 2001, Iglesias, 2002, Okano et al., 2008] and also in chapters 2 and 5 of this thesis.

8.3.2 Results Obtained

For the case when we have only one system input/output update at each sampling instant of time :

1. Traces of the Inverse of Reachability Gramians (first column) with respect to the Scheduling Sequences (second column) are shown in Table 8.1.

Table 8.1: Traces of Inverse Reachability Gramians W_r^{-1} w.r.t. Scheduling Sequences $\mathcal{S}_i$

$tr(W_r^{-1})$ $(\times 10^4)$	Sequence $\mathcal{S}_i$	Sequence No.
0.0115	$3 \rightarrow 2 \rightarrow 3 \rightarrow 1$	1
0.0096	$1 \rightarrow 3 \rightarrow 2 \rightarrow 3$	2
0.0088	$3 \rightarrow 1 \rightarrow 3 \rightarrow 2$	3
0.0160	$2 \rightarrow 3 \rightarrow 1 \rightarrow 3$	4
0.0094	$3 \rightarrow 3 \rightarrow 2 \rightarrow 1$	5
0.0086	$3 \rightarrow 3 \rightarrow 1 \rightarrow 2$	6
0.0111	$1 \rightarrow 2 \rightarrow 3 \rightarrow 3$	7
0.0157	$2 \rightarrow 1 \rightarrow 3 \rightarrow 3$	8
0.0162	$2 \rightarrow 1 \rightarrow 3 \rightarrow 1$	9
0.0117	$1 \rightarrow 2 \rightarrow 1 \rightarrow 3$	10

2. Determinant of the Inverse Reachability Gramian (first column) with respect to the Scheduling Sequences (second column) is as shown below :

$$\det W_r^{-1} = 10^6 \times [1.8429 \;\; 0.7493 \;\; 0.3046 \;\; 4.5329 \;\; 0.2490 \;\; 0.1012 \;\; 0.6125 \;\; 1.5065 \;\; 1.2736 \;\; 0.9435]$$

3. Let us denote the eigenvalues, maximal moduli of eigenvalues and minimal moduli of eigenvalues of Inverse Reachability Gramians by $\Lambda_i, \max|\Lambda_i(W_r^{-1})|$ and $\min|\Lambda_i(W_r^{-1})|$ respectively. Representing the differences between Max-Min of Moduli of eigenvalue by $\Delta\Lambda_{maxmin}$. The maximal moduli and the minimal moduli of the eigenvalues of the Inverse Reachability Gramians (first column and second column) with respect to the Scheduling Sequences (fourth column) are as presented in Table (8.2).

Table 8.2: Maxima-Minima Moduli of Eigenvalues of Inverse Reachability Gramians W_r^{-1} w.r.t. Scheduling Sequences $\mathcal{S}_i$

$\max\|\Lambda_i(W_r^{-1})\|$	$\min\|\Lambda_i(W_r^{-1})\|$	$\Delta\Lambda_{maxmin}$	Sequence $\mathcal{S}_i$	Sequence No.
0.5637	0.0192	0.5445	$3\rightarrow 2\rightarrow 3\rightarrow 1$	1
1.0694	0.0248	1.0446	$1\rightarrow 3\rightarrow 2\rightarrow 3$	2
0.7101	0.0231	0.6870	$3\rightarrow 1\rightarrow 3\rightarrow 2$	3
0.5632	0.0101	0.5531	$2\rightarrow 3\rightarrow 1\rightarrow 3$	4
1.4147	0.0248	1.3899	$3\rightarrow 3\rightarrow 2\rightarrow 1$	5
1.4147	0.0231	1.3916	$3\rightarrow 3\rightarrow 1\rightarrow 2$	6
1.0694	0.0192	1.0502	$1\rightarrow 2\rightarrow 3\rightarrow 3$	7
0.7101	0.0101	0.7000	$2\rightarrow 1\rightarrow 3\rightarrow 3$	8
1.8168	0.0101	1.8067	$2\rightarrow 1\rightarrow 3\rightarrow 1$	9
1.5711	0.0192	1.5519	$1\rightarrow 2\rightarrow 1\rightarrow 3$	10

4. For a Dual System i.e. ($A = A^T, B = C^T, C = B^T$ and $D = D^T$) we can treat reachability and observability as dual to each other and hence Inverse of Observability Gramian $W_0^{-1} = \bar{W}_r^{-1}$, where $\bar{W}_r^{-1}$ is the inverse of the reachability gramian of the dual system.
Hence,Traces of the Inverses of Observability Gramians (first column) with respect to the Scheduling Sequences (second column) are as given in Table (8.3).

5. Determinant of the Inverse of Observability Gramian (first column) with respect to the Scheduling Sequences (second column) is as given below :
$\det W_o^{-1} = 10^6\times[1.8429\ \ 0.7493\ \ 0.3046\ \ 4.5329\ \ 0.2490\ \ 0.1012\ \ 0.6125\ \ 1.5065\ \ 1.2736\ \ 0.9435]$

6. The maximal moduli and minimal moduli of the eigenvalues of the inverse of Observability Gramians (first column and second column) along with their differences (third column) with respect to the scheduling sequences (fourth column) are

Table 8.3: Traces of Inverse Observability Gramians W_0^{-1} w.r.t. Scheduling Sequences $\mathcal{S}_i$

$tr(W_0^{-1})$ $(\times 10^4)$	Sequence $\mathcal{S}_i$	Sequence No.
0.0185	$3 \rightarrow 2 \rightarrow 3 \rightarrow 1$	1
0.0124	$1 \rightarrow 3 \rightarrow 2 \rightarrow 3$	2
0.0077	$3 \rightarrow 1 \rightarrow 3 \rightarrow 2$	3
0.0326	$2 \rightarrow 3 \rightarrow 1 \rightarrow 3$	4
0.0111	$3 \rightarrow 3 \rightarrow 2 \rightarrow 1$	5
0.0073	$3 \rightarrow 3 \rightarrow 1 \rightarrow 2$	6
0.0191	$1 \rightarrow 2 \rightarrow 3 \rightarrow 3$	7
0.0324	$2 \rightarrow 1 \rightarrow 3 \rightarrow 3$	8
0.0322	$2 \rightarrow 1 \rightarrow 3 \rightarrow 1$	9
0.0193	$1 \rightarrow 2 \rightarrow 1 \rightarrow 3$	10

as given in Table (8.4).

Table 8.4: Maxima-Minima Moduli of Eigenvalues of Inverse Observability Gramians W_o^{-1} w.r.t. Scheduling Sequences $\mathcal{S}_i$

$\max \lvert\Lambda_i (W_o^{-1})\rvert$	$\min \lvert\Lambda_i (W_o^{-1})\rvert$	$\Delta\Lambda_{maxmin}$	Sequence $\mathcal{S}_i$	Sequence No.
0.8852	0.0078	0.8774	$3 \rightarrow 2 \rightarrow 3 \rightarrow 1$	1
0.7698	0.0162	0.7537	$1 \rightarrow 3 \rightarrow 2 \rightarrow 3$	2
0.6866	0.0321	0.6544	$3 \rightarrow 1 \rightarrow 3 \rightarrow 2$	3
0.7471	0.0039	0.7433	$2 \rightarrow 3 \rightarrow 1 \rightarrow 3$	4
0.8852	0.0162	0.8691	$3 \rightarrow 3 \rightarrow 2 \rightarrow 1$	5
0.8651	0.0321	0.8330	$3 \rightarrow 3 \rightarrow 1 \rightarrow 2$	6
0.8994	0.0078	0.8916	$1 \rightarrow 2 \rightarrow 3 \rightarrow 3$	7
0.8994	0.0039	0.8956	$2 \rightarrow 1 \rightarrow 3 \rightarrow 3$	8
3.4541	0.0039	3.4502	$2 \rightarrow 1 \rightarrow 3 \rightarrow 1$	9
3.9048	0.0078	3.8969	$1 \rightarrow 2 \rightarrow 1 \rightarrow 3$	10

The values in red colour correspond to the sequence which is the best one for the criteria (trace, determinant and the min / max of eigenvalues of the inverse of reachability and observability gramians) used for the inverse of reachability and observability gramians. The blue ones are the second best values and the green ones correspond to the third best values. It is shown here that a combination of these criteria can make a reliable metric selection for the best sequence of the message. So, we may say that the Sequence No. 3 is the best one here.

8.4 Optimal On-Line Control and Scheduling of Control Signals

Suppose that the system (8.6) given by :

$$\bar{x}_\iota(q+1) = \bar{A}_\iota \bar{x}_\iota(q) + \bar{B}_\iota \bar{v}_\iota(q)$$

is preliminarily stabilized by a controller [Bittanti et al., 1991] :

$$\bar{v}_\iota(q) = -\tilde{K}(\iota)\bar{x}_\iota(q) + \bar{v}_{s\iota}(q) \tag{8.17}$$

From (8.6) and (8.17) we obtain the stabilized linear system :

$$\bar{x}_\iota(q+1) = \bar{A}_{s\iota}\bar{x}_\iota(q) + \bar{B}_\iota \bar{v}_{s\iota}(q). \tag{8.18}$$

where $\bar{A}_{s\iota} = \bar{A}_\iota - \tilde{K}(\iota)\bar{B}_\iota$.

Problem 8.3. Given an initial state $x(0)$, a final time N, a target state $X_d \equiv 0$ and the static periodic scheduling function $\delta(k), k \in 0, \ldots, T_\delta - 1$, find the optimal real-time scheduling function $\delta^*(k), (k \in \{1, \ldots, N, \ldots\})$ and the optimal control sequence $v^N = [v(0), v(1), \ldots, v(N)]$ which resolves the optimization problem as follows :

$$F_r(\iota^*, \bar{v}^*_{s\iota}) \quad = \quad \min_{\{\bar{v}_{s\iota}(q))\}} \sum_{q=0}^{q=\mathcal{H}} \bar{v}'_{s\iota}(q)\bar{v}_{s\iota(q)} \tag{8.19}$$

$\Rightarrow (\overline{x}_\iota(0) = x(0), \overline{x}_\iota(\mathcal{H}) = X_d)$,

where, $\iota \in \{0, \ldots, (T_\delta - 1)\}$

$\bar{x}_\iota(q+1) = \bar{A}_{s\iota}\bar{x}_\iota(q) + \bar{B}_\iota \bar{v}_{s\iota}(q), \iota \in \{0, \ldots, T_\delta - 1\}$

Remark 4: The *Problem - 8.3* has two types of optimization variables. The first one concerns with the scheduling function which is a periodic one, proper to each periodic system $\bar{A}_{s\iota}, \bar{B}_\iota, (\iota \in \{0, \ldots, T_\delta - 1\})$. Second optimization parameter is the control value(s) calculated for all the horizons $q = 0$ to $q = \mathcal{H} - 1$.

The following theorem gives us the solution of the *Problem - 8.3* :

Theorem 8.3. *The optimal solution of the optimization problem :*

$$\begin{aligned} F_r(\iota^*, \bar{v}^*_{s\iota}) &= \min_{\{\bar{v}_{s\iota}(q),\iota\}} \sum_{q=0}^{q=\mathcal{H}} \bar{v}'_{s\iota}(k)\bar{v}_{s\iota}(q) \\ \Rightarrow \overline{x}_\iota(0) &= x(0), \overline{x}_\iota(\mathcal{H}) = X_d, \end{aligned} \tag{8.20}$$

$$\bar{x}_\iota(q+1) = \bar{A}_\iota \bar{x}_\iota(q) + \bar{B}_\iota \bar{v}_\iota(q), \iota \in \{0, \ldots, T_\delta - 1\}$$

is given by control law :

$$\bar{v}^*_{s\iota^*}(q) = \bar{B}'_{\iota^*}(\bar{A}'_{s\iota^*})^{(\mathcal{H}-q)} W_r^{-1}(\iota^*, \mathcal{H})(X_d - (\bar{A}'_{s\iota^*})^{\mathcal{H}} x(0)) \tag{8.21}$$

where, $q = \{0, 1, \ldots, \mathcal{H}\}$ which ensures the asymptotic convergence to the target states in N sampling periods.

Proof. For each value of $\iota \in \{0, \ldots, T_\delta - 1\}$ the optimality of control value given by (8.21) is the solution of classical optimal control energy problem which transfers the system from the initial state $x(0)$ to the final state X_d. So it ensures the convergence of the initial state to the target or final state in N sampling period or $\mathcal{H}$ (under sampling) periods. Its structure is :

$$\bar{v}^*_{s\iota^*} = [\bar{v}^*_{s\iota^*}(0, \iota^*), \bar{v}^*_{s\iota^*}(1, \iota^*), \ldots, \bar{v}^*_{s\iota^*}(\mathcal{H} - 1, \iota^*), 0, 0 \ldots] \tag{8.22}$$

Recall that the control structure applied to our system is given by (see (8.17))

$$\bar{v}_\iota(q) = -\tilde{K}(\iota)\bar{x}_\iota(q) + \bar{v}_{s\iota}(q)$$

From (8.22) we see that the values of $\bar{v}^*_{s\iota^*}$ are equal to zero when the state of the system hits the target state or after N sampling periods. So the only control applied beyond N sampling periods is the asymptotic stabilizing one with respect to the target state (in our case target state is zero state) given by :

$$\bar{v}_\iota(q) = -\tilde{K}(\iota)\bar{x}_\iota(q) + 0 \tag{8.23}$$

From (8.17) and for each index value of $\iota \in \{0, \ldots, T_\delta - 1\})$ the control law (8.23) is asymptotically stabilizing one.

Minimal energy criterion when applied over whole family of systems depends on the index value ι ($\iota \in \{0, \ldots, T_\delta - 1\}$) which ensures optimality of the degree of reachability / controllability. □

The optimal control law $\bar{v}^*_{s\iota^*}(q)$ obtained by the Theorem 8.3 is given by a set of N control signals and the associated optimal scheduling functions from the initial state x_0 to the target state x_d. The new control signal $\bar{v}^*_{s\iota^*}(q)$ which is dependent on initial state $x(0)$ or the disturbance effect (and in general of the final state x_d) corresponds to open-loop control signals with respect to system (8.18). The optimal periodic scheduling function $\delta^*(k), k \in \{0, \ldots, (T_\delta - 1)\}$ given by its optimal index value ι^* and its periodic

property is denoted by $\delta^*(\iota^*)$. It modifies the application order of periodic feedback control signals obtained from (8.17). Practically, it only chooses to update b_w control signals in the order given by $\delta^*(\iota^*)$. We note that this order modification is done only once during the transfer of the initial state x_0 to the target state x_d.

To enhance the performance, we can modify the value of scheduling function $\delta^*(\iota^*)$ in each period during the transfer of the system from the initial state to the target state. This means that we have to integrate the information of the current state and formulate a Receding Horizon Control (refer Appendix - B) problem. This implies that only the current value of optimal control $\bar{v}_{\iota^*(q)}$ given by (8.17) will be applied. The Receding Horizon formulation and its solution is given in the following Theorem :

Theorem 8.4. *The optimal solution of the optimization problem :*

$$\begin{aligned} F_r(\iota^*, \bar{v}^*_{s\iota}(q)) &= \min_{(\bar{v}_{s\iota}(q),\iota)} \{ \sum_{q=l}^{q=\mathcal{H}+l} \bar{v}'_{s\iota}(q)\bar{v}_{s\iota}(q)\} \\ &\Rightarrow \ (\bar{x}_\iota(l) = x(l), \bar{x}_\iota(\mathcal{H}+l) = X_d) \\ \bar{x}_\iota(q+1) &= \bar{A}_{s\iota}\bar{x}_\iota(q) + \bar{B}_\iota\bar{v}_{s\iota}(q), \iota \in \{0, \ldots, (T_\delta - 1)\} \end{aligned} \tag{8.24}$$

is given by control law :

$$\bar{v}^*_{s\iota^*}(l, \iota^*(l)) = \bar{B}'_{\iota^*}(\bar{A}'_{s\iota^*})^{((\mathcal{H}+l)-l)}W_r^{-1}(\iota^*(l), \mathcal{H})(X_d - (\bar{A}'_{s\iota^*})^{((\mathcal{H}+l)-l)}x(l)) \tag{8.25}$$

which ensures the asymptotic convergence to the target states.

Proof. Let us choose the Lyapunov function :

$$\mathcal{V}(\iota^*(l), \bar{x}_\iota(l)) = (X_d - (\bar{A}'_{s\iota^*})^{\mathcal{H}}x(l))'W_r^{-1}(\iota^*(l), \mathcal{H})(X_d - (\bar{A}'_{s\iota^*})^{\mathcal{H}}x(l)) \tag{8.26}$$

We may see that the expression of Lyapunov function is the control energy necessary to transfer the state of the system, for each period l, from current state $x(l)$ to the final state X_d. Let us express the difference between the two successive values of the Lyapunov function in the following. It will be given in function of optimal control variables $\bar{v}^*_{s\iota^*}(l, \iota^*(l))$ calculated for two successive instances l and $l+1$ and the respective optimal index values, $\iota^*(l)$ and $\iota^*(l+1)$.

$$\begin{aligned} &\mathcal{V}(\iota^*(l+1), \bar{x}_{\iota^*}(l+1)) - \mathcal{V}(\iota^*(l), \bar{x}_{\iota^*}(l)) \\ &= \sum_{q=l+1}^{q=\mathcal{H}+l+1} \bar{v}^*_{s\iota^*}(q, \iota^*(l+1))'\bar{v}^*_{s\iota^*}(q, \iota^*(l+1)) - \sum_{q=l}^{q=\mathcal{H}+l} \bar{v}^*_{s\iota^*}(q, \iota^*(l))'\bar{v}^{*l}_{s\iota^*}(q, \iota^*(l)) \\ &= \textstyle\sum_{q=l+1}^{q=\mathcal{H}+l+1} \bar{v}^*_{s\iota^*}(q, \iota^*(l+1))'\bar{v}^*_{s\iota^*}(q, \iota^*(l+1)) \end{aligned} \tag{8.27}$$

$-\sum_{q=l+1}^{q=\mathcal{H}+l} \bar{v}^*_{s\iota^*}(q,\iota^*(l))'\bar{v}^*_{s\iota^*}(q,\iota^*(l)) - \bar{v}^*_{s\iota^*}(l,\iota^*(l))'\bar{v}^*_{s\iota^*}(l,\iota^*(l))$

$$\leq -\bar{v}^*_{s\iota^*}(l,\iota^*(l))'\bar{v}^*_{s\iota^*}(l,\iota^*(l)) \tag{8.28}$$

The relation (8.28) is induced from the fact that in each period we select the index value which minimizes the transfer energy from current state $\bar{x}_{\iota^*}(l)$ to the target state X_d. So the sum of the differences in (8.28) is negative or equal to zero.

From (8.25) and (8.27) we may write :

$$\mathcal{V}(\iota^*(l+1),\bar{x}_\iota(l+1)) - \mathcal{V}(\iota^*(l),\bar{x}_\iota(l)) \leq -\gamma \left\|\bar{x}_\iota(l)\right\| \tag{8.29}$$

Secondly from theorem (8.3) it is clear that, for each current and optimal index value, there exists a sequence $\{\bar{x}_\iota(l)\}$ (the optimal one) which assures that $\lim_{l\to+\infty} \mathcal{V}(\iota^*,\bar{x}_\iota(q)) = 0$ (the sequence $\{\bar{x}_\iota(l),\ldots,\bar{x}_\iota(l+\mathcal{H}), X_d = 0,\ldots\}$) which implies that the states converge towards zero. From [Goodwin et al., 2005] the optimal control and scheduling obtained from the theorem (8.4) assures the asymptotic stability. □

8.5 Conclusions and Perspectives

In this chapter, the problem of optimal integrated control/observation and scheduling under communication constraints is addressed based on our work [Çela et al., 2010]. These constraints are expressed by the number or subset of control/sensor signals we may send from the controller node to the actuator node and from sensor node to the controller node respectively in each sampling period. Thus, based on our paper [Çela et al., 2010] we propose a detailed design methodology in the case of constrained communication of control signals.

The solution of the optimal integrated control and scheduling under communication constraints is constructed in two phases. From the first phase or *off line phase* we obtain a static solution depending only on the intrinsic properties of the system and maximization of the *degree of reachability/observability* [Antoulas, 2005]. It consists of a periodic scheduling function of a reduced set of control/sensor signals and its induced feedback periodic controller. The second phase solution or *real-time solution* is obtained based on the periodic properties of the system induced by the periodic scheduling function. In this phase we modify, for each period, the current index value ι ($\iota \in \{0,1,\ldots,T_\delta - 1\}$) or equivalently choosing one of the T_δ linear arrangements of the reduced set control signals to send from the controller node to the actuator node. An optimization problem is formulated and resolved. Its criteria are related to the *degree of reachability* [Antoulas, 2005] for each configuration of hybrid system. So in runtime we construct a new periodic scheduling function by choosing the current index value depending on the current state of the system or the perturbation influence. The problem treated in this chapter is much in common to the problem of optimally locating the actuators (sensors) proposed in [Georges, 1995] and related papers. Here, we propose a sub-optimal real-time solution of the actuators and sensors placement / ordering in time-domain modifying their optimal static scheduling.

Model reduction due to the reduced set of control signals applied in each sampling period combined with more efficient update algorithm of scheduling function will constitute the direction of our future research.

9 Conclusion

9.1 Summary

In this thesis, a new approach for control of distributed embedded system under communication constraints was proposed and analyzed using information theoretic explanations and view points. Information-theoretic views of constrained control was studies and analyzed using Shannon Entropy, Mutual-Information and Fisher Information. Controllability parameter in terms of Controllability Gramian matrix was related with Shannon information and entropy. It was also related with Fisher Information Matrix (**FIM**). Thus information theory, estimation theory and control theory were converged in a novel way.

We have studied and analyzed distributed embedded system control under communication constraints with special emphasis on information-theoretic aspects and convergence of control theory, information theory and estimation theory.

We have also discussed the Quantized Control especially in the light of Information Theory.

We have surveyed literatures dealing with information-theoretic control in view of the control under communication constraints, which vast majority of researchers in conventional control theory have overlooked.

We have analyzed and discussed the Degree of Reachability and Observability as Control-theoretic Metrics for Optimal Integrated Control and Scheduling of Networked Control Systems (NCS). We have proposed a detailed design methodology in the case of constrained communication of control signals.

9.2 Future work

The following research directions represent possible extensions to the work presented in the thesis :

Experimental and Simulation Results supporting the Information-Theoretic view of Control: Although we have analyzed and correlated the informational aspects with control theoretic aspects, we need to validate these theoretic results experimentally and by simulations in future.

Enhancing Zooming Algorithm and Optimal Bandwidth Allocation: Future work in this direction would be also to propose an information-theoretic analysis for enhancing the zooming algorithm proposed in [Ben Gaid and Çela, 2006] and optimal allocation of communication bandwidth which maximizes the systems' performances based on Controllability Grammians.

Cooperation with Limited Information: We observe that information transmission / processing and feedback are central to generating cooperative behavior. Without information transmission / processing or feedback, there is no cooperation or cooperative behavior. However, in a cooperative system with multiple agents, it is not easy to separate the notions of information and control : The transmission of information can be viewed as feedback among agents; information processing is intertwined with estimation and detection; and the processed information is utilized for decision making and generating control commands. Therefore, it is necessary and advantageous to study information, estimation, and control jointly in a cooperative system. We would like to see the broader impact of our study on cooperation problems in future.
We emphasize that the introduction of noise is important for studying the interplay between information / feedback and the cooperative behavior, since it allows us, first, to keep track of how information is conveyed, processed, and utilized in such a system, and second, to understand how local information exchange may be used to generate global behaviors.

Information-Theoretic Analysis of Degree of Controllability / Observability : As Degree of Controllability and Degree of Observability are control metrics of practical importance, we would like to extend our work to the information-theoretic analysis of them and also to provide a metric (measure) definition of Degree of Controllability / Observability from information-theoretic point of view.

Appendices

A

Analysis of First-Order Information

Here, we discuss the classical analysis of state-space models including the system abilities like controllability and observability based on [Skelton et al., 1998]. We consider a linear time-varying continuous dynamic system as well as discrete one to analyze the controllability and observability of the system.

Solution of Linear Differential Equations for Continuous Systems

Considering the linear time-varying continuous dynamic system of the form :

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) + D(t)u(t) \end{aligned} \tag{A.1}$$

where $A(t), B(t), C(t)$ and $D(t)$ are the system matrices which may be functions of time t and $x(t)$ is the state vector, $u(t)$ is the control input and $y(t)$ is the output. $y(t)$ is called the First-Order Information. The solution to (A.1) is given by

$x(t) = \Phi(t, t_0)x(t_0) + \int_{t_0}^{t} \Phi(t, \sigma)B(\sigma)u(\sigma)d\sigma$

where $\Phi(t, t_0)$ is called the state transition matrix which is generated by solving the differential equation

$\frac{d}{dt}\Phi(t, t_0) = A(t)\Phi(t, t_0),\ \Phi(t, t_0) = \mathbf{I}$; $\mathbf{I}$ being the Identity Matrix.

If A is a constant matrix, then

$$\Phi(\tau + t_0, t_0) = \Phi(\tau, 0) = e^{A\tau} \triangleq \sum_{i=0}^{\infty} \frac{A^i \tau^i}{i!} \tag{A.2}$$

and

$$x(t) = e^{A(t-t_0)} x(t_0) + \int_{t_0}^{t} e^{A(t-\sigma)} B u(\sigma) d\sigma \tag{A.3}$$

Solution of Linear Difference Equations for Discrete Systems

Considering the linear discrete-time dynamic system of the form :

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k \\ y_k &= C_k x_k + D_k u_k \end{aligned} \tag{A.4}$$

where $A_k, B_k, C_k, D_k, x_k, y_k, u_k$ denote matrices and vectors that are functions of the time index k, i.e. at time t_k, $A_k = A(t_k)$, etc. The solution to (A.4) for x_k is

$$\begin{aligned} x_k &= \Phi_{k0} x_0 + \sum_{i=1}^{k} \Phi_{ki} B_{i-1} u_{i-1} \\ \Phi_{kk} &= \mathbf{I} \Phi_{ki} \triangleq \prod_{\alpha=1}^{\alpha=k-1} A_\alpha \end{aligned}$$

If A and B are constants then $A_\alpha = A$ for all α and the solution reduces to

$$x_k = A^k x_0 + \sum_{i=1}^{k} A^{k-i} B u_{i-1} \tag{A.5}$$

Controllability and Observability of Continuous-Time Systems

Controllability :

Consider the system

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{A.6}$$

Definition A.1. System (A.6) is said to be completely state controllable at time $t = t_0$ if there exists a time $t_f > t_0$ and a control $u(t)$, $t \in [t_0, t_f]$ such that the state is transferred from an arbitrary initial state $x(t_0) = x_0$ to an arbitrarily specified $x(t_f) = x_f$ in a finite time $t_f < \infty$.

Suppose we wish to know whether the system (A.6) is completely controllable at t_0. Then it is equivalent to asking whether there exists a $u(\sigma), \sigma \in [t_0, t_f]$ such that

$$\int_{t_0}^{t_f} \Phi(t_f, \sigma) B(\sigma) u(\sigma) d\sigma = \tilde{x} \tag{A.7}$$

for some $t_f < \infty$ and for any specified $\tilde{x} \triangleq x_f - \Phi(t_f, t_0) x_0$. Since every element of the vector $\tilde{x}$ is arbitrary, the rows of the matrix

$$R(\sigma) \triangleq \Phi(t_f, \sigma) B(\sigma)$$

must be linearly independent on the interval $\sigma \in [t_0, t_f]$. This is equivalent to

$$X(t_f) \triangleq \int_{t_0}^{t_f} R(\sigma) R^T(\sigma) d\sigma > 0. \tag{A.8}$$

Replacing t_f by t and differentiating $X(t_f)$ with respect to t in (A.8), we get

$$\begin{aligned} \dot{X}(t) &= X(t) A^T(t) + A(t) X(t) + B(t) B^T(t) \\ X(t_0) &= 0 \\ X(t_f) &> 0. \end{aligned} \tag{A.9}$$

Theorem A.1. *The system $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ is completely state controllable at time t_0 if and only if there exists $t_f < \infty$ such that (A.9) holds.*

Now, suppose that A, B, C are constant matrices hereafter, then equation (A.8) can also be written as follows :

$$\begin{aligned} X(t_f) &= \int_{t_0}^{t_f} e^{A(t_f - \sigma)} B B^T e^{A^T (t_f - \sigma)} d\sigma \\ &= -\int_{t_f - t_0}^{0} e^{A\tau} B B^T e^{A^T \tau} d\tau \\ &= \int_{t_0}^{t_f - t_0} e^{A\tau} B B^T e^{A^T \tau} d\tau \end{aligned}$$

Since the integrand $e^{A\tau} B B^T e^{A^T \tau} d\tau \geq 0$, it follows that $X(t_2) \geq X(t_1)$ if $t_2 \geq t_1$ for any given t_0.
Hence, the existence of a t_f such that $X(t_f) > 0$ does not depend upon the choice of t_0.

Now supposing $X(t_f) > 0$ for some $t_f < \infty$, then $X(\tilde{t}) > 0$ for every $\tilde{t} \geq t_f$, including the limiting case $t_f = \infty$. Likewise, if $X(\infty)$ is not positive definite then $X(\tilde{t})$ also is not positive definite for any $t_f < \infty$. This proves the following corollary.

Corollary A.1. *The linear time-invariant system $\dot{x} = Ax + Bu$ is completely state controllable if and only if there exists some $t_f < \infty$ such that*

$$X(t_f) = \int_{t_0}^{t_f} e^{A\tau} B B^T e^{A^T \tau} d\tau > 0.$$

or equivalently (A.9) holds for $t_0 = 0$ and A, B constant.

Noting the fact that $X(t_f)$ always exists for $t_f < \infty$ but might not exist for $t_f = \infty$. Now defining $X(\infty)$ by (if it exists)

$$X(\infty) = \int_0^\infty e^{A\tau} B B^T e^{A^T \tau} d\tau. \tag{A.10}$$

This matrix is called the *Controllability Gramian* which is denoted by $\mathbf{W}_{c\infty}$ in infinite form or $\mathbf{W}_c$ in the finite form $(t = t_f)$ in the rest of the chapters of this thesis.

Corollary A.2. *$X(\infty)$ exists if and only if the controllable modes of $\dot{x} = Ax + Bu$ are asymptotically stable. If $X(\infty)$ exists, then $X(\infty) > 0$ if and only if (A, B) is controllable pair [Skelton et al., 1998]. Here, the "modes" are characterized by the eigenvalue, eigenvector pairs and "controllable mode" i that is asymptotically stable corresponds to negative real part of the eigenvalue.*

Corollary A.3. *If the controllable modes of $\dot{x} = Ax + Bu$ are asymptotically stable, the following statements are equivalent :*

1. *The system is completely state controllable.*
2. *$\int_0^\infty e^{A\tau} B B^T e^{A^T \tau} d\tau > 0$, i.e. the controllability gramian is positive definite.*
3. *$X > 0, AX + XA^T + BB^T = 0$*
4. *$rank \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix} = n$ (dimension of x).*

Observability :

Now consider the system

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{A.11}$$

Let us determine $x(t_0)$ given the observed data $y(t), t_0 \le t \le t_f$. Note that the knowledge of $x(t_0)$ is equivalent to the knowledge of $x(t)$ for any t, since $\Phi(t, t_0)$ is invertible and $x(t) = \Phi(t, t_0)x(t_0)$. Using equation (A.11) and the value of $x(t)$ we get

$$y(t) = C(t)\Phi(t, t_0)x(t_0) \tag{A.12}$$

Definition A.2. The system (A.11) is said to be completely observable at time $t_f > t_0$ if the data $y(t), t \in [t_0, t_f]$ yields a unique solution $x(t_0)$ to (A.12).

Now considering $y(t) = C(t)x(t)$ and some given data $y(t)$ over an interval $t_0 \le t \le t_f$, we can say that for unique solution to $x(t_0)$ in equation (A.12) the columns of $C(t)\Phi(t, t_0)$ must be linearly independent on the interval $[t_0, t_f]$. This means that

$K(t_0) \triangleq \int_{t_0}^{t_f} \Phi^T(\sigma, t_0) C^T(\sigma) C(\sigma) \Phi(\sigma, t_0) d\sigma > 0$

(where $\dot{\Phi}(t, t_0) = A(t)\Phi(t, t_0)$, $\Phi(t_0, t_0) = \mathbf{I}$), or equivalently

$$\left.\begin{array}{l} -\dot{K}(t) = K(t)A(t) + A^T(t)K(t) + C^T(t)C(t) \\ K(t_f) = 0 \\ K(t_0) > 0 \text{ for some } t_0 < t_f \end{array}\right\} \tag{A.13}$$

These results are summarized as follows :

Theorem A.2. *The system $\dot{x}(t) = A(t)x(t), y(t) = C(t)x(t)$ is completely observable at time t_f if and only if there exists $0 < t_0 < t_f$ such that $K(t_0) > 0$ where*

$$\begin{aligned} -\dot{K}(t) &= K(t)A + A^T K(t) + C^T C \\ K(t_f) &= 0. \end{aligned} \tag{A.14}$$

The time-invariant cases follow in a natural way from the above theorems by setting $\dot{K}(t)$ to zero. The matrix K below is called the *Observability Gramian*

$$K = \int_0^\infty e^{A^T \tau} C^T C e^{A\tau} d\tau. \tag{A.15}$$

Corollary A.4. *If the observable modes of $\dot{x} = Ax, y = Cx$ are asymptotically stable, the following statements are equivalent :*

1. *The system is completely observable.*
2. $\int_0^\infty e^{A^T \tau} C^T C e^{A\tau} d\tau > 0.$
3. $K > 0$, $KA + A^T K + C^T C = 0.$
4. *rank*$\left[\begin{array}{cccc} C^T & A^T C^T & \dots & A^{(n-1)T} C^T \end{array}\right] = n$ *(dimension of x)*

Controllability and Observability of Discrete-Time Systems

Controllability :

Consider now the discrete-time system (A.4) as defined earlier.

Definition A.3. The system (A.4) is called "output controllable at time k_0" if there exists an integer k_f and a control sequence $\{u_{k_0}, u_{k_0+1}, u_{k_0+2}, \dots, u_{k_f}\}$ such that $y_{k_f} = y_f$ for an arbitrarily specified y_f, for any given initial state x_{k_0}.

When (A, B, C, D) are constant matrices, the "at time k_0" can be deleted in the definition and $k_0 = 0$ can be substituted without loss. When $C = \mathbf{I}$, y is replaced by x and C, D need not be stated, where output controllability reduces to state controllability in the definition.

Theorem A.3. *These statements are equivalent :*

1. *The matrix time-varying tuple* (A_k, B_k, C_k, D_k) *is output controllable at time* k_0.
2. *There exists* $k_f > k_0$ *such that*

$$\begin{aligned} X_{k+1} &= A_k X_k A_k^T + B_k B_k^T && \text{(A.16)} \\ X_{k_0} &= 0 \\ C_{k_f} X_{k_f} C_{k_f}^T + D_f D_f^T &> 0. && \text{(A.17)} \end{aligned}$$

Theorem A.4. *Let* A, B, C *be constant and suppose* X *exists satisfying*

$$X = AXA^T + BB^T. \tag{A.18}$$

Then the following statements are equivalent

1. *The system (A.4) is output controllable.*
2. $CXC^T + DD^T > 0$.
3. *The system (A.4) is state controllable.*
4. $X > 0$.

The solution to (A.18), if it exists, is

$$X = \sum_{i=0}^{\infty} A^i B B^T (A^T)^i. \tag{A.19}$$

as proved by direct substitution into (A.18). From (A.19)

$$X = \Omega\Omega^T, \text{ where } \Omega \triangleq \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix}. \tag{A.20}$$

making it clear that rank X = rank Ω.

Theorem A.5. *If* X *in (A.19) exists , then the following statements are equivalent :*

1. *The matrix pair* (A, B) *is state controllable.*
2. *rank* X = *rank* $\begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix} = n$ = *dimension of* A.
3. $X > 0$.

Observability :

Let us define the system

$$\begin{aligned} x_k &= A_k x_k + B_k u_k \\ y_k &= C_k x_k + D_k u_k \end{aligned} \tag{A.21}$$

Definition A.4. We say that (A.21) is observable at time p if there exists a time $q \leq p$ such that the knowledge of $\{u(k), y(k) q \leq p\}$ allows a unique solution for $x(q)$.

From (A.21) we can write

$$\begin{bmatrix} y_q \\ y_{q+1} \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} C_q \\ C_{q+1}A_q \\ \vdots \\ C_p A_{p-1} \dots A_q \end{bmatrix} x_q + \begin{bmatrix} D_q & \dots & \dots & \dots \\ C_{q+1}B_q & D_{q+1} & \ddots & \dots \\ \vdots & \dots & \dots & \dots \\ C_p A_{p-2} \dots A_{q-1} B_q & \dots\dots & \dots & D_p \end{bmatrix} \begin{bmatrix} u_q \\ \vdots \\ \vdots \\ u_p \end{bmatrix}$$

or,simply,

$$\tilde{y}(p,q) = \tilde{C}(p,q) x_q + \tilde{B}(p,q) \tilde{u}(p,q). \tag{A.22}$$

Hence, observability at time p is equivalent to the existence of a unique x_q satisfying (A.22), given the matrix $\tilde{C}(p,q)$, and the vector $\hat{y}(p,q) \triangleq \tilde{y}(p,q) - \tilde{B}(p,q)\tilde{u}(p,q)$. This linear algebra problem has solution

$$x_q = \tilde{C^+}(p,q)\hat{Y}(p,q) + (\mathbf{I} - \tilde{C^+}(p,q)\tilde{C}(p,q))z.$$

(where the superscript $+$ represents the Moore-Penrose Inverse)

if the following existence condition holds

$$(\mathbf{I} - \tilde{C}(p,q)\tilde{C^+}(p,q))\hat{y}(p,q) = 0. \tag{A.23}$$

The solution x_q is unique if the columns of $\tilde{C}(p,q)$ are linearly independent, in which case

$$\mathbf{I} - \tilde{C^+}(p,q)\tilde{C}(p,q) = 0. \tag{A.24}$$

There exists at least one solution for $x(q)$ if (A.23) holds and there exists a solution for arbitrary input/output data $\hat{Y}(p,q)$ if and only if the rows of $\tilde{C}(p,q)$ are linearly independent so that

$$\mathbf{I} - \tilde{C}(p,q)\tilde{C^+}(p,q) = 0. \tag{A.25}$$

Since $\tilde{C} \in \mathcal{R}^{n_y(p-q+1)\times n_x}$, uniqueness requires $n_y(p-q+1) \geq n_x$. Specifically, we require a left inverse of $\tilde{C}(p,q)$, or equivalently

$$W_o(p,q) \triangleq \tilde{C^T}(p,q)\tilde{C}(p,q) > 0.$$

W_o being the Observability Gramian.

or

$$W_o(p,q) = \sum_{i=0}^{p-q} A^{Ti} C^T C A^i. \tag{A.26}$$

Since $W_o(p_2, q_2) \geq W_o(p_1, q_1)$ where $p_2 - q_2 \geq p_1 - q_1$, we test observability in the time-invariant case by the condition

$$W_o \triangleq \sum_{i=0}^{\infty} A^{Ti} C^T C A^i > 0. \tag{A.27}$$

Since observability is a function only of the matrix pair (A, C) we may say, relative to (A.21) that the "matrix pair (A, C) is observable" (or not).

Theorem A.6. *The following statements are equivalent*

1. *The time-varying system (A.21) is observable at time p.*
2. *There exists q such that*

$$\begin{aligned} W_{ok} &= A_k^T W_{ok+1} A_k + C_k^T C_k, W_{op} = 0 \\ W_{oq} &> 0. \end{aligned} \tag{A.28}$$

If (A, C) is a pair of constant matrices and if W_{oq} exists from (A.28) for $p = \infty$, the following statements are equivalent :

3. *The time-varying system (A.21) is observable.*
4. *$W_o = A^T W_o A + C^T C$, $W_o > 0$.*

For the time-invariant case using Cayley-Hamilton Theorem we can write

rank W_o = rank $\sum_{i=0}^{\infty} A^{iT} C^T C A^i$ = rank $\begin{bmatrix} C^T & A^T C & \dots \end{bmatrix} \begin{bmatrix} C^T & A^T C^T & \dots \end{bmatrix}^T$

= rank $\begin{bmatrix} C^T & A^T C^T & \dots A^{(n-1)T} C^T \end{bmatrix}$.

Hence, observability is equivalent to

rank $\begin{bmatrix} C^T & A^T C^T & \dots A^{(n-1)T} C^T \end{bmatrix} = n$ = dimension of A.

B

Receding Horizon Optimal Control with Constraints

B.1 Introduction

Receding Horizon Control (*RHC*) strategies have become quite popular recently. This interest is partly due to the availability of faster and cheaper computers as well as efficient numerical algorithms for solving optimization problems. Another key advantage of these strategies is the potential ability to handle control saturations. Many of the successful applications of receding horizon control methods have been in the area of chemical process control. Several researchers have attempted to address the problem of stability for receding horizon control to allow its application in stability critical areas such as unmanned aerial vehicles (UAV).

Here, we discuss the principle of Receding Horizon Optimal Control with Constraints based on [Goodwin et al., 2005]. Receding Horizon Control (*RHC*) is usually represented by a state feedback control if states are available. However, full states may not be available, since measurement of all states may be expensive or impossible from resource constraints point of view. In RHC we can utilize the measured data in recent finite time $[k - N, k]$, (here k being the time instant and N being the size of the horizon and is a design parameter) and obtain an estimated state by a linear combination of the measured inputs and outputs over the receding finite horizon with some weighting gains to be chosen such that error between the actual state and the estimated one is minimized.

The idea is to start with a fixed optimization horizon of length N using the current state of the plant as the initial state. We then optimize the performance function over this

fixed interval considering only the constraints, getting an optimal sequence of N control moves and applying only the first control move to the plant. Time then increments one step and the same N-step optimization problem is considered using the new state of the plant as the initial state. Thus one continuously revises the current control action based on the current state considering the constraints over an optimization horizon of length N.

B.2 The Receding Horizon Control Principle

Fixed horizon optimization leads to a control sequence $\{u_i, \ldots, u_{i+N-1}\}$ which begins at the current time i and ends at some future time $i + N - 1$. This fixed horizon solution suffers from two potential drawbacks :

- Something unexpected may happen to the system at some time over the future interval $[i, i + N - 1]$ that was not predicted by (or included in) the model. This would render the fixed control choices $\{u_i, \ldots, u_{i+N-1}\}$ obsolete.
- As one approaches the final time $i+N-1$, the control law typically gives up trying since there is too little time to go to achieve anything useful in terms of objective function reduction.

The above two problems are addressed by the idea of receding horizon optimization. This idea can be summarized as follows :

1. At current time i with the current state x_i, solve an optimal control problem over a fixed future interval, say $[i, i + N - 1]$, taking into account the current and future constraints.
2. Apply only the step-1 in the resulting optimal control sequence.
3. Measure the state x_{i+1} reached at time $i + 1$.
4. Repeat the fixed horizon optimization at time $i + 1$ over the future interval $[i + 1, i + N]$, starting from (now) the current state x_{i+1}.

In the absence of disturbances, the state measured in step-3 will be the same as that predicted by the model. Nonetheless, it seems prudent to use the measured state rather than the predicted state for surity. The above description assumes that the state is measured at time $i + 1$. In practice, one would use some form of observer to estimate x_{i+1} based on the available data. For the moment, we assume that the full state vector is measured and the impact of disturbances is ignored.

If the model and objective function are time invariant, then the same input u_i will result whenever the state takes the same value. That is, the receding horizon optimization strategy is really a particular time-invariant state feedback control law illustrated in Fig. (B.1) :

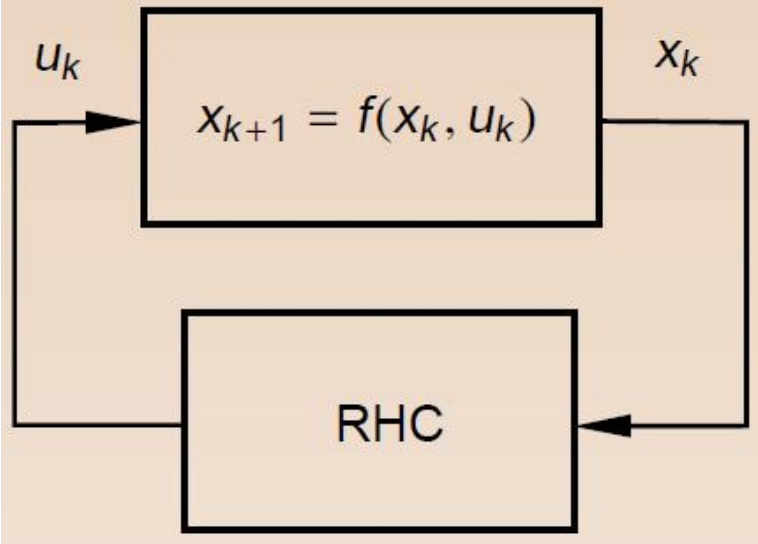


Figure B.1: Receding Horizon Control

In particular, we can set $i = 0$ in the formulation of the open loop control problem. More precisely, at the current time i, and for the current state x, we can solve :

$$\begin{aligned}
\mathcal{P}_N(x) : V_N^{opt}(x) &\triangleq \min V_N\left(\{x_k\}, \{u_k\}\right); && \text{(B.1)}\\
\text{subject to :}& \\
x_{k+1} &= f(x_k, u_k) \text{ for } k = 0, \ldots, N-1 && \text{(B.2)}\\
x_0 &= x; && \text{(B.3)}\\
u_k &\in \mathbb{U} \quad \text{for } k = 0, \ldots, N-1; && \text{(B.4)}\\
x_k &\in \mathbb{X} \quad \text{for } k = 0, \ldots, N; && \text{(B.5)}\\
x_N &\in \mathbb{X}_f \subset \mathbb{X}; && \text{(B.6)}\\
\text{where the objective function, } V_N\left(\{x_k\}, \{u_k\}\right) &\triangleq F(x_N) + \sum_{k=0}^{N-1} L(x_k, u_k); && \text{(B.7)}
\end{aligned}$$

The sets $\mathbb{U} \subset \mathbb{R}^m$, $\mathbb{X} \subset \mathbb{R}^n$, and $\mathbb{X}_f \subset \mathbb{R}^n$ are the input, state and terminal constraint set, respectively. All sequences $\{u_k\} = \{u_0, \ldots, u_{N-1}\}$ and $\{x_k\} = \{x_0, \ldots, x_N\}$ are satisfying the constraints (B.2) - (B.6) are called feasible sequences. A pair of feasible sequences $\{u_0, \ldots, u_{N-1}\}$ and $\{x_0, \ldots, x_N\}$ constitute a feasible solution. The functions F and L in the objective function (B.7) are the terminal state weighting and the per-stage weighting, respectively.

In the sequel, we make the following assumptions :

- f, F and L are continuous functions of their arguments;
- $\mathbb{U} \subset \mathbb{R}^m$ is a compact set, $\mathbb{X} \subset \mathbb{R}^n$, and $\mathbb{X}_f \subset \mathbb{R}^n$ are closed sets; Here, a compact set is the one that is both bounded and closed.

- there exists a feasible solution to problem (B.1) - (B.7).

Because N is finite, these assumptions are sufficient to ensure the existence of a minimum by *Weierstrass' theorem*.
Weierstrass' Theorem : *Let $\mathcal{S} \subset \mathbb{R}^n$ be a non-empty set, and let $f : \mathcal{S} \to \mathbb{R}$ be a continuous function on $\mathcal{S}$, i.e., there exists a minimizing solution to the problem* $\min\{f(x) : x \in \mathcal{S}\}$.

Typical choices for the weighting functions F and L are quadratic functions of the form :

$F(x) = x^T P x$ and $L(x,u) = x^T Q x + u^T R u$;

where $P = P^T \geq 0, Q = Q^T \geq 0$ and $R = R^T > 0$.

More generally, one could use functions of the form :

$F(x) = \|Px\|_p$ and $L(x,u) = \|Qx\|_p + \|Ru\|_p$;

where $\|y\|_p$ with $p = 1, 2, \ldots, \infty$. is the $p-$norm of the vector y.

Denoting the minimizing control sequence, which is a function of the current state x_i, by

$$\mathscr{U}_{x_i}^{opt} = \left\{u_0^{opt}, u_1^{opt}, \ldots, u_{N-1}^{opt}\right\}; \tag{B.8}$$

we can say that the control applied to the plant at time i is the first element of this sequence, i.e.,

$$u_i = u_0^{opt}. \tag{B.9}$$

Time is then incremented one instant, and the above procedure is repeated for another $N-$step-ahead optimization horizon.
The first element of the new $N-$step input sequence is then applied, and so on.

The above procedure is called Receding Horizon Control (*RHC*). The Fig. (B.2) illustrates the RHC principle for horizon $N = 5$. Each plot shows the minimising control sequence $\mathscr{U}_{x_i}^{opt}$ given in (B.8), computed at time $i = 0, 1, 2$. Note that only the shaded inputs are actually applied to the system. We can see that we are continually looking ahead to judge the impact of current and future decisions on the future response.

The above receding horizon procedure implicitly defines a time-invariant control policy $\mathcal{K}_N : \mathbb{X} \to \mathbb{U}$ of the form :

$$\mathcal{K}_N(x) = u_0^{opt}. \tag{B.10}$$

The receding horizon controller is implemented in closed loop as follows in Fig. (B.3). Note that the strict definition of the function $\mathcal{K}_N(.)$ requires the minimizer to be unique. Most of the problems treated in this context are convex and hence satisfy this condition. It is common in receding horizon control applications to compute numerically, at time

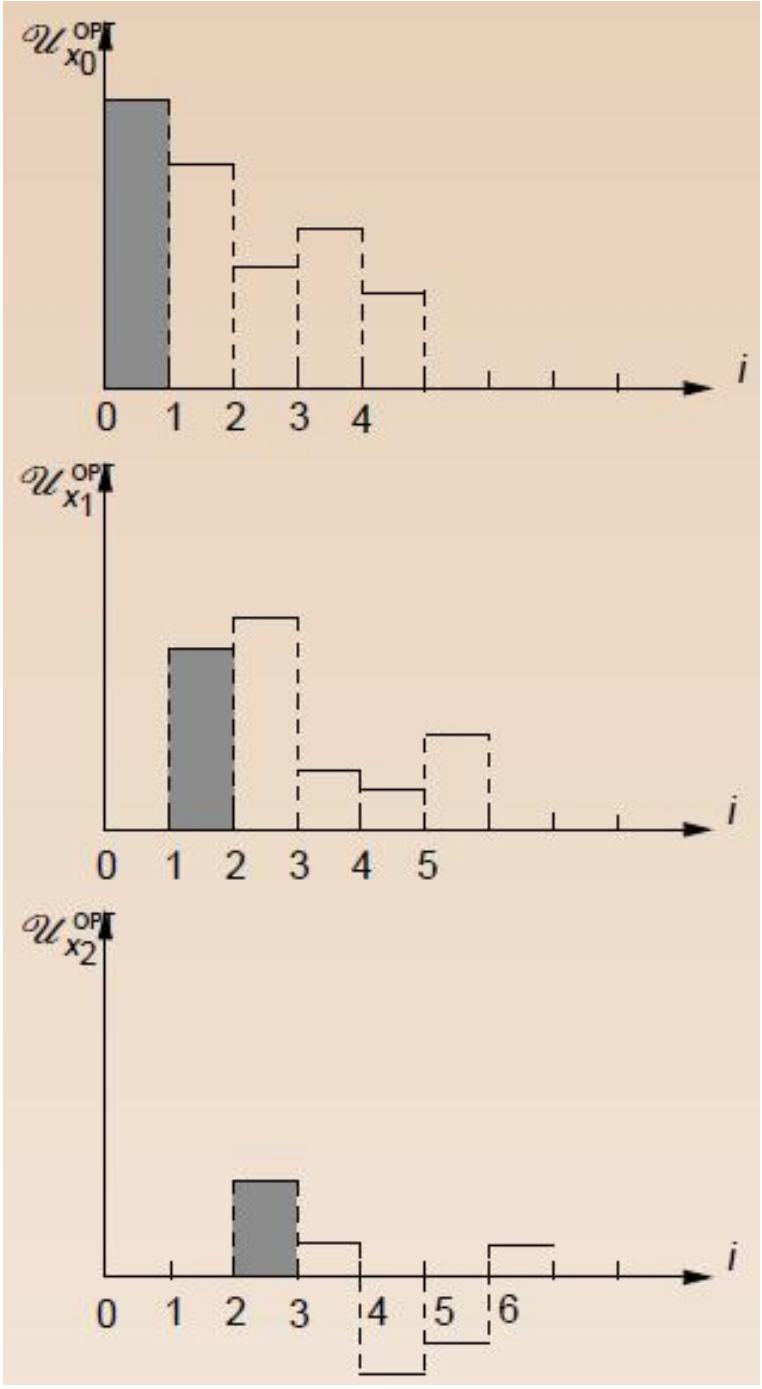


Figure B.2: Receding Horizon Control (RHC) Principle

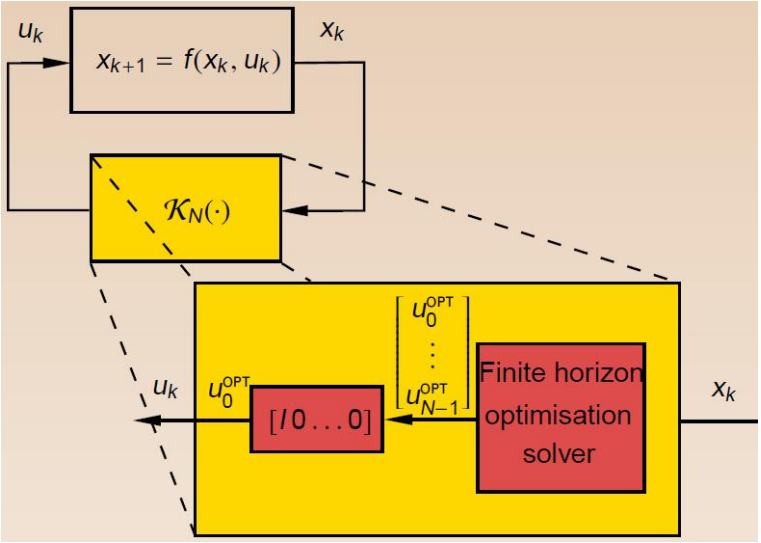


Figure B.3: Receding Horizon Control in Closed Loop

i, and for the current state $x_i = x$, the optimal control move $\mathcal{K}_N(x)$. In this case, we call it an implicit receding horizon optimal policy. In some cases, we can explicitly evaluate the control law $\mathcal{K}_N(.)$. In this case, we say that we have an explicit receding horizon optimal policy.

B.3 Stability of Receding Horizon Optimal Control

Optimality can be turned into a notion of stability by utilizing the value function ($V_N^{opt}(x)$ in equation (B.1), which is a function of initial state only) as a Lyapunov function. But, the difficulty lies in the fact that optimization problems that we are solving are only defined over a finite future horizon whereas the stability is a property that must hold over an infinite future horizon. To resolve this problem an appropriate weighting is added to the terminal state in the finite horizon problem so as to take care the impact of events that lie beyond the end of the fixed horizon. This effectively turns the fixed horizon problem into an infinite horizon one.
Hence, a terminal control law and an associated terminal state weighting in the objective function (that captures the impact of using the terminal control law over an infinite time) can be defined. Usually, the chosen terminal control laws are simple and only feasible in a restricted (local) region. There are three factors on which the stability depends:

- a *terminal constraint set* $\mathbb{X}_f$ in the state space which is invariant under terminal control law,
- a feasible *terminal control law* $\mathcal{K}_f$ that holds in the terminal constraint set,

- a *terminal state weighting* F on the finite horizon optimization problem, usually corresponding to the objective function value generated by the use of the terminal control law over an infinite time.

Definition B.1. The set $\mathbb{S}_N$ of feasible initial states is the set of initial states $x \in \mathbb{X}$ for which there exists feasible state and control sequences for the fixed horizon optimal control problem $\mathcal{P}_N(x)$ in (B.1) - (B.7).

Definition B.2. The set $\mathbb{S} \subset \mathbb{R}^n$ is said to be positively invariant for the system $x_{i+1} = f(x_i, u_i)$ under the control $u_i = \mathcal{K}(x_i)$ if $f(x, \mathcal{K}(x)) \in \mathbb{S}$ for all $x \in \mathbb{S}$.

We make the following assumptions on the data of problem $\mathcal{P}_N(x)$ in (B.1) - (B.7) :

A1 The terminal constraint set in (B.6) is the origin, i.e. $\mathbb{X}_f = \{0\}$.

A2 The control constraint set in B.4 contains the origin, i.e. $0 \in \mathbb{U}$.

A3 $L(x, u)$ in (B.7) satisfies $L(0, 0) = 0$ and $L(x, u) \geq \gamma(\|x\|)$ for all $x \in \mathbb{S}_N, u \in \mathbb{U}$, where a function $\gamma : [0, \infty) \to [0, \infty)$ is continuous, $\gamma(t) > 0$ for all $t > 0$ and $\lim_{t\to\infty} \gamma(t) = \infty$.

A4 There is no terminal state weighting in the objective function, i.e., $F(x) \equiv 0$ in (B.7).

Under the above conditions the following stability result can be obtained (refer [Goodwin et al., 2005] for proof) :

Theorem B.1. *Consider the system*

$$x_{i+1} = f(x_i, u_i) \text{ for } i \geq 0, \quad f(0, 0) = 0; \tag{B.11}$$

*controlled by the receding horizon algorithm (B.1) - (B.9) and subject to Assumptions **A1-A4** as stated above. Then,*

1. *The set $\mathbb{S}_N$ of feasible initial states is positively invariant for the closed loop system.*
2. *The origin is globally attractive in $\mathbb{S}_N$ for the closed loop system.*
3. *If, in addition to **A1-A4**, $0 \in$ int $\mathbb{S}_N$ and the value function V_N^{opt} in (B.1) is continuous on some neighbourhood of the origin, then the origin is asymptotically stable in $\mathbb{S}_N$ for the closed loop system.*

B.4 Conditions for Stability

Following are the more general terminal conditions :

B1 The per-stage weighting $L(x,u)$ in (B.7) satisfies $L(0,0)=0$ and $L(x,u)\geq\gamma(\|x\|)$ for all $x\in\mathbb{S}_N, u\in\mathbb{U}$, where $\gamma:[0,\infty)\rightarrow[0,\infty)$ is continuous, $\gamma(t)>0$ for all $t>0$ and $\lim_{t\rightarrow\infty}\gamma(t)=\infty$.

B2 The terminal state weighting $F(x)$ in (B.7) satisfies $F(0)=0$, $F(x)\geq 0$ for all $x\in\mathbb{X}_f$ and there exists a terminal control law $\mathcal{K}_f:\mathbb{X}_f\rightarrow\mathbb{U}$ such that $F\left(f(x,\mathcal{K}_f(x))\right)-F(x)\leq -L(x,\mathcal{K}_f(x))$ for all $x\in\mathbb{X}_f$.

B3 The set $\mathbb{X}_f$ is positively invariant for the system (B.11) under $\mathcal{K}_f(x)$ i.e. $f(x,\mathcal{K}_f(x))\in\mathbb{X}_f$ for all $x\in\mathbb{X}_f$.

B4 The terminal control $\mathcal{K}_f(x)$ satisfies the control constraints in $\mathbb{X}_f$, i.e. $\mathcal{K}_f(x)\in\mathbb{U}$ for all $x\in\mathbb{X}_f$.

B5 The sets $\mathbb{U}$ and $\mathbb{X}_f$ contain the origin of their respective spaces.

Using the above conditions, which are more general on terminal triple $(\mathbb{X}_f,\mathcal{K}_f,F)$, more general theorem on stability of RHC can be written as follows :

Theorem B.2. ***(Stability of Receding Horizon Control)***
*Consider the closed loop system formed by system (B.11), controlled by the receding horizon algorithm (B.1) - (B.9), and suppose that conditions **B1-B5** are satisfied. Then,*

1. *The set $\mathbb{S}_N$ of feasible initial states is positively invariant for the closed loop system.*
2. *The origin is globally attractive in $\mathbb{S}_N$ for the closed loop system.*
3. *If, in addition to **B1-B5**, $0\in int\mathbb{S}_N$ and the value function $V_N^{opt}(.)$ in (B.1) is continuous on some neighbourhood of the origin, then the origin is asymptotically stable in $\mathbb{S}_N$ for the closed loop system.*
4. *If, in addition to **B1-B5**, $0\in int\ \mathbb{X}_f$ and $\mathbb{S}_N$ is compact, $\gamma(t)\geq at^\sigma$ in **B1**, $F(x)\leq b\|x\|^\sigma$ for all $x\in\mathbb{X}_f$ in **B2**, where $a>0, b>0$ and $\sigma>0$ are some real constants, and the value function $V_N^{opt}(.)$ in (B.1) is continuous on $\mathbb{S}_N$, then the origin is exponentially stable in $\mathbb{S}_N$ for the closed loop system.*

B.5 Terminal Conditions for Stability

Considering the possible choices for the terminal triple $(\mathbb{X}_f,\mathcal{K}_f,F)$ that satisfy the conditions **B1-B5** of Theorem B.2 we proceed as follows :

One choice for the terminal state weighting $F(x)$ is the value function $V_\infty^{opt}(x)$ for the associated infinite horizon constrained optimal control problem, defined as follows :

$$\mathcal{P}_\infty(x) : V_\infty^{opt}(x) \triangleq \min V_\infty\left(\{x_k\}, \{u_k\}\right); \tag{B.12}$$

subject to :

$x_{k+1} = f(x_k, u_k)$ for $k = 0, \ldots,$

$x_0 = x$;

$u_k \in \mathbb{U}$ for $k = 0, \ldots,$

$x_k \in \mathbb{X}$ for $k = 0, \ldots,$

where $\{x_k\}$ and $\{u_k\}$ are infinite sequences and

$$V_\infty\left(\{x_k\}, \{u_k\}\right) \triangleq \sum_{k=0}^{\infty} L(x_k, u_k). \tag{B.13}$$

Note that $\mathcal{P}_\infty(x)$ does not have either a terminal state weighting nor a terminal state constraint; both are irrelevant since, if a solution to the problem exists, the state must converge to zero as $k \to \infty$ (since L is assumed to satisfy condition **B1**). In this case, it follows from the principle of optimality that the finite horizon value function for problem $\mathcal{P}_N(x)$ in (B.1) is $V_N^{opt}(x) = V_\infty^{opt}(x)$. With this choice , on-line optimization is unnecessary and hence the infinite horizon problem is automatically advantageous. But, the constraints render this approach impossible. Usually, then, $\mathbb{X}_f$ is then chosen to be in the vicinity of the origin in which case $V_\infty^{opt}(x)$ is exactly or approximately known and $F(x) = V_\infty^{opt}(x)$ or its approximation.

B.6 Conclusion

Since we have used the Receding Horizon Control (*RHC*) in chapter- 8, we have discussed here the Receding Horizon Control (*RHC*) under constraints like communication constraints in real-time systems. *RHC* Principle along with the stability conditions are discussed. For linear systems with input and state constraints that are common in industrial problems, RHC can be easily implemented using mathematical programming like quadratic and semidefinite programming. Even for non-linear systems, such as unmanned aerial vehicles, RHC can handle input and state constraints numerically in many cases due to the optimization over finite horizon. It guarantees stability under weak conditions under input and state constraints for linear as well as non-linear systems. Optimal control on infinite horizon i.e. the steady-state optimal control can also be an alternative.

Bibliography

[Antoulas, 2005] Antoulas, A. (2005). *Approximation of Large-Scale Dynamical Systems*. SIAM.

[Antsaklis and Baillieul, 2007] Antsaklis, P. and Baillieul, J. (2007). Special issue on the technology of networked control systems. *Proceedings of the IEEE*, 95(1).

[Anzai, 1974] Anzai, Y. (1974). A note on reachability of discrete-time quantized control systems. *IEEE Transactions on Automatic Control*, 19(5):575–577.

[Årzén et al., 2006] Årzén, K., Bicchi, A., Hailes, S., Johansson, K., and Lygeros, J. (2006). On the design and control of wireless networked embedded systems. In *Proc. of IEEE Computer Aided Control Systems Design Symposium*.

[Årzén et al., 2000] Årzén, K.-E., Cervin, A., Eker, J., and Sha, L. (2000). An introduction to control and real-time scheduling co-design. In *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia.

[Åström and Wittenmark, 1997] Åström, K. and Wittenmark, B. (1997). *Computer-Controlled Systems : Theory and Design, 3ed.* Prentice Hall.

[Baillieul, 1999] Baillieul, J. (1999). Feedback designs for controlling device arrays with communication channel bandwidth constraints. In *Proceedings of ARO Workshop on Smart Structures*, Pennsylvania State Univ.

[Baillieul, 2001] Baillieul, J. (2001). Feedback designs in information based control. In Pasik-Duncan, B., editor, *Proc. of a Workshop Held in Lawrence, Kansas*, Stochastic Theory and Control, pages 35–57. Springer-Verlag, New York, USA.

[Ben Gaid, 2006] Ben Gaid, M. (2006). *Optimal Scheduling and Control for Distributed Real-time Systems*. PhD thesis, Université d'Evry Val d'Essone, France.

[Ben Gaid and Çela, 2006] Ben Gaid, M. and Çela, A. (2006). Trading quantization precision for sampling rates in networked systems with limited communication. In *Proceedings of the 45th IEEE Conference on Decision & Control*, San Diego, CA, USA.

[Ben Gaid et al., 2006a] Ben Gaid, M., Çela, A., and Hamam, Y. (2006a). Optimal integrated control and scheduling of networked control systems with communication

constraints : Application to a car suspension system. *IEEE Transactions on Control Systems Technology*, 14(4).

[Ben Gaid et al., 2006b] Ben Gaid, M.-M., Çela, A., Hamam, Y., and Ionete, C. (2006b). Optimal scheduling of control tasks with state feedback resource allocation. In *Proceedings of the 2006 American Control Conference*, Minneapolis, Minnesota, USA.

[Bertsekas, 2000] Bertsekas, D. (2000). *Dynamic Programming and Optimal Control, Vol.1, 2ed.* Athena Press, MA, USA.

[Bicchi et al., 2002] Bicchi, A., Marigo, A., and Piccoli, B. (2002). On the reachability of quantized control systems. *IEEE Transactions on Automatic Control*, 47(4):546–563.

[Bittanti et al., 1991] Bittanti, S., Laub, A. J., and Willems, J., editors (1991). *The Riccati Equation*, chapter The periodic Riccati Equation. Springer-Verlag.

[Bode, 1945] Bode, H. (1945). *Network Analysis and Feedback Amplifier Design*. D., Van Nostrand.

[Borkar and Mitter, 1997] Borkar, V. and Mitter, S. (1997). LQG control with communication constraints. In *Communications, Computation, Control and Signal Processing*, pages 365–373. Norwell, MA: Kluwer.

[Branicky et al., 2000] Branicky, M., Phillips, S., and Zhang, W. (2000). Stability of networked control systems: Explicit analysis of delay. In *Proc. of American Control Conference*.

[Brockett, 1997] Brockett, R. (1997). Minimum attention control. In *Proceedings of the 36th IEEE Conference on Decision and Control*, San Diego, California, USA.

[Brockett and Liberzon, 2000] Brockett, R. and Liberzon, D. (2000). Quantized feedback stabilization of linear systems. *IEEE Transactions on Automatic Control*, (7):1279–1289.

[Brockett, 1970] Brockett, R. W. (1970). *Finite Dimensional Linear Systems*. Wiley, New York, USA.

[Çela et al., 2010] Çela, A., Reama, A., Niculescu, S., and Roy, P. (2010). Degree of reachability / observability as a metric for optimal integrated control and scheduling of networked control systems. In *submitted to the IFAC Symposium on System, Structure and Control*, Ancona, Italy.

[Cervin et al., 2002] Cervin, A., Eker, J., Bernhardsson, B., and Årzén, K.-E. (2002). Feedback-feedforward scheduling of control tasks. *Real-Time Systems*, 23(1):25–53.

[Conant, 1976] Conant, R. (1976). Laws of information which govern systems. *IEEE Trans. on Syst., Man, Cybern.*, 6:240–255.

[Cover and Thomas, 2006] Cover and Thomas (2006). *Elements of Information Theory, 2ed*. John Wiley and Sons.

[Curry, 1970] Curry, R. (1970). *Estimation and Control With Quantized Measurements*. MIT Press,Cambridge,USA.

[Delchamps, 1989] Delchamps, D. (1989). Extracting state information from a quantized output record. *Systems and Control Letters*, 13:365–372.

[Delchamps, 1990] Delchamps, D. (1990). Stabilizing a linear system with quantized state feedback. *IEEE Transactions on Automatic Control*, 35(8):916–924.

[Delvenne, 2006] Delvenne, J. (2006). An optimal quantized feedback strategy for scalar linear systems. *IEEE Transactions on Automatic Control*, 51(2):298–303.

[Egerstedt and Hu, 2002] Egerstedt, M. and Hu, X. (2002). A hybrid control approach to action coordination for mobile robots. *Automatica*, 38:125–130.

[Elia, 2000] Elia, N. (2000). Design of hybrid systems with guaranteed performance. In *Proc. 39th IEEE Conf. on Decision and Control*.

[Elia, 2002] Elia, N. (2002). Coarsest quantizer density for quadratic stabilization of two-input linear systems. In *Proceedings of the 10th Mediterranean Conference on Control and Automation*.

[Elia, 2004] Elia, N. (2004). When Bode meets Shannon : Control-oriented feedback communication schemes. *IEEE Transactions on Automatic Control*, 49(9).

[Elia and Mitter, 2001] Elia, N. and Mitter, S. (2001). Stabilization of linear systems with limited information. *IEEE Transactions on Automatic Control*, 46(9):1384–1400.

[Fagnani, 2004] Fagnani, F. (2004). Chaotic quantized feedback stabilizers: The scalar case. *Commun. Inf. Syst.*, 4(1):53–72.

[Fagnani and Zampieri, 2003] Fagnani, F. and Zampieri, S. (2003). Stability analysis and synthesis for scalar linear systems with a quantized feedback. *IEEE Trans. Autom. Control*, 48(9):1569–1584.

[Fagnani and Zampieri, 2004] Fagnani, F. and Zampieri, S. (2004). Quantized stabilization of linear systems: Complexity versus performances. *IEEE Transactions on Automatic Control*, 49(9):1534–1548.

[Fagnani and Zampieri, 2005] Fagnani, F. and Zampieri, S. (2005). A symbolic dynamics approach to performance analysis of quantized feedback systems: The scalar case. *SIAM J. Control Optim.*, 44(3):816–866.

[Fernando and Nicholson, 1981] Fernando, K. and Nicholson, H. (1981). Degree of controllability due to individual inputs. *IET Electronic Letters*, 17(9).

[Freudenberg et al., 2007] Freudenberg, J., Middleton, R., and Braslavsky, J. (2007). Stabilization with disturbance attenuation over a gaussian channel. In *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans,LA,USA.

[Georges, 1995] Georges, D. (1995). The use of observability and controllability gramians or functions for optimal sensors and actuator location in finite-dimensional. In *Proceedings of the 34th IEEE Conference on Decision and Control*, New Orleans, LA, USA.

[Gish and Pierce, 1968] Gish, H. and Pierce, J. (1968). Asymptotically efficient quantizing. *IEEE Transactions on Information Theory*, 14(5):676–683.

[Goodwin et al., 2005] Goodwin, G., Seron, M., and De Doná, J. (2005). *Constrained Control and Estimation : An Optimization Approach*. Springer Verlag.

[Gray and Neuhoff, 1998] Gray, R. and Neuhoff, D. (1998). Quantization. *IEEE Transactions on Information Theory*, 44(6).

[Guo et al., 2005] Guo, D., Shamai(Shitz), S., and Verdú, S. (2005). Mutual information and minimum mean-square error in Gaussian channels. *IEEE Transactions on Information Theory*, 51(4).

[Gupta et al., 2006] Gupta, V., Dana, A., Murray, R., and Hassibi, B. (2006). On the effect of quantization on performance at high rates. In *Proc. of American Control Conference*, Minneapolis, Minnesota, USA.

[Haykin, 2001] Haykin, S. (2001). *Communication Systems, 4ed*. John Wiley and Sons.

[Hespanha et al., 2007] Hespanha, J., Naghshtabrizi, P., and Xu, Y. (2007). A survey of recent results in networked control systems. *Proc. of the IEEE, Spl. Issue on NCS*, 95(1):138–162.

[Hristu, 1999] Hristu, D. (1999). *Optimal control with limited communication*. PhD thesis, Division of Engineering and Applied Sciences, Harvard University.

[Hristu, 2007] Hristu, D. (2007). On the period of communication policies for networked control systems, and the question of zero-order holding. In *Proceedings of the 46st IEEE Conference on Decision and Control*, New Orleans, USA.

[Iglesias et al., 1990] Iglesias, Mustafa, and Glover (1990). Discrete-time H_∞ controllers satisfying a minimum entropy criteria. *Systems and Control Letters*, 14(4):275–286.

[Iglesias, 2001] Iglesias, P. (2001). Trade-offs in linear time-varying systems : An analogue of Bode's sensitivity integral. *Automatica*, 37(10):1541–1550.

[Iglesias, 2002] Iglesias, P. (2002). Logarithmic integrals and system dynamics : An analogue of Bode's sensitivity integral for continuous-time time-varying systems. *Linear Alg. Appl.*, 343–344(2):451–471.

[Ionete and Çela, 2006] Ionete, C. and Çela, A. (2006). Structural properties and stabilization of NCS with medium access constraints.

[Jaynes, 1957] Jaynes, E. (1957). Information theory and statistical mechanics. *Physics Review*, 106.

[Kalman, 1960] Kalman, R. (1960). On the general theory of control systems. In *Proc. of the First International Congress on Automatic Control*, Moscow.

[Kalman et al., 1963] Kalman, R., Ho, Y., and Narendra, K. (1963). Controllability of linear dynamical systems. *Contributions to Differential Equations*, 1(2):189–213.

[Lee et al., 2005] Lee, K., Lee, S., and Lee, M. (2005). Qos-based remote control of networked control systems via profibus token passing protocol. *IEEE Transactions on Industrial Informatics*, 1(3):183–191.

[Liberzon, 2003] Liberzon, D. (2003). On stabilization of linear systems with limited information. *IEEE Trans. Autom. Control*, (2):304–307.

[Liou and Ray, 1991a] Liou, L. and Ray, A. (1991a). A stochastic regulator for integrated communication and control systems: Part-I - formulation of control law. *ASME Journal of Dynamic Systems, Measurement and Control*, 113(4):604–611.

[Liou and Ray, 1991b] Liou, L. and Ray, A. (1991b). A stochastic regulator for integrated communication and control systems: Part-II - numerical analysis and simulation. *ASME Journal of Dynamic Systems, Measurement, and Control*, 113(4):612–619.

[Liu and Elia, 2004] Liu, J. and Elia, N. (2004). Quantized feedback stabilization of nonlinear affine systems. *International Journal of Control*, 77(3):239–249.

[Liu and Elia, 2006] Liu, J. and Elia, N. (2006). Convergence of fundamental limitations in information, estimation,and control. In *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego,CA,USA.

[Lloyd, 1988] Lloyd, S. (1988). *Black Holes, Demons and the Loss of Coherence : How complex systems get information and what they do with it*. PhD thesis, The Rockefeller University, USA.

[Luck and Ray, 1990] Luck, R. and Ray, A. (1990). An observer-based compensator for distributed delays. *Automatica*, 26(5):903–908.

[Martí et al., 2004] Martí, P., Yépez, J., Velasco, M., Villà, R., and Fuertes, J. (2004). Managing quality-of-control in network-based control systems by controller and message scheduling co-design. *IEEE Transactions on Industrial Electronics*, 51(6):1159–1167.

[Martins et al., 2007] Martins, N., Dahleh, M., and Doyle, J. (2007). Fundamental limitations of disturbance attenuation in the presence of side information. *IEEE Trans. Autom. Control*, 52(1):56–66.

[Martins and Daleh, 2005] Martins, N. and Daleh, M. (2005). Fundamental limitations of performance in the presence of finite capacity feedback. In *Proc. of American Control Conference*.

[Martins and Daleh, 2008] Martins, N. and Daleh, M. (2008). Feedback control in the presence of noisy channels: Bode-like fundamental limitations of performance. *IEEE Trans. on Automatic Control*.

[Massey, 1990] Massey, J. (1990). Causality, feedback and directed information. In *Proceedings of the International Symposium on Information Theory and its Applications*, Waikiki,Hawaii,USA.

[Matveev and Savkin, 2004] Matveev, A. and Savkin, A. (2004). The problem of LQG optimal control via a limited capacity communication channel. *Sys. Control Lett.*, 53(1):51–64.

[Mehta et al., 2006] Mehta, P., Vaidya, U., and Banaszuk, A. (2006). Markov chains, entropy, and fundamental limitations in nonlinear stabilization. In *Proc. of the 45th IEEE Conference on Decision & Control*, San Diego, USA.

[Middleton, 1960] Middleton, D. (1960). *An Introduction to Statistical Communication Theory*. McGraw-Hill.

[Minero et al., 2009] Minero, P., Franceschetti, M., Dey, S., and Nair, G. (2009). Data rate theorem for stabilization over time-varying feedback channels. *IEEE Trans. on Automatic Control*, 54(2):243–255.

[Mitra, 1969] Mitra, D. (1969). W-matrix and the geometry of model equivalence and reduction. *Proc. of the IEE*, 116(6):1101–1106.

[Mitter, 2000] Mitter, S. (2000). Control with limited information. In *Proc. of International Symposium on Information Theory*, Italy.

[Mitter and Newton, 2005] Mitter, S. and Newton, N. (2005). Information and entropy flow in the Kalman-Bucy filter. *J. of Stat. Phys.*, 118(1-2):145–176.

[Montestruque and Antsaklis, 2002a] Montestruque, L. and Antsaklis, P. (2002a). Model-based networked control systems: Necessary and sufficient conditions for stability. In *Proc. of 10th Mediterranean Contol Conference*.

[Montestruque and Antsaklis, 2002b] Montestruque, L. and Antsaklis, P. (2002b). State and output feedback control in model-based networked control systems. In *Proc. of IEEE Conference on Decision and Control*.

[Montestruque and Antsaklis, 2003] Montestruque, L. and Antsaklis, P. (2003). On the model-based control of networked systems. *Automatica*, 39(10):1837–1843.

[Montestruque and Antsaklis, 2004] Montestruque, L. and Antsaklis, P. (2004). Stability of model based network control systems with time varying transmission times. *IEEE Trans. on Automatic Control*, 49(9).

[Moore, 1981] Moore, B. (1981). Principal component analysis in linear systems: Controllability, observability, and model reduction. In *IEEE Transaction on Automatic Control*, volume AC-26, New Orleans, LA, USA.

[Murray et al., 2003] Murray, R., Åström, K., Boyd, S., Brockett, R., and Stein, G. (2003). Future directions in control in an information-rich world. *IEEE Control Systems Magazine*, 23(2):20–33.

[Mustafa and Glover, 1990] Mustafa, D. and Glover, K. (1990). *Minimum Entropy H_∞ Control*. Springer-Verlag, Lecture Notes in Control Information Sciences, Vol. 146, Berlin, Heidelberg.

[Nair et al.,] Nair, G., Dey, S., and Evans, R. Infimum data rates for stabilising Markov jump linear systems. In *Proc. of 42nd IEEE Conference on Decision & Control*.

[Nair and Evans, 2000] Nair, G. and Evans, R. (2000). Communication-limited stabilization of linear systems. In *Proc. of IEEE Conference on Decision and Control*.

[Nair and Evans, 2003] Nair, G. and Evans, R. (2003). Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica*, 39(4):585–593.

[Nair and Evans, 2004] Nair, G. and Evans, R. (2004). Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM J. Control Optim.*, 43(2):413–436.

[Nair et al., 2004] Nair, G., Evans, R., Mareels, I., and Moran.W. (2004). Topological feedback entropy and nonlinear stabilization. *IEEE Trans. on Automatic Control*, 49(9):1585–1597.

[Nair et al., 2007] Nair, G., Fagnani, F., Zampieri, S., and Evans, R. (2007). Feedback under data rate constraints: An overview. *Proc. IEEE, Spl. Issue on Emerg. Tech. NCS*, 95(1):108–137.

[Nilsson, 1998] Nilsson, J. (1998). *Real-Time Control Systems with Delays*. PhD thesis, Department of Automatic Control, Lund Institute of Technology, Sweden.

[Okano et al., 2008] Okano, K., Hara, S., and Ishii, H. (2008). Characterization of a complementary sensitivity property in feedback control: An information theoretic approach. In *Proc. of the 17th IFAC-World Congress*, Seoul, Korea.

[Petersen and Savkin, 2001] Petersen, I. and Savkin, A. (2001). Multi-rate stabilization of multivariable discrete-time linear systems via a limited capacity communication channel. In *Proc. of the 40th IEEE Conf. Dec. Control*, pages 304–309.

[Phat et al., 2004] Phat, V., Jiang, J., Savkin, A., and Petersen, I. (2004). Robust stabilization of linear uncertain discrete-time systems via a limited capacity communication channel. *Systems and Control Letters*, 53(5):347–360.

[Picasso et al., 2002] Picasso, B., Gouaisbaut, F., and Bicchi, A. (2002). Construction of invariant and attractive sets for quantized-input linear systems. In *Proc. 41st IEEE Conference on Decision & Control*, pages 824–829.

[Plastino et al., 1998] Plastino, A., Casas, M., and Plastino, A. (1998). Fisher's information, Kullback's measure, and H-theorems. *Physics Letters A*, 24(6):498–504.

[Qiu and Chen, 1999] Qiu, L. and Chen, T. (1999). Multirate sampled-data systems : all H_∞ suboptimal controllers and the minimum entropy controller. *IEEE Transactions on Automatic Control*, 44(3):537–550.

[Rao, 1965] Rao, C. (1965). *Linear Statistical Inference and its Applications*. John Wiley & sons.

[Renyi, 1959] Renyi, A. (1959). On the dimension and entropy of probability distributions. *Acta Math. Acad. Sci. Hungar.*, 10:193–215.

[Roy and Çela, 2007] Roy, P. and Çela, A. (2007). A survey on control under communication constraints. In *Proc. of the 6th International RoEduNet Conference*, University of Craiova, Romania.

[Roy et al., 2009a] Roy, P., Çela, A., and Hamam, Y. (2009a). An information-theoretic view of control. In *Proc. of the 6th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2009)*, Milan,Italy.

[Roy et al., 2009b] Roy, P., Çela, A., and Hamam, Y. (2009b). On the relation of FIM and controllability gramian. In *Proc. of IEEE International Symposium on Industrial Embedded Systems (SIES 2009)*, EPFL, Lausanne, Switzerland.

[Rugh, 1996] Rugh, W. J. (1996). *Linear System Theory*. Prentice Hall.

[Sahai, 2001] Sahai, A. (2001). *Anytime Information Theory*. PhD thesis, Massachusetts Institute of Technology.

[Sandberg and Bernhardsson, 2005] Sandberg, H. and Bernhardsson, B. (2005). A Bode sensitivity integral for linear time-periodic systems. *IEEE Transactions on Automatic Control*, 50(12).

[Saridis, 1988] Saridis, G. (1988). Entropy formulation of optimal and adaptive control. *IEEE Transactions on Automatic Control*, 33(8).

[Savkin, 2006] Savkin, A. (2006). Analysis and synthesis of networked control systems: Topological entropy, observability, robustness and optimal control. *Automatica*, 42:51–62.

[Seron et al., 1997] Seron, M., Braslavsky, J., and Goodwin, G. (1997). *Fundamental Limitations in Filtering and Control*. Springer Verlag.

[Shannon, 1948] Shannon, C. (1948). The mathematical theory of communication. *Bell Systems Tech. J.*, 27:379–423,623–656.

[Skelton et al., 1998] Skelton, R., Iwasaki, T., and Grigoriadis, K. (1998). *A Unified Algebraic Approach to Linear Control Design*. Taylor & Francis.

[Sontag, 1990] Sontag, E. (1990). *Mathematical Control Theory*. Springer-Verlag, New York, USA.

[Stam, 1959] Stam, A. (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *J. of Information and Control*, 2(2):101–112.

[Tatikonda, 2000] Tatikonda, S. (2000). *Control Under Communication Constraints*. PhD thesis, Massachusetts Institute of Technology.

[Tatikonda and Mitter, 2004] Tatikonda, S. and Mitter, S. (2004). Control under communication constraints. *IEEE Trans. Autom. Control*, 49(7):1056–1068.

[Tatikonda et al., 2004] Tatikonda, S., Sahai, A., and Mitter, S. (2004). Stochastic linear control over a communication channel. *IEEE Transactions on Automatic Control*, 49(9):1549–1561.

[Tsumura and Maciejowski, 2003] Tsumura, K. and Maciejowski, J. (2003). Stability of SISO control systems under constraints of channel capacities. In *Proc. 42nd IEEE Conference on Decision & Control*.

[Water and Willems, 1981] Water, H. and Willems, J. (1981). The certainty equivalence property in stochastic control theory. *IEEE Transactions on Automatic Control*, 26(5):1080–1087.

[Witsenhausen, 1971] Witsenhausen, H. (1971). Separation of estimation and control for discrete time systems. *Proceedings of the IEEE*, 59(11):1557–1566.

[Wong and Brockett, 1999] Wong, W. and Brockett, R. (1999). Systems with finite communication bandwidth constraints II: Stabilization with limited information feedback. *IEEE Transactions on Automatic Control*, 44(5):1049–1053.

[Wu and Jonckheere, 1992] Wu, B. and Jonckheere, E. (1992). A simplified approach to Bode's theorem for continuous-time and discrete-time systems. *IEEE Transactions on Automatic Control*, 37(11).

[Yüksel and Başar, 2006] Yüksel, S. and Başar, T. (2006). Minimum rate coding for LTI systems over noiseless channels. *IEEE Transactions on Automatic Control*, 51(12):1878–1887.

[Zador, 1966] Zador, P. (1966). Topics in the asymptotic quantization of continuous random variables. *Bell Lab. Tech. Memo.*

[Zang and Iglesias, 2003] Zang, G. and Iglesias, P. (2003). Nonlinear extension of Bode's integral based on an information theoretic interpretation. *Systems and Control Letters*, 50:11–29.

[Zhang and Liu, 2008] Zhang, C. and Liu, X. (2008). Trade-off between the sampling rate and the data accuracy. In *Proc. of American Control Conference*, Seattle, Washington, USA.

[Zhang and Sun, 2003] Zhang, H. and Sun, Y. (2003). Information theoretic limit and bound of disturbance rejection in LTI systems : Shannon entropy and H_∞ entropy. In *Proc. of IEEE IEEE Conference on Systems, Man and Cybernetics*, volume 2, pages 1378–1383.

[Zhang and Hristu-Varsakelis, 2006] Zhang, L. and Hristu-Varsakelis, D. (2006). Communication and control codesign for networked control systems. In *Automatica*, volume 42, pages 953–958.

[Zhou, 1998] Zhou, K. (1998). *Essentials of Robust Control*. Prentice-Hall, USA.