**DOCTOR OF SCIENCES**

# The LQ-optimal control problem for invariant linear systems

Beauthier, Charlotte

*Award date:*
2011

*Awarding institution:*
University of Namur

Facultés Universitaires Notre-Dame de la Paix, Namur

Faculté des Sciences

Département de Mathématiques

# The LQ-Optimal Control Problem for Invariant Linear Systems

Thèse présentée par
**Charlotte Beauthier**
en vue de l'obtention du grade
de Docteur en Sciences

**Composition du Jury :**

Frank Callier
Denis Dochain
Daniela Iacoviello
Annick Sartenaer
Joseph Winkin (Promoteur)

Mai 2011

Rempart de la Vierge, 13
B-5000 Namur (Belgique)



Imprimé en Belgique

ISBN : 978-2-87037-714-7
Dépôt légal: D / 2011 / 1881 / 19

Facultés Universitaires Notre-Dame de la Paix
Faculté des Sciences
rue de Bruxelles, 61, B-5000 Namur, Belgium

Facultés Universitaires Notre-Dame de la Paix
Faculté des Sciences
Rue de Bruxelles, 61, B-5000 Namur, Belgium

**Le problème de commande optimale linéaire quadratique pour les systèmes linéaires invariants**

par Charlotte Beauthier

**Résumé :** Ce travail a pour objet l'étude du problème de commande optimale au sens linéaire quadratique (LQ) pour des systèmes linéaires avec contraintes d'inégalité affines sur les trajectoires d'état et/ou d'entrée, et en particulier pour des systèmes linéaires entrée/état-invariants. L'étude de ces systèmes est motivée notamment par le problème de coexistence dans un modèle de chémostat où, pour des raisons biologiques, il est important de chercher à forcer les trajectoires d'état et d'entrée de rester dans un cône. Des conditions nécessaires et suffisantes d'optimalité sont établies pour le problème LQ invariant entrée/état en utilisant le principe du maximum avec contraintes sur l'état et l'entrée et à l'aide de l'admissibilité de la solution du problème LQ standard. Des résultats similaires et spécifiques sont obtenus pour le problème LQ appliqué aux systèmes positifs, qui sont caractérisés par l'invariance de l'orthant non négatif de l'espace d'état. Les méthodes développées dans cette thèse sont appliquées au modèle de chémostat via l'étude des systèmes non linéaires localement entrée/état-invariants. Les principaux résultats de ce travail sont illustrés par des exemples numériques.

**The LQ-optimal control problem for invariant linear systems**

by Charlotte Beauthier

**Abstract :** This work is concerned with the study of the linear quadratic (LQ) optimal control problem for linear systems with affine inequality constraints on the state and/or the input trajectories, and in particular for input/state-invariant linear systems. The study of such systems is motivated notably by the coexistence problem in a chemostat model where, for biological reasons, it is meaningful to aim at forcing the state and the input trajectories to remain in a cone. Necessary and sufficient optimality conditions are established for the input/state-invariant LQ problem by using the maximum principle with state and input constraints and by using the admissibility of the solution of the standard LQ problem. Similar and specific results are obtained for the particular LQ problem for positive systems, which are characterized by the invariance of the nonnegative orthant of the state space. The methods developed in this thesis are applied to the chemostat model via the study of locally positively input/state-invariant nonlinear systems. The main results of this work are illustrated by some numerical examples.

Thèse de doctorat en Sciences mathématiques (Ph.D. thesis in Mathematics)
Date : 12-05-2011
Département de Mathématiques
Promoteur (Advisor) : Prof. J. J. WINKIN

# *Remerciements*

Cette thèse n'aurait pu voir le jour sans l'aide de nombreuses personnes qui m'ont guidée, soutenue et encouragée tout au long de mon doctorat.

Tout d'abord, je tiens à remercier mon promoteur, Joseph Winkin, pour son encadrement et ses conseils tout au long de ces années. Son soutien, son expérience ainsi que ses remarques pertinentes ont favorisé l'aboutissement de ce projet de recherche. Je le remercie également de m'avoir toujours permis d'orienter notre recherche selon mes affinités ainsi que de m'avoir donné l'opportunité de participer à plusieurs conférences internationales enrichissantes, au cours desquelles j'ai pu présenter mes travaux.

Je souhaite aussi adresser un merci particulier à Denis Dochain, professeur ordinaire au pôle en ingénierie mathématique (INMA), à l'université catholique de Louvain (UCL), avec qui j'ai pu collaborer sur une application dans le domaine de la biologie microbienne. Merci de m'avoir guidée dans ce domaine et d'avoir toujours pris le temps de répondre à mes questions. Cette collaboration me fut très enrichissante. Je remercie également Frank Callier, professeur émérite du département de Mathématiques (FUNDP), d'une part, pour l'enseignement prodigué en contrôle durant mes années de licence en mathématiques au département et d'autre part, d'avoir accepté de faire partie de mon jury de DEA et de thèse de doctorat. Merci pour tous les conseils et remarques intéressantes formulées lors des lectures de ces deux manuscrits. Je souhaite aussi remercier Daniela Iacoviello, professeur à l'université Sapienza de Rome, et Annick Sartenaer, professeur ordinaire au département de Mathématiques (FUNDP), d'avoir accepté de faire partie du jury de cette thèse. Je tiens également à tous les remercier pour leurs remarques constructives et leurs conseils pertinents qui ont permis d'améliorer ce manuscrit.

C'est avec une pensée particulière que je remercie chaleureusement celles qui m'ont accompagnées au quotidien : mes collègues de bureau, Emilie Wanufelle et Martine DeVleeschouwer. Merci pour leur amitié, leur bonne humeur, leur écoute, leurs encouragements et leurs conseils pendant ces différentes années dans le même bureau.

Mon merci suivant s'adresse à tous mes collègues, anciens ou actuels, avec qui j'ai partagé de bons moments et sans qui ces années auraient été sans nul doute fort différentes. Pour ces agréables souvenirs et pour leur bonne humeur et leur amitié, je voudrais remercier les "copains de midi", notamment, Florent Deleflie, Sandrine d'Hoedt, Julien Dufey, Katia Demaseure, Caroline Sainvitu et Melissa Weber, ainsi que les amis du whist, ceux du vendredi soir, Nicolas Delsate, Benoit Noyelles et Dimitri Tomanos pour ces soirées inoubliables ; mais également ceux du temps de midi, anciens et nouveaux, Joffray Baune, Jehan Boreux, Audrey Compere, Charles Hubaux, Patrick Laloyaux, Anne-Sophie Libert, Charlotte Tannier, Emilie Verheylewegen et Sebastian Xhonneux pour toutes ces parties mémorables partagées ensemble qui m'ont permis de décompresser. Un merci particulier aux secrétaires, Pascale Hermans et Martine Van Caenegem pour leur aide précieuse dans les démarches administratives. Mes remerciements vont aussi vers tous les autres membres du département de Mathématiques que j'ai pu cotoyer durant ces années pour leur accueil et leur convivialité.

Merci aussi à tous mes amis, mathématiciens ou non, qui m'ont aidée, par tous ces moments de détente passés en leur compagnie.

Je tiens également à remercier toute ma famille pour son soutien et ses encouragements pendant ces années. Un merci particulier à ma soeur, Catherine, pour ses relectures et son aide lors des préparations de mes premières présentations en anglais.

Mon dernier remerciement, mais non le moindre, s'adresse à mon cher mari, Fabrice, d'avoir toujours été présent à mes côtés, de m'avoir toujours écoutée et soutenue et d'avoir toujours fait tout ce qui était possible pour me rendre heureuse. Merci pour sa présence, son amour, ses encouragements et sa confiance en moi qui m'ont aidée à persévérer dans les moments difficiles et qui ont été indispensables pour mener à bien ce projet. Pour tout cela et pour tout le reste, je le remercie de tout mon coeur.

Enfin, j'adresse également un dernier remerciement à mon petit garçon, ma petite merveille, Léopold, qui me comble de bonheur jour après jour. Même s'il est trop petit pour s'en rendre compte, sa présence, ses sourires et sa joie de vivre m'ont beaucoup aidée durant cette dernière année de thèse.

Merci à tous ceux et toutes celles qui ont contribué, de près ou de loin, à l'aboutissement de ce travail. A vous tous, merci pour tout ce que vous m'avez apporté.

Charlotte.

# Contents

# Introduction

## A world involved in systems and control

Control theory is an interdisciplinary branch of science which has its origin in engineering and mathematics and which deals with influencing the behavior of dynamical systems. It focuses on the modeling of a diverse range of dynamical systems (e.g. mechanical, biological or economical systems) and the design of controllers that will cause these systems to behave in a desired manner.

Let us consider, for example, a car with the system of cruise control, which is a device designed to maintain the vehicle's speed at a constant desired value provided by the driver. The *control* (or *input*) is the engine's throttle position which determines the power of the engine and the *output* is the car's speed. A first way to implement cruise control is simply to lock the engine's throttle position when the driver engages cruise control. This is called an *open-loop* design because no measurement of the *output* (the car's speed) is used to modify the *control* (or *input*) (the engine's throttle position). As a result, the controller can not compensate for changes acting on the car, like a change in the slope of the road. In a *closed-loop* design, a sensor measures the output (the car's speed) and transmits the data to a controller which adjusts the input (the engine's throttle position) as necessary to maintain the desired output (match the car's speed to the desired speed). Also, the (minimal or maximal) speed limit on the road can be seen as *bound constraints* on a *state* component (the car's speed).

More formally, in *systems and control* theory, one is interested in governing the *state* of a dynamical system by using *control*. The dynamical behavior of the system is the manner in which the state changes under the influence of the control and is often described by an ordinary differential equation. *Optimal control* deals with the problem of finding a control law for a given system such that a certain optimality criterion is achieved under constraints. An optimal control problem includes a cost functional that is a function of the state and control variables : the objective of optimal control theory is to determine a control law that will cause a process to satisfy the physical constraints and at the same time, minimize (or maximize) a cost functional, i.e. a performance criterion, see e.g. [HSV95]. For example, if we want to keep the car's speed, denoted by $x(t)$, near a constant value $\alpha$ on a time interval $[0,\, t_f]$, then this question can be formalized as the problem of finding a control (engine's throttle position), denoted by $u(t)$, which minimizes the cost functional $\int_0^{t_f} (x(t) - \alpha)^2 \,\mathrm{dt}$.

The most common optimal control problem is the *linear quadratic* (LQ) optimal control problem, see e.g. [KS72, CD91]. This problem consists of minimizing a *quadratic* cost functional subject to *linear* dynamical constraints described by a set of linear differential equations. One of the most salient features of the LQ control is the fact that it is of *state-feedback* type. That means that the optimal control $u$ can be written in terms of linear combinations of the state components $x$, i.e. $u = K\,x$, where $K$ is called the *feedback* matrix.

The principle of feedback is simple : feedback is a process that is looped back to control a system within itself. It is the process in which part of the output of a system is returned back to its input in order to regulate its further output. The term feedback can also be seen as the situation in which two (or more) dynamical systems are connected together such that each system influences the other and their dynamics are thus strongly coupled. A system is said to be a *closed-loop* system if the compound systems are interconnected in a cycle (Figure 2). If the interconnection is broken, the system is said to be an *open-loop* system (Figure 1).

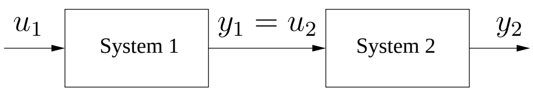



Figure 1: Open design.

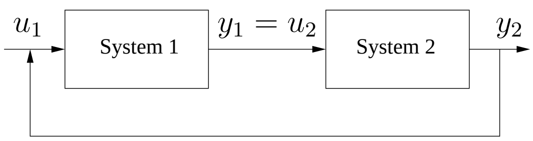



Figure 2: Closed-loop design.

Feedback has many interesting properties that can be exploited in designing systems : notably, feedback allows a system to be insensitive both to external disturbances and to variations in its individual elements, see [AM90]. Another use of feedback is to change the dynamics of a system. Through feedback, one can alter the behavior of a system to meet the needs of an application : for example, systems that are unstable can be stabilized.

Let us mention that another well-known optimal control problem deals with *model predictive control* (MPC), see e.g. [Zhe10] and [AZ00]. MPC is a control strategy in which the applied input is determined on-line at each sampling instant by the solution of an open-loop optimal control problem using the current (estimated) state as initial state. The solution of the optimization problem yields an optimal input signal from which only the first part is implemented until the next measurement becomes available.

This thesis is devoted to the LQ-optimal control problem for linear systems with affine inequality constraints on the state and/or the input trajectories, and in particular for *input/state-invariant* linear systems, which are characterized by the fact that the input and the state trajectories should remain in a cone. The study of such systems is motivated notably by the problem of *coexistence of species* in a *chemostat* (i.e. a continuous stirred tank reactor). The concept of coexistence of species means that the concentration of the species should remain strictly positive. From a mathematical point of view, the study of the problem of coexistence of species can be performed by means of input/state-invariant systems. Indeed, for physical or biological reasons, it is meaningful to aim at forcing the state and/or the input trajectories of such systems to remain in a cone. On the other hand, dynamical models of many biological and pharmacological processes, such as metabolic systems or biochemical reactions, are derived from mass and energy balance considerations that involve states whose values are nonnegative. Hence it follows from physical considerations that the state trajectories of such models should remain in the nonnegative orthant of the state space for nonnegative initial conditions. This motivates the study of the LQ-optimal control problem for the particular class of *positive* systems, which are characterized by the invariance of the nonnegative orthant of the state space.

In the literature, the concept of invariance of linear systems is an important topic in systems and control, (see e.g. [Bla99]), as well as the positivity of linear systems, see e.g. [FR00] and [HCH10] for an overview. This class of systems is very interesting for the study of applications, see e.g. [HCH10, God83, Van08] and there are many contributions which are devoted to such systems, see e.g. [BF04, HCH10]. In the framework of the LQ problem (see [CD91, AM90]), the constrained problem has already been studied when only considering nonnegative constraints, either on the state or on the input (see e.g. [HVS98] for the LQ problem with positive controls, [Ka 02] for the minimal energy positive control problem for positive systems and the recent book [HCH10] and the references therein). There is also a large literature devoted to modifying the chemostat model to ensure coexistence of the organisms, see e.g. [BHW85, Smi95, SFA79, Hsu80] and [DS03] where feedback control of the dilution rate is studied. See [SW95] for an overview on the chemostat model. The interest of applying an LQ control to a chemostat model is to benefit of its specific properties in order to get a model for which the coexistence of species is guaranteed.

The contributions of this thesis with respect to the literature are summarized after the concluding section.

# Structure of the document

First of all, in this work, the *continuous time* case is considered in all chapters except for the last one, which is devoted to similar and also specific results in *discrete time*. This thesis is then divided in three main parts for the continuous time case.

The first part describes properties of time-varying and time-invariant input/state-invariant linear systems (Chapter 1), and well-known properties of positive linear systems (Chapter 2).

The second part deals with the study of the input/state-invariant linear quadratic problem, first in *finite horizon* (Chapters 3 and 4), and next in *infinite horizon* (Chapters 5 to 7). In Chapter 3, optimality conditions are established for the input/state-invariant LQ problem, which are based on the maximum principle (see [HSV95]) and on the admissibility of the solution of the standard LQ problem. Similar results have been obtained for the particular LQ problem for positive systems, which is studied in Chapter 4. Moreover, specific results are stated in terms of the matrix solution of the Riccati differential equation (RDE) and the particular problem of minimal energy control with penalization of the final state is also studied. This chapter is completed by illustrative numerical examples. In the second half of this part (Chapters 5 to 7), the input/state-invariant LQ problem is studied for the infinite horizon case. Chapter 5 briefly analyzes a receding horizon approach. Criteria for the existence of a solution to the positive LQ problem are established in Chapter 6, by using a Newton-type iterative scheme. In addition, positivity criteria are stated in terms of the solution of the algebraic Riccati equation (ARE) and in terms of the Hamiltonian matrix $H$. Chapter 7 is devoted to the inverse input/state-invariant LQ problem which consists of finding, for a fixed invariant stabilizing matrix $K$, weighting matrices such that the feedback control is optimal for the resulting LQ problem.

The third part is devoted to the application of the LQ problem to *locally positively invariant* nonlinear systems. First, properties of a locally positively invariant nonlinear system together with the link with its linear approximation around an equilibrium, are described (Chapter 8). Then, the application of the LQ problem in order to solve the problem of coexistence of species in a chemostat model is studied (Chapter 9). In the latter, we first describe the framework of the chemostat model and the problem of coexistence of species which are therein in competition for one substrate. The theory developed so far for the input/state-invariant LQ problem is then applied to guarantee the local positive invariance of the chemostat model, which is described by a nonlinear system. Numerical simulations have also been performed to complete this study.

Finally, in the last chapter, several results are stated for positive systems in *discrete time* and the corresponding positive LQ problem in finite horizon is studied. These results are similar to the ones obtained in the continuous time case. Moreover, a specific result is derived for the particular class of monomial systems, which can be seen as inverse time positive systems. At last, an algorithm, both in a vector and a matrix form, is developed by using Hamiltonian systems. The main results of this chapter are illustrated by some numerical examples.

We conclude this thesis by summarizing our approach and the obtained results and by suggesting some perspectives for future work. A summary of our contributions and tables containing the main notations and abbreviations used in this thesis can be found after the conclusion.

Notice that, in the sequel, definitions, theorems (including lemmas, corollaries and propositions), remarks and examples are numbered with respect to the current section for a given chapter. For example, Theorem 10.2.6 denotes the 6th theorem in the 2nd section of Chapter 10.

# A. Continuous Time Case

# Part I

# Invariant Linear Systems

In this first part, the theory of invariant continuous-time linear systems is presented. Conditions for the invariance property are stated in terms of the matrices defining the system dynamics. The concept of invariance of linear systems means that, under some conditions on the initial state and the input trajectories, their state trajectories remain in a (shifted) cone. This is an important topic in system theory, see e.g. [Bla99] and [BNS89]. Furthermore, the stability and the stabilizability of such systems are studied. Finally, similar results and some additional results are presented for the particular class of positive systems. These systems encompass dynamical models where all the variables should remain nonnegative for any nonnegative initial condition and for any nonnegative input trajectory. This class of systems is much studied in the literature and there is a large class of applications in this field. Some typical examples of positive systems are economics models, chemical processes, compartmental systems and biological systems. A lot of theoretic problems have already been investigated for positive systems. See e.g. [FR00] and [HCH10] for an overview of the state of the art in this topic.

# Chapter 1

# Invariant Linear Systems

Set invariance is an important and extensively studied topic in systems and control, see e.g. [Bla99]. Invariant sets play an important role for example in constrained control, robustness analysis, synthesis and optimization. This chapter is devoted to the study of invariant linear systems, i.e. systems where, under some conditions on the initial state and the input trajectories, the state trajectories remain in a (shifted) cone, see e.g. [Bla99] and [BNS89, Chapter 4]. For these systems, characterization of the invariance can be described in terms of the matrices defining the system dynamics. The stability of such systems is also studied. Moreover the problem of invariant stabilization is studied, which consists of finding a stabilizing state feedback which ensures the invariance property for the resulting closed-loop system. The next chapter is devoted to the particular class of positive systems.

## 1.1 Main concepts and results

### 1.1.1 Invariant LTV systems

Let $X$ and $Y$ be matrices in $\mathbb{R}^{p\times q}$. The property that, for all $i = 1,\ldots,p$ and for all $j = 1,\ldots,q$, $x_{ij} \geq y_{ij}$, ($x_{ij} > y_{ij}$, respectively), is denoted by $X \geq Y$, ( $X \gg Y$, respectively). Finally, $X > Y$ means that $X \geq Y$ and $X \neq Y$.

**Definition 1.1.1** *Let $M$ be a matrix in $\mathbb{R}^{p\times q}$.*

- *$M$ is said to be* **nonnegative** *if $M \geq 0$.*
- *$M$ is said to be* **strictly positive** *if $M \gg 0$.*
- *$M$ is said to be* **positive** *if $M > 0$.*
- *A square matrix $M \in \mathbb{R}^{p\times p}$ is said to be* **Metzler** *if all its off-diagonal components are nonnegative, i.e.*

  $$\forall\, i,\ j \in \{1,\ldots,p\} \text{ such that } i \neq j,\ m_{ij} \geq 0,$$

  *i.e. $M + \alpha\, I_p$ is a nonnegative matrix for some $\alpha \in \mathbb{R}$, (see Appendix A), where $I_p$ denotes the identity matrix of dimension $p$.*

In particular, these notations and definitions obviously apply to the case $q = 1$, i.e. to vectors $x \in \mathbb{R}^p$.

Consider the following linear time-varying (LTV) homogeneous system, for $t \in [t_0,\, t_f]$,

$$\dot{x}(t) = A(t)\, x(t), \quad x(t_0) = x_0, \tag{1.1}$$

where the state $x(t) \in \mathbb{R}^n$, $A(t) \in \mathbb{R}^{n \times n}$ is a piecewise continuous real matrix function, $x_0 \in \mathbb{R}^n$ denotes any fixed initial state and $[t_0,\, t_f]$ is an arbitrarily fixed time interval. A first property of the fundamental matrix of such system is the equivalence between its nonnegativity and the Metzler property of the matrix $A(t)$.

**Lemma 1.1.1** *The matrix $A(t)$ of system (1.1) is a Metzler matrix for all $t \in [t_0,\, t_f]$ if and only if $\Phi(t,\, t_0)$ is nonnegative for all $t \in [t_0,\, t_f]$ where $\Phi(t,\, t_0)$ is called the* **fundamental matrix** *and satisfies the following homogeneous equation :*

$$\frac{\partial \Phi}{\partial t}(t,\, t_0) = A(t)\, \Phi(t,\, t_0), \quad \forall\, t \in [t_0,\, t_f], \tag{1.2}$$

*with the initial condition $\Phi(t_0,\, t_0) = I_n$, see [CD91, pp. 10-11].*

**Proof :**

Necessity : Consider $\alpha(t) \leq \min\{a_{ii}(t)\}_{i=1}^n$ where $\alpha(\cdot)$ is a (piecewise) continuous function. Set $\bar{A}(t) = A(t) - \alpha(t)\, I_n$. Then, $\bar{A}(t) \geq 0$. Consider the following differential equation :

$$\dot{z}(t) = \bar{A}(t)\, z(t), \quad \text{with } z(t_0) = x(t_0)$$

such that

$$z(t) = \bar{\Phi}(t,\, t_0)\, x(t_0),$$

where $\bar{\Phi}(t,\, t_0)$ is the corresponding fundamental matrix. Therefore, one has :

$$\Phi(t,\, t_0) = exp\left(\int_{t_0}^{t} \alpha(\tau)\, \mathrm{d}\tau\right) \bar{\Phi}(t,\, t_0). \tag{1.3}$$

Let

$$\bar{\Phi}(t,\, t_0) = \lim_{m \to \infty} X_m(t), \quad t \in [t_0,\, t_f], \tag{1.4}$$

where $X_m(t)$, $m \in \mathbb{N}$, denote the Picard's iterates, (see [CD91, p. 13 and pp. 471-475]), that are defined, for all $t \in [t_0,\, t_f]$, by :

$$\left\{ \begin{array}{rcl} X_0(t) & = & I_n \\ X_{m+1} & = & I_n + \displaystyle\int_{t_0}^{t} \bar{A}(\tau)\, X_m(\tau)\, \mathrm{d}\tau. \end{array} \right.$$

Observe that $X_0(t) \geq 0$ and that if $X_m(t) \geq 0$ then $X_{m+1}(t) \geq 0$. Hence,

$$\forall\, t \geq t_0, \forall\, m \,:\, X_m(t) \geq 0.$$

Consequently, by (1.4), $\bar{\Phi}(t,\, t_0) \geq 0$ and then by (1.3), $\Phi(t,\, t_0) \geq 0$ for all $t \in [t_0,\, t_f]$.

Sufficiency : Assume that for all $t \in [t_0,\, t_f]$, $\Phi(t,\, t_0) \geq 0$. Let $i \neq j$ and $h > 0$. With $e_i$ denoting the $i^{\text{th}}$ vector of the canonical basis of $\mathbb{R}^n$, one has :

$$
\begin{aligned}
0 &\leq \frac{< \Phi(t+h,\, t)\, e_i,\, e_j >}{h} \\
&= \frac{< \Phi(t+h,\, t)\, e_i,\, e_j > - < e_i,\, e_j >}{h} \\
&= \frac{< (\Phi(t+h,\, t) - I_n)\, e_i,\, e_j >}{h} \\
&= \frac{< \displaystyle\int_t^{t+h} A(\sigma)\, \Phi(\sigma,\, t)\, \mathrm{d}\sigma \; e_i,\, e_j >}{h} \\
&= \frac{1}{h} \int_t^{t+h} < A(\sigma)\, \Phi(\sigma,\, t)\, e_i,\, e_j > \, \mathrm{d}\sigma
\end{aligned}
$$

since $\Phi(t,\, t_0)$ is the solution of equation (1.2). Then, by the meanvalue theorem applied to the continuous function $< A(\sigma)\, \Phi(\sigma,\, t)\, e_i,\, e_j >$ on $[t,\, t+h]$ for $h$ sufficiently small, it follows that :

$$0 \leq < A(t)\, e_i,\, e_j > = a_{ij}(t), \qquad \text{for } i \neq j.$$

Therefore, $A(t)$ is a Metzler matrix for all $t \in [t_0,\, t_f]$.

□

Let $\bar{x} \ll 0$ be a fixed state and consider an initial condition $x(t_0) := x_0 \geq \bar{x}$. Such a fixed state $\bar{x}$ is used in Chapter 9, which is devoted to the chemostat model where several species are in competition for a single nutrient. The model is described by a nonlinear system which is studied by means of its linearization around an equilibrium $x_e := -\bar{x}$. This equilibrium state $x_e$ is assumed to be strictly positive and therefore it guarantees the coexistence of species. The results developed in the current chapter also hold for $\bar{x} = 0$, see Chapter 2 on positive systems. Now, consider the following shifted cone $C_{\bar{x}}$ :

$$C_{\bar{x}} := \{x \in \mathbb{R}^n \,:\, x \geq \bar{x}\} = \mathbb{R}^n_+ + \bar{x}. \tag{1.5}$$

**Definition 1.1.2** *The cone $C_{\bar{x}}$ is said to be* **invariant with respect to (w.r.t.) system (1.1) on $[t_0,\, t_f]$** *if $C_{\bar{x}}$ is $\Phi(t,\, t_0)$-invariant on $[t_0,\, t_f]$, i.e.*

$$\forall\, t \in [t_0,\, t_f], \quad \Phi(t,\, t_0)\, C_{\bar{x}} \subset C_{\bar{x}},$$

*or equivalently*

$$\forall\, t \in [t_0,\, t_f], \forall\, x_0 \in C_{\bar{x}}, \quad x(t) := \Phi(t,\, t_0)\, x_0 \in C_{\bar{x}}.$$

*In this case, system (1.1) is said to be* **state-invariant w.r.t. $C_{\bar{x}}$ on $[t_0,\, t_f]$**.

**Theorem 1.1.2** *The cone $C_{\bar{x}}$ is invariant w.r.t. system (1.1) on $[t_0,\, t_f]$ if and only if the following conditions hold :*

$$\forall\, t \in [t_0,\, t_f], \quad A(t) \textit{ is a Metzler matrix} \tag{1.6}$$

*and*

$$\forall\, t \in [t_0,\, t_f], \quad \Phi(t,\, t_0)\, \bar{x} \geq \bar{x} \tag{1.7}$$

**Remarks 1.1.1** *a) Observe that condition (1.7) implies that*

$$A(t_0)\,\bar{x} \geq 0. \tag{1.8}$$

*Indeed, this follows from the identity*

$$A(t_0)\,\bar{x} = \lim_{t \to t_{0+}} \frac{(\Phi(t, t_0) - I_n)\,\bar{x}}{t - t_0}.$$

*b) Condition (1.7) seems to be difficult to check. However, it is not necessary to compute $\Phi(t,\, t_0)$ for all time, but only the state trajectories from the initial condition $\bar{x}$ and not from all $x_0$. Moreover, it is shown in Theorem 1.1.5 that condition (1.7) can be translated only in terms of $A$ for linear time-invariant systems.*

**Proof of Theorem 1.1.2 :**
Necessity : Consider the initial state $x_0 := \bar{x}$. Then $x(t) = \Phi(t,\, t_0)\,x_0 = \Phi(t, t_0)\,\bar{x}$ with $x(t) \geq \bar{x}$ for all $t \in [t_0,\, t_f]$ by assumption. Therefore condition (1.7) holds. Now taking $x_0 := \alpha_j\,\bar{x}_j\,e_j \geq \bar{x}$ for $\alpha_j < 0$ gives $x(t) = \Phi(t,\, t_0)\,x_0 = \Phi(t,\, t_0)\,\alpha_j\,\bar{x}_j\,e_j$. In particular, for all $t \in [t_0,\, t_f]$ and for all $i,\, j$,

$$x_i(t) = \Phi_{ij}(t,\, t_0)\,\alpha_j\,\bar{x}_j \geq \bar{x}_i$$

or equivalently

$$\Phi_{ij}(t,\, t_0) \geq \frac{\bar{x}_i}{\bar{x}_j\,\alpha_j}, \quad \text{where } \alpha_j < 0 \text{ and } \bar{x}_i,\, \bar{x}_j < 0.$$

Hence letting $\alpha_j \to -\infty$, it follows that $\Phi_{ij}(t,\, t_0) \geq 0$ for all $t \in [t_0,\, t_f]$ and for all $i, j$. Then by Lemma 1.1.1, condition (1.6) holds.

Sufficiency : Let $t \in [t_0,\, t_f]$ and $x_0 \geq \bar{x}$ be arbitrarily fixed. Then

$$\begin{aligned} x(t) = \Phi(t,\, t_0)\,x_0 &\geq \Phi(t,\, t_0)\bar{x} && \text{since } \Phi(t,\, t_0) \geq 0 \text{ by (1.6) and Lemma 1.1.1} \\ &\geq \bar{x} && \text{by condition (1.7).} \end{aligned}$$

□

Now consider the following LTV system description denoted by $R = [A(\cdot),\, B(\cdot)]$, for $t \in [t_0,\, t_f]$ :

$$\dot{x}(t) = A(t)\,x(t) + B(t)\,u(t), \quad x(t_0) = x_0 \tag{1.9}$$

where the state $x(t) \in \mathbb{R}^n$ and the control $u(t) \in \mathcal{U}$ where $\mathcal{U}$ is the set of piecewise continuous functions from $[t_0,\, t_f]$ to $\mathbb{R}^m$, $A(t)$ and $B(t)$ are piecewise continuous real matrix functions of compatible sizes and $x_0 \in \mathbb{R}^n$ denotes any fixed initial state. In the sequel, unless otherwise stated, these conditions are assumed to hold for such systems. Consider an input $u(t) = K(t)\,x(t)$ for $t \in [t_0,\, t_f]$, where $K(\cdot) \in \mathbb{R}^{m\times n}$ is a piecewise continuous state feedback function. Therefore we consider the following LTV closed-loop system, denoted by $R = [A + B\,K(\cdot),\, 0]$, for $t \in [t_0,\, t_f]$ :

$$\dot{x}(t) = (A + B\,K(t))\,x(t), \quad x(t_0) = x_0. \tag{1.10}$$

Let an input $\bar{u} \leq 0$ be fixed. Such a fixed input $\bar{u}$ is also used in Chapter 9 as an equilibrium input of the considered nonlinear system. Consider the following shifted cone $C_{\bar{u}}$ :

$$C_{\bar{u}} := \{u \in \mathbb{R}^m \ : \ u \geq \bar{u}\} = \mathbb{R}_+^m + \bar{u}. \tag{1.11}$$

**Definition 1.1.3** *The cone $C_{\bar{u}}$ is said to be* **invariant with respect to system (1.10) on $[t_0,\, t_f]$** *if*

$$\forall\, t \in [t_0,\, t_f], \forall\, x_0 \ \textit{such that} \ u(t_0) = K(t_0)\, x_0 \in C_{\bar{u}},$$
$$u(t) := K(t)\, x(t) = K(t)\, \Phi_K(t,\, t_0)\, x_0 \in C_{\bar{u}},$$

*where $\Phi_K(t,\, t_0)$ is the fundamental matrix which satisfies the following differential equation*

$$\frac{\partial}{\partial t}\Phi_K(t,\, t_0) = (A + B\, K(t))\, \Phi_K(t,\, t_0), \quad \forall\, t \in [t_0,\, t_f]$$

*with the initial condition $\Phi_K(t_0,\, t_0) = I_n$.*
*In this case, system (1.10) is said to be* **input-invariant w.r.t. $C_{\bar{u}}$ on $[t_0,\, t_f]$**.

**Theorem 1.1.3**
*a) Let $K(t)$ be a state feedback of system (1.10) of full column rank and $m \geq n$. If the following conditions hold :*

$$\forall\, t \in [t_0,\, t_f], \quad K(t)\, \Phi_K(t,\, t_0)\, K_l^+(t_0) \geq 0 \tag{1.12}$$

*and*

$$\forall\, t \in [t_0,\, t_f], \quad K(t)\, \Phi_K(t,\, t_0)\, K_l^+(t_0)\, \bar{u} \geq \bar{u} \tag{1.13}$$

*where $K_l^+ \in \mathbb{R}^{n\times m}$ denotes the left pseudo-inverse of $K$, i.e. $K_l^+ := (K^T K)^{-1} K^T$ such that $K_l^+\, K = I_n$. Then the cone $C_{\bar{u}}$ is invariant w.r.t. system (1.10) on $[t_0,\, t_f]$.*

*b) Conversely, if the cone $C_{\bar{u}}$ is invariant w.r.t. system (1.10) on $[t_0,\, t_f]$, with a state feedback $K(t)$ of full row rank and $m \leq n$. Then the following conditions hold*

$$\forall\, t \in [t_0,\, t_f], \quad K(t)\, \Phi_K(t,\, t_0)\, K_r^+(t_0) \geq 0 \tag{1.14}$$

*and*

$$\forall\, t \in [t_0,\, t_f], \quad K(t)\, \Phi_K(t,\, t_0)\, K_r^+(t_0)\, \bar{u} \geq \bar{u} \tag{1.15}$$

*where $K_r^+ \in \mathbb{R}^{n\times m}$ denotes the right pseudo-inverse of $K$, i.e. $K_r^+ := K^T (K\, K^T)^{-1}$ such that $K\, K_r^+ = I_m$.*

**Proof :** a) Assume that (1.12) and (1.13) hold, i.e., with $V(t,\, t_0) := K(t)\, \Phi_K(t,\, t_0)\, K_l^+(t_0)$,

$$\forall\, t \in [t_0,\, t_f], \quad V(t,\, t_0) \geq 0 \text{ and } V(t,\, t_0)\, \bar{u} \geq \bar{u}.$$

Then, for all $t \in [t_0,\, t_f]$, we have, with $u(t_0) := K(t_0)\, x_0 \geq \bar{u}$ :

$$\begin{array}{rcl} u(t) & = & K(t)\, x(t) = K(t)\, \Phi_K(t,\, t_0)\, x_0 \\ & = & K(t)\, \Phi_K(t,\, t_0)\, K_l^+(t_0)\, K(t_0)\, x_0 \\ & = & V(t,\, t_0)\, K(t_0)\, x_0 \\ & \geq & V(t,\, t_0)\, \bar{u} \\ & \geq & \bar{u} \end{array}$$

b) Let $x_0 = K_r^+(t_0)\,\bar{u}$ such that $u(t_0) = K(t_0)\,x_0 = K(t_0)\,K_r^+(t_0)\,\bar{u} = \bar{u}$. Then $u(t) = K(t)\,x(t) = K(t)\,\Phi_K(t,\,t_0)\,x_0 = K(t)\,\Phi_K(t,\,t_0)\,K_r^+(t_0)\,\bar{u}$. Therefore since $u(t) \geq \bar{u}$ for all $t \in [t_0,\,t_f]$ by assumption, condition (1.15) holds. Now taking $x_0 = K_r^+(t_0)\,\alpha_j\,\bar{u}_j\,e_j$ such that $u(t_0) = K(t_0)\,x_0 = \alpha_j\,\bar{u}_j\,e_j \geq \bar{u}$ for $\alpha_j < 0$ gives $u(t) = K(t)\,x(t) = K(t)\,\Phi_K(t,\,t_0)\,x_0 = \underbrace{K(t)\,\Phi_K(t,\,t_0)\,K_r^+(t_0)}_{=:V(t,\,t_0)}\,\underbrace{(\alpha_j\,\bar{u}_j\,e_j)}_{=:\varepsilon} \geq \bar{u}$. In particular, for all $t \in [t_0,\,t_f]$ and for all $i,\,j$,

$$u_i(t) = [K\,x]_i(t) = \sum_{k=1}^{m} v_{ik}(t,\,t_0)\,\varepsilon_k \geq \bar{u}_i$$

where

$$\epsilon_k = \begin{cases} \alpha_j\,\bar{u}_j & \text{if } k = j \\ 0 & \text{otherwise} \end{cases}$$

that is $v_{ij}(t,\,t_0)\,\alpha_j \bar{u}_j \geq \bar{u}_i$ or equivalently

$$v_{ij}(t,\,t_0) \geq \frac{\bar{u}_i}{\bar{u}_j}\,\frac{1}{\alpha_j}, \ \text{ with } \ \alpha_j < 0.$$

Hence letting $\alpha_j \to -\infty$, it follows that $v_{ij}(t,\,t_0) \geq 0$ for all $i,\,j$ and for all $t \in [t_0,\,t_f]$. Since $i$ and $j$ were arbitrarily fixed, one can conclude that (1.14) holds.

□

Now the following corollary considers the case where $K(t)$ is of full rank with $m = n$ (so $K(t)$ is invertible). Then the pseudo-inverse of $K(t)$ is the inverse of $K(t)$ :

**Corollary 1.1.4** *The cone $C_{\bar{u}}$ is invariant with respect to system (1.10) on $[t_0,\,t_f]$, with a state feedback $K(t)$ of full rank and $m = n$ if and only if the following conditions hold :*

$$\forall\, t \in [t_0,\,t_f], \quad K(t)\,\Phi_K(t,\,t_0)\,K(t_0)^{-1} \geq 0 \tag{1.16}$$

*and*

$$\forall\, t \in [t_0,\,t_f], \quad K(t)\,\Phi_K(t,\,t_0)\,K(t_0)^{-1}\,\bar{u} \geq \bar{u} \tag{1.17}$$

### 1.1.2 Invariant LTI systems

In this subsection, the particular case of state and input-invariance of linear time-invariant (LTI) homogeneous systems is studied. First, we study separately the concept of invariance of LTI system, on $\mathbb{R}_+$, with self-contained proofs. Actually, for LTI systems, the proofs of the conditions of invariance are more algebraic than in the case of LTV systems. In these proofs, a specific lemma is used, which describes the fact that the invariance implies that whenever one component of $x(t)$ reaches the boundary of the cone (i.e. $x_i(t) = 0$), it is redirected to the interior of the cone (i.e $\dot{x}_i(t) > 0$). Then, we compare the results with those obtained for LTV systems by applying the conditions of Theorems 1.1.2 and 1.1.3 for LTI systems on a fixed interval $[t_0,\,t_f]$.

**A. State-invariance**

Consider the following LTI homogeneous system, denoted by $R = [A,\, 0]$, for $t \in [t_0,\, t_f]$ :

$$\dot{x}(t) = A\, x(t), \quad x(0) = x_0. \tag{1.18}$$

Let $\bar{x} \ll 0$ be a fixed state. Consider the shifted cone $C_{\bar{x}}$ defined previously, see equation (1.5).

**Definition 1.1.4** *The cone $C_{\bar{x}}$ is said to be* **invariant with respect to system (1.18)** *if $C_{\bar{x}}$ is $e^{A\,t}$-invariant, i.e.*

$$\forall\, t \geq 0, \quad e^{A\,t}\, C_{\bar{x}} \subset C_{\bar{x}},$$

*or equivalently*

$$\forall\, t \geq 0, \forall\, x_0 \in C_{\bar{x}}, \quad x(t) := e^{A\,t}\, x_0 \in C_{\bar{x}}.$$

*In this case, system (1.18) is said to be* **state-invariant w.r.t. $\boldsymbol{C_{\bar{x}}}$**.

**Theorem 1.1.5** *The cone $C_{\bar{x}}$ is invariant with respect to system (1.18) if and only if the following conditions hold :*

$$A \text{ is a Metzler matrix} \tag{1.19}$$

*and*

$$A\, \bar{x} \geq 0. \tag{1.20}$$

**Remarks 1.1.2** *a) In view of the assumption that $\bar{x} \ll 0$, conditions (1.19)-(1.20) imply that*

$$\forall\, i = 1, \ldots, n, \quad a_{ii} \leq 0. \tag{1.21}$$

*Indeed, these conditions imply that for all $i = 1, \ldots, n$ :*

$$\begin{array}{rrcl} & (A\,\bar{x})_i & = & \displaystyle\sum_{j=1}^{n} a_{ij}\, \bar{x}_j = a_{ii}\, \bar{x}_i + \sum_{j \neq i} a_{ij}\, \bar{x}_j \geq 0 \\ \Leftrightarrow & a_{ii}\, \bar{x}_i & \geq & \displaystyle -\sum_{j \neq i} a_{ij}\, \bar{x}_j \\ \Leftrightarrow & a_{ii} & \leq & \displaystyle -\sum_{j \neq i} a_{ij}\, \frac{\bar{x}_j}{\bar{x}_i} \leq 0 \end{array}$$

*where $a_{ij} \geq 0$ for $j \neq i$ and where $\bar{x}_i$ and $\bar{x}_j$ are negative.*
*b) Condition (1.20) with condition (1.21) can be seen as a weighted diagonal dominance condition, namely*

$$\forall\, i = 1, \ldots, n, \quad |a_{ii}| \geq \sum_{j \neq i} a_{ij}\, \frac{\bar{x}_j}{\bar{x}_i} \geq 0. \tag{1.22}$$

*c) The concept of state-invariance of a LTI homogeneous is also studied in [BNS89, Chapter 4]. Holdability of closed convex sets are considered by means of subtangentiality of control linear systems (by using graphs and geometric considerations).*

To prove Theorem 1.1.5, the following lemma is needed :

**Lemma 1.1.6** *If $C_{\bar{x}}$ is invariant with respect to system (1.18), if for all $t \geq 0$, $x(t) = e^{At}x_0$, where $x_0$ is any initial state in $C_{\bar{x}}$ and if there exists $i \in \{1, \ldots, n\}$ such that $x_i(t) = \bar{x}_i$, then $\dot{x}_i(t) \geq 0$.*

**Proof :**
For the sake of contradiction, assume that for some $x_0 \geq \bar{x}$, there exists at least one coordinate $i = 1, \ldots, n$ and a time $t \geq 0$ such that $x_i(t) = \bar{x}_i$ and $\dot{x}_i(t) < 0$. Now, by assumption, for all $x_0 \geq \bar{x}$, $x(t) \geq \bar{x}$ where $x(t)$ is solution of system (1.18). In particular, $x_i(t) \geq -\bar{x}_i$, $\forall\, t \geq 0$. Moreover, $\dot{x}(t) = A\, x(t) = A\, e^{A\,t}\, x_0 = e^{A\,t}\, A\, x_0$, $t \geq 0$. Then since the function $\dot{x}_i(\cdot)$ is continuous on $\mathbb{R}_+$, $\dot{x}_i(t) < 0$ implies that there exists $t_1 > t$ such that for all $\tau \in [t, t_1]$, $\dot{x}_i(\tau) < 0$, that means $x_i(\cdot)$ is strictly decreasing on $[t, t_1]$ with $x_i(t) = \bar{x}_i$. Therefore $x_i(t_1) < x_i(t) = \bar{x}_i$, for all $\tau \in [t, t_1]$. It follows that $x(t_1) \notin C_{\bar{x}}$. On the other hand, since $x(t) \in C_{\bar{x}}$, for all $t \geq 0$, $x(t_1) = e^{A\,(t_1-t)}x(t) \in C_{\bar{x}}$. This clearly contradicts the fact that $x(t_1) \notin C_{\bar{x}}$. Thus $\dot{x}_i(t) \geq 0$.

□

**Proof of Theorem 1.1.5 :**
Necessity : Since $\forall\, x_0 \geq \bar{x},\ x(t) \geq \bar{x}$, for all $t \geq 0$, by Lemma 1.1.6, $\forall\, i = 1, \ldots, n$, such that $x_i(t) = \bar{x}_i$, $\dot{x}_i(t) = \displaystyle\sum_{j=1}^{n} a_{ij}x_j(t) \geq 0$. First, observe that $x(0) = x_0 := \bar{x} \in C_{\bar{x}}$. It follows from Lemma 1.1.6 applied in $t = 0$ that $\forall\, i = 1, \ldots, n,\ \dot{x}_i(0) \geq 0$, or equivalently $\dot{x}(0) = A\,\bar{x} \geq 0$, i.e. condition (1.20) holds. Now remark that for any $y \in \mathbb{R}^n$, $y \in C_{\bar{x}}$ if and only if $y$ is on the form

$$y = \sum_{j=1}^{n} \alpha_j\, \bar{x}_j\, e_j := \Delta\, \bar{x} \tag{1.23}$$

for some (unique) diagonal matrix $\Delta = \mathrm{diag}[\alpha_i]_{i=1}^{n}$ where $\forall\, i = 1, \ldots, n,\ \alpha_i \leq 1$.
Then consider $x(0) = y$ of the form (1.23) where $\alpha_i = 1$ and $\alpha_j < 0$ for some arbitrarily fixed $i, j = 1, \ldots, n$ such that $i \neq j$ and $\forall\, k \neq i$ and $k \neq j,\ \alpha_k = 0$. Therefore,

$$\dot{x}_i(0) = a_{ii}\, \bar{x}_i + a_{ij}\alpha_j\, \bar{x}_j \geq 0$$

or equivalently

$$a_{ij} \geq -\frac{\bar{x}_i}{\bar{x}_j}\, a_{ii}\, \frac{1}{\alpha_j}, \ \text{ with } \ \alpha_j < 0,\ i \neq j.$$

Hence letting $\alpha_j \to -\infty$, it follows that $a_{ij} \geq 0$. Since $i$ and $j$ were arbitrarily fixed, one can conclude that (1.19) holds.

Sufficiency : Assume that (1.19) and (1.20) hold, i.e.

$$\forall\, t \geq 0,\ e^{A\,t} \geq 0 \ \text{ and } \ A\,\bar{x} \geq 0.$$

Then, for all $t \geq 0$,

$$e^{A\,t}\,\bar{x} = \bar{x} + \int_0^t e^{A\,\tau} A\,\bar{x}\, \mathrm{d}\tau,$$

where $\forall\, \tau \in [0,\, t],\ e^{A\,\tau}\, A\, \bar{x} \geq 0$. Therefore $e^{A\,t}\, \bar{x} \geq \bar{x}$. Hence, for all $t \geq 0$ and for every $x_0 \in \mathbb{R}^n$ such that $x_0 \geq \bar{x}$,

$$e^{A\,t} x_0 \geq e^{A\,t} \bar{x} \geq \bar{x},$$

that is $x(t) \geq \bar{x}$.

□

Now, in order to link the concept of state-invariance on a fixed interval $[t_0,\, t_f]$ with the state-invariance (on $\mathbb{R}_+$), the following proposition is useful :

**Proposition 1.1.7** *For a LTI system $R = [A,\, 0]$, the conditions of Theorem 1.1.2 on any interval $[t_0,\, t_f]$, i.e.*

$$\forall\, t \in [t_0,\, t_f], \quad A(t) \text{ is a Metzler matrix} \tag{1.6}$$

*and*

$$\forall\, t \in [t_0,\, t_f], \quad \Phi(t,\, t_0)\, \bar{x} \geq \bar{x}, \tag{1.7}$$

*are equivalent to*

$$A \text{ is a Metzler matrix} \tag{1.24}$$

*and*

$$A\, \bar{x} \geq 0. \tag{1.25}$$

**Proof :** First condition (1.6) is clearly equivalent to (1.24) when $A$ is a constant matrix. Now assume that condition (1.7) holds. Then since $\forall t \in [t_0,\, t_f]$, $\Phi(t,\, t_0) = e^{A\,(t-t_0)}$, condition (1.7) can be rewritten as $(e^{A\,(t-t_0)} - I_n)\, \bar{x} \geq 0$. Now observe that

$$A = \lim_{t \to t_{0+}} \frac{e^{A\,(t-t_0)} - I_n}{t - t_0}.$$

Hence,

$$A\, \bar{x} = \lim_{t \to t_{0+}} \frac{(e^{A\,(t-t_0)} - I_n)\, \bar{x}}{t - t_0} \geq 0$$

i.e. condition (1.25) holds. Conversely, assume that $A$ is a Metzler matrix (i.e. $e^{A\,t} \geq 0$ for all $t \geq 0$, see Proposition A.1.3) and $A\, \bar{x} \geq 0$. Since for $t \in [t_0,\, t_f]$,

$$(e^{A\,(t-t_0)} - I_n)\, \bar{x} = \int_{t_0}^{t} e^{A\,(\tau - t_0)}\, A\, \bar{x}\, \mathrm{d}\tau \geq 0,$$

it follows that condition (1.7) holds on $[t_0,\, t_f]$, where $\Phi(t,\, t_0) = e^{A\,(t-t_0)}$.

□

These considerations lead to the following corollary :

**Corollary 1.1.8** *The cone $C_{\bar{x}}$ is invariant with respect to system (1.18) if and only if for any interval $[t_0,\, t_f] \subseteq \mathbb{R}_+$, the cone $C_{\bar{x}}$ is invariant with respect to system (1.1) on $[t_0,\, t_f]$.*

**Proof :** The result follows directly from Proposition 1.1.7 and Theorems 1.1.2 and 1.1.5.

□

Therefore, for LTI systems, we obtain the equivalence of the concepts of state-invariance on any interval $[t_0,\, t_f] \subseteq \mathbb{R}_+$ and on $\mathbb{R}_+$.

**B. Input-invariance**

Consider a LTI system $R = [A,\, B]$ with an input $u(t) = K\, x(t)$ for $t \in [t_0,\, t_f]$, where $K$ is a state feedback. Therefore we consider the following LTI closed-loop system, denoted by $R = [A + B\, K,\, 0]$, for $t \geq 0$,

$$\dot{x}(t) = (A + B\, K)\, x(t), \quad x(0) = x_0. \tag{1.26}$$

Let $\bar{u} \leq 0$ be a fixed input. Consider the shifted cone $C_{\bar{u}}$ defined previously, see equation (1.11).

**Definition 1.1.5** *The cone $C_{\bar{u}}$ is said to be* **invariant with respect to system (1.26)** *if $C_{\bar{u}}$ is $K\, e^{(A+B\,K)\,t}$-invariant, i.e.*

$$\forall\, t \geq 0, \quad K\, e^{(A+B\,K)\,t}\, C_{\bar{u}} \subset C_{\bar{u}},$$

*or equivalently*

$$\forall\, t \geq 0, \forall\, x_0 \ \textit{such that} \ u(0) = K\, x_0 \in C_{\bar{u}}, \quad u(t) := K\, x(t) = K\, e^{(A+B\,K)\,t}\, x_0 \in C_{\bar{u}}.$$

*In this case, system (1.26) is said to be* **input-invariant w.r.t. $\boldsymbol{C_{\bar{u}}}$**.

**Theorem 1.1.9** *The cone $C_{\bar{u}}$ is invariant with respect to system (1.26), with a state feedback $K$ of full rank and $m = n$ if and only if the following conditions hold :*

$$K\,(A + B\,K)K^{-1} \ \textit{is a Metzler matrix} \tag{1.27}$$

*and*

$$K\,(A + B\,K)K^{-1}\,\bar{u} \geq 0 \tag{1.28}$$

As for Theorem 1.1.5, an additional lemma is needed to prove this theorem.

**Lemma 1.1.10** *If $C_{\bar{u}}$ is invariant with respect to system (1.26), if for all $t \geq 0$, $x(t) = e^{A\,t}x_0$, where $x_0$ is any initial state such that $K\, x_0 \in C_{\bar{u}}$ and if there exists $i \in \{1, \dots, m\}$ such that $u_i(t) := [K\, x]_i(t) = \bar{u}_i$, then $\dot{u}_i(t) = [\dot{K\, x}]_i(t) \geq 0$.*

**Remark 1.1.3** *This result holds for any $x_0 \in \mathbb{R}^n$. Here it is not needed to assume that $x_0 \geq \bar{x}$ and that for all $t \geq 0$, $x(t) \in C_{\bar{x}}$.*

**Proof :** For the sake of contradiction, assume that for some $x_0$ such that $u(0) = K\, x_0 \geq \bar{u}$, there exists at least one $i = 1, \dots, m$ and a time $t \geq 0$ such that $u_i(t) = [K\, x]_i(t) = \bar{u}_i$ and $[\dot{K\, x}]_i(t) < 0$. Then by continuity, there exists $t_1 > t$ such that for all $\tau \in [t,\, t_1]$, $[\dot{K\, x}]_i(\tau) < 0$, that means that the function $u_i(\cdot) = [K\, x]_i(\cdot)$ is strictly decreasing on $[t,\, t_1]$ with $u_i(t) = [K\, x]_i(t) = \bar{u}_i$. Therefore $[K\, x]_i(t_1) < [K\, x]_i(t) = \bar{u}_i$, for all $\tau \in [t,\, t_1]$. It follows that $u(t_1) \notin C_{\bar{u}}$. On the other hand, since $u(t) = K\, x(t) \in C_{\bar{u}}$, for all $t \geq 0$, $u(t_1) = K\, e^{(A+B\,K)\,(t_1-t)}x(t) \in C_{\bar{u}}$. This clearly contradicts the fact that $u(t_1) \notin C_{\bar{u}}$. Thus $[\dot{K\, x}]_i(t) \geq 0$. $\square$

**Proof of Theorem 1.1.9 :**
Necessity : The proof is similar to the one of Theorem 1.1.5. Since for all $x_0$ such that $u(0) = K\,x_0 \geq \bar{u}$, $u(t) = K\,x(t) \geq \bar{u}$, for all $t \geq 0$, by Lemma 1.1.10, $\forall\, i = 1, \ldots, m$, such that $[K\,x]_i(t) = \bar{u}_i$, $\dot{[K\,x]}_i(t) \geq 0$. First, let $x_0 = K^{-1}\,\bar{u}$ such that $u(0) = K\,x_0 = K\,K^{-1}\,\bar{u} = \bar{u}$. It follows by Lemma 1.1.10 applied at time $t = 0$ that $\forall\, i = 1, \ldots, m,\ \dot{u}_i(0) = \dot{[K\,x]}_i(0) \geq 0$, or equivalently,

$$K\,\dot{x}(0) = K\,(A + B\,K)\,x(0) = K\,(A + B\,K)\,K^{-1}\bar{u} \geq 0,$$

i.e. condition (1.28) holds. Now observe that for any $y \in \mathbb{R}^m$, $y \in C_{\bar{u}}$ if and only if $y$ is of the form

$$y = \sum_{j=1}^{m} \alpha_j\,\bar{u}_j\,e_j := \Delta'\,\bar{u} \tag{1.29}$$

for some (unique) diagonal matrix $\Delta' = \mathrm{diag}[\alpha_i]_{i=1}^m$ where $\forall\, i = 1, \ldots, m,\ \alpha_i \leq 1$.
Then consider $x_0 = K^{-1}\,y$ such that $u(0) = K\,x_0 = y$ of the form (1.29) where $\alpha_i = 1$ and $\alpha_j < 0$ for some arbitrarily fixed $i,\ j = 1, \ldots, m$ such that $i \neq j$ and $\forall k \neq i$ and $k \neq j,\ \alpha_k = 0$. Therefore, by Lemma 1.1.10,

$$\dot{u}_i(0) = \dot{[K\,x]}_i(0) \geq 0 \quad \text{i.e.} \quad [K\,(A{+}B\,K)\,x(0)]_i = [\underbrace{K\,(A + B\,K)\,K^{-1}}_{:=V}\,\underbrace{(\bar{u}_i\,e_i + \alpha_j\,\bar{u}_j\,e_j)}_{:=\varepsilon}]_i \geq 0$$

or equivalently

$$[V\,\varepsilon]_i = \sum_{k=1}^{m} v_{ik}\,\varepsilon_k \geq 0$$

where

$$\epsilon_k = \begin{cases} \bar{u}_i & \text{if } k = i \\ \alpha_j\,\bar{u}_j & \text{if } k = j \neq i \\ 0 & \text{otherwise} \end{cases}$$

that is $v_{ii}\,\bar{u}_i + v_{ij}\,\alpha_j\bar{u}_j \geq 0$ or equivalently

$$v_{ij} \geq \frac{\bar{u}_i}{\bar{u}_j}\,v_{ii}\,\frac{1}{\alpha_j},\ \text{ with } \alpha_j < 0,\ i \neq j.$$

Hence letting $\alpha_j \to -\infty$, it follows that $v_{ij} \geq 0$. Since $i$ and $j$ were arbitrarily fixed, one can conclude that (1.27) holds.

Sufficiency : Assume that (1.27) and (1.28) hold, i.e., with $V := K\,(A + B\,K)\,K^{-1}$,

$$V\,\bar{u} \geq 0 \text{ and } \forall\, t \geq 0,\ e^{V\,t} \geq 0.$$

Recall that, for all $t \geq 0$,

$$e^{V\,t}\,\bar{u} = \bar{u} + \int_0^t e^{V\,\tau}\,V\,\bar{u}\,\mathrm{d}\tau,$$

where $\forall\, \tau \in [0,\,t],\ e^{V\,\tau}\,V\,\bar{u} \geq 0$. Therefore $e^{V\,t}\,\bar{u} \geq \bar{u}$. Hence, for all $t \geq 0$ and for every $x_0 \in \mathbb{R}^n$ such that $K\,x_0 \geq \bar{u}$, one has :

$$e^{V\,t}\,K\,x_0 \geq e^{V\,t}\,\bar{u} \geq \bar{u},$$

that is, with $e^{Vt} = e^{K\,(A+B\,K)\,K^{-1}\,t} = K\, e^{(A+B\,K)\,t} K^{-1}$,

$$K\, \underbrace{e^{(A+B\,K)\,t} K^{-1} K\, x_0}_{=x(t)} \geq \bar{u}$$

Then the cone $C_{\bar{u}}$ is invariant with respect to system (1.26).

□

Now, in order to link the concept of input-invariance on a fixed interval $[t_0,\, t_f]$ with the input-invariance (on $\mathbb{R}_+$), the following proposition is needed :

**Proposition 1.1.11** *For a LTI system $R = [A + B\,K,\, 0]$ with a state feedback $K(t)$ of full rank and $m = n$, the conditions of Corollary 1.1.4, i.e.*

$$\forall\, t \in [t_0,\, t_f], \quad K(t)\, \Phi_K(t,\, t_0)\, K(t_0)^{-1} \geq 0 \tag{1.16}$$

*and*

$$\forall\, t \in [t_0,\, t_f], \quad K(t)\, \Phi_K(t,\, t_0)\, K(t_0)^{-1}\, \bar{u} \geq \bar{u}, \tag{1.17}$$

*are equivalent to*

$$K\,(A + B\,K)\, K^{-1} \text{ is a Metzler matrix} \tag{1.30}$$

*and*

$$K\,(A + B\,K)\, K^{-1}\, \bar{u} \geq 0. \tag{1.31}$$

**Proof :**
Necessity : Assume that condition (1.16) holds, that is, with $\Phi_K(t,\, t_0) := e^{(A+B\,K)(t-t_0)}$ for $t \in [t_0,\, t_f]$,

$$\begin{array}{rll} & K\, e^{(A+B\,K)(t-t_0)} K^{-1} & \geq & 0 \\ \Leftrightarrow & e^{(K\,(A+B\,K)\,K^{-1})(t-t_0)} & \geq & 0. \end{array}$$

Then $K\,(A + B\,K)\, K^{-1}$ is a Metzler matrix (by Proposition A.1.3) and condition (1.30) holds. Moreover, if condition (1.17) holds, then for $t \in [t_0,\, t_f]$,

$$\begin{array}{rlll} & K\, e^{(A+B\,K)(t-t_0)} K^{-1} \bar{u} - \bar{u} & \geq & 0 \\ \Leftrightarrow & (K\, e^{(A+B\,K)(t-t_0)} K^{-1} - I_m)\, \bar{u} & \geq & 0 \\ \Leftrightarrow & (e^{(K\,(A+B\,K)\,K^{-1})(t-t_0)} - I_m)\, \bar{u} & \geq & 0. \end{array}$$

Hence, with $V := K\,(A + B\,K)\, K^{-1}$,

$$V = \lim_{t \to t_{0+}} \frac{e^{V(t-t_0)} - I_m}{t - t_0},$$

whence

$$V\, \bar{u} = \lim_{t \to t_{0+}} \frac{(e^{V(t-t_0)} - I_m)\, \bar{u}}{t - t_0} \geq 0,$$

i.e. condition (1.31) holds.

<u>Sufficiency</u> : Assume that conditions (1.30)-(1.31) hold. Condition (1.30) also reads

$$e^{(K\,(A+B\,K)\,K^{-1})(t-t_0)} \geq 0 \qquad \text{for all } t \in [t_0,\, t_f],$$

i.e.

$$K\, e^{(A+B\,K))(t-t_0)}\, K^{-1} \geq 0.$$

Then condition (1.16) holds. Now with $V := K\,(A + B\,K)\,K^{-1}$, conditions (1.30)-(1.31) are equivalent to $e^{V(t-t_0)} \geq 0$ for all $t \in [t_0,\, t_f]$ and $V\,\bar{u} \geq 0$. It follows that

$$(e^{V(t-t_0)} - I_m)\,\bar{u} = \int_{t_0}^{t} e^{V\,(\tau-t_0)}\, V\,\bar{u}\,\mathrm{d}\tau \geq 0,$$

i.e. $e^{V(t-t_0)}\bar{u} \geq \bar{u}$, or equivalently condition (1.17) holds.

□

These considerations lead to the following corollary :

**Corollary 1.1.12** *The cone $C_{\bar{u}}$ is invariant with respect to system (1.26), with a state feedback $K(t)$ of full rank and $m = n$, if and only if for any interval $[t_0,\, t_f] \subseteq \mathbb{R}_+$, the cone $C_{\bar{u}}$ is invariant with respect to system (1.10) on $[t_0,\, t_f]$.*

**Proof :** The result follows directly from Proposition 1.1.11, Corollary 1.1.4 and Theorem 1.1.9.

□

Therefore, for LTI systems, we obtain the equivalence of the concepts of input-invariance on any interval $[t_0,\, t_f] \subseteq \mathbb{R}_+$ and on $\mathbb{R}_+$.

### C. Input/state-invariance

In this part, we consider the problem of invariance of the state and the input. For this purpose, we define a new cone $C_{\bar{x},\bar{u}}$ which joins the two previous cases as follows :

$$C_{\bar{x},\bar{u}} := \left\{ x \in \mathbb{R}^n \;:\; \begin{bmatrix} I_n \\ K \end{bmatrix} x \geq \begin{bmatrix} \bar{x} \\ \bar{u} \end{bmatrix} \right\} = C_{\bar{x}} \cap C_{\bar{u}}. \tag{1.32}$$

**Definition 1.1.6** *The cone $C_{\bar{x},\bar{u}}$ is said to be* **invariant with respect to system (1.26)** *if $C_{\bar{x},\bar{u}}$ is $e^{(A+B\,K)\,t}$-invariant, i.e.*

$$\forall\, t \geq 0, \quad e^{(A+B\,K)\,t}\, C_{\bar{x},\bar{u}} \subset C_{\bar{x},\bar{u}},$$

*or equivalently*

$$\forall\, t \geq 0, \forall\, x_0 \in C_{\bar{x},\bar{u}}, \quad x(t) = e^{(A+B\,K)\,t}\, x_0 \in C_{\bar{x},\bar{u}}.$$

*In this case, system (1.26) is said to be* **input/state-invariant w.r.t. $\boldsymbol{C_{\bar{x},\bar{u}}}$.**

So the state feedback $K$ is such that for all $x_0 \geq \bar{x}$, the cone $C_{\bar{x},\bar{u}}$ is invariant with respect to system (1.26), that is such that for all $t \geq 0$, $x(t) \geq \bar{x}$ and $u(t) = K\,x(t) \geq \bar{u}$. The following result gives a characterization of such a $K$ and is an adapted version of [CH93, Proposition 1, p. 1681], which is summarized in Lemma 1.1.14 below.

**Theorem 1.1.13** *The cone $C_{\bar{x},\bar{u}}$ is invariant with respect to system (1.26) if and only if there exists a Metzler matrix $\mathcal{H} \in I\!R^{(m+n)\times(m+n)}$ such that*

$$\left\{ \begin{array}{l} \begin{bmatrix} -I_n \\ -K \end{bmatrix} (A + B\,K) - \mathcal{H} \begin{bmatrix} -I_n \\ -K \end{bmatrix} = \begin{bmatrix} 0_{n\times n} \\ 0_{m\times n} \end{bmatrix} \\ \mathcal{H} \begin{bmatrix} -\bar{x} \\ -\bar{u} \end{bmatrix} \leq 0_{(m+n)\times 1} \end{array} \right. \tag{1.33}$$

**Lemma 1.1.14** *The set $R[Q,\rho] := \{x \in I\!R^n \ : \ Q\,x \leq \rho\}$, with $Q \in I\!R^{r\times n}$ and $\rho \in I\!R^r$ is invariant with respect to system (1.18) if and only if there exists a Metzler matrix $\mathcal{H} \in I\!R^{r\times r}$ such that*

$$Q\,A - \mathcal{H}\,Q = 0$$
$$\mathcal{H}\,\rho \leq 0$$

**Proof of Theorem 1.1.13 :** The result follows from Lemma 1.1.14 applied to $A := A + B\,K$ with the following identifications : $Q := \begin{bmatrix} -I_n \\ -K \end{bmatrix}$, $\rho := \begin{bmatrix} -\bar{x} \\ -\bar{u} \end{bmatrix}$ and $r := m + n$. $\square$

**Remark 1.1.4** *The invariance conditions (1.33) do not require any particular assumption on the matrix $Q := \begin{bmatrix} -I_n \\ -K \end{bmatrix}$ and on the vector $\rho := \begin{bmatrix} -\bar{x} \\ -\bar{u} \end{bmatrix}$. However, if $\rho \gg 0$, condition $\mathcal{H}\,\rho \leq 0$ with $\mathcal{H}$ a Metzler matrix implies that $-\mathcal{H}$ is a M-matrix, see Definition A.2.1. Then the real-parts of the eigenvalues of $\mathcal{H}$ are nonpositive, see Theorem A.2.2.*

Now by the previous analysis on the invariance of the cone $C_{\bar{x}}$ and $C_{\bar{u}}$, we obtain the following sufficient condition, by choosing an appropriate matrix $\mathcal{H}$ in Theorem 1.1.13.

**Corollary 1.1.15** *If there exists a state feedback $K$ such that the following conditions hold :*

$$\left\{ \begin{array}{l} A + B\,K \text{ is a Metzler matrix} \\ (A + B\,K)\,\bar{x} \geq 0 \\ K\,A \geq 0 \\ K\,B \text{ is a Metzler matrix} \\ K\,(A\,\bar{x} + B\,\bar{u}) \geq 0, \end{array} \right. \tag{1.34}$$

*then the cone $C_{\bar{x},\bar{u}}$ is invariant with respect to system (1.26).*

**Proof :** By Theorem 1.1.13, a necessary and sufficient condition of invariance of the cone $C_{\bar{x},\,\bar{u}}$ is the existence of a matrix $\mathcal{H} = \begin{bmatrix} \mathcal{H}_1 & \mathcal{H}_2 \\ \mathcal{H}_3 & \mathcal{H}_4 \end{bmatrix}$ with $\mathcal{H}_1 \in \mathbb{R}^{n\times n}$, $\mathcal{H}_2 \in \mathbb{R}^{n\times m}$, $\mathcal{H}_3 \in \mathbb{R}^{m\times n}$ and $\mathcal{H}_4 \in \mathbb{R}^{m\times m}$ such that

$$\left\{ \begin{array}{l} -(A + B\,K) + \mathcal{H}_1 + \mathcal{H}_2\,K = 0_{n\times n} \\ -K\,(A + B\,K) + \mathcal{H}_3 + \mathcal{H}_4\,K = 0_{m\times n} \\ -\mathcal{H}_1\,\bar{x} - \mathcal{H}_2\,\bar{u} \leq 0_{n\times 1} \\ -\mathcal{H}_3\,\bar{x} - \mathcal{H}_4\,\bar{u} \leq 0_{m\times 1} \\ \mathcal{H}_1,\, \mathcal{H}_4 \text{ are Metzler matrices} \\ \mathcal{H}_2,\, \mathcal{H}_3 \geq 0. \end{array} \right. \tag{1.35}$$

Choosing $\mathcal{H}_1 = A + B\,K$, $\mathcal{H}_2 = 0_{n\times m}$, $\mathcal{H}_3 = K\,A$ and $\mathcal{H}_4 = K\,B$ in equations (1.35) leads easily to the sufficient conditions (1.34). $\square$

## 1.2 Stability of invariant LTI homogeneous systems

In this section the stability of invariant LTI homogeneous system (1.18) is studied.

### 1.2.1 Reminders on linear systems

First recall the definition of stability for LTI homogeneous systems and some useful results, see e.g. [CD91].

#### A. Definition and characterization of stability

**Definition 1.2.1**

- *A LTI homogeneous system (1.18) is said to be* **asymptotically stable** *if for all* $x_0 \in \mathbb{R}^n$, $x(t)$ *tends to zero as* $t$ *tends to infinity.*
- *A LTI homogeneous system (1.18) is said to be* **exponentially stable** *if,* $\exists\, \alpha,\ \beta > 0$ *such that for all* $t \geq 0$, $\|e^{A\,t}\| \leq \beta\, e^{-\alpha\, t}$.

**Theorem 1.2.1 (Asymptotic stability)** *A LTI homogeneous system (1.18) is asymptotically stable if and only if* $e^{A\,t} \;\rightarrow\; 0$ *as* $t \;\rightarrow\; \infty$.

**Theorem 1.2.2 (Exponential stability)** *A LTI homogeneous system (1.18) is exponentially stable if and only if every eigenvalue of* $A$ *has negative real part, i.e.*

$$\forall\, \lambda \in \sigma(A) \;:\; \mathcal{R}e(\lambda) < 0. \tag{1.36}$$

**Remark 1.2.1** *These two concepts of stability are equivalent, see e.g. [CD91]. Therefore in the sequel, the terms "asymptotic" and "exponential" are omitted. Moreover, the abuse of language "$A$ is stable" is also used instead of "system (1.18) is stable" whenever the characterization (1.36) is used to prove the stability of a system.*

#### B. Stability and Lyapunov equation

Consider the Lyapunov equation

$$A^T P + P\,A = -Q \tag{1.37}$$

where $A \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{n\times n}$ is a symmetric positive definite and a unique symmetric positive definite solution $P$ is to be found for (1.37). The solvability of the Lyapunov equation relates directly to the stability of system (1.18), see [CD91, pp. 186-188].

**Theorem 1.2.3** *A LTI homogeneous system (1.18) is stable if and only if for all symmetric positive definite matrix $Q$, the Lyapunov equation (1.37) has a unique symmetric positive definite solution $P$ given by*

$$P = \int_0^{\infty} e^{A^T t} Q\, e^{A\,t}\, dt.$$

### 1.2.2 Stability of invariant LTI systems

For the particular case of state-invariant LTI homogeneous systems, we obtain the following result which is a direct consequence of Theorem 1.2.2 together with Theorem 1.1.5 and the spectral property of a Metzler matrix, see Theorem A.1.5 :

**Theorem 1.2.4** *A state-invariant LTI homogeneous system (1.18) is stable if the Frobenius eigenvalue of $A$ (i.e. the dominant eigenvalue of $A$, see Theorem A.1.5) is negative.*

## 1.3 Invariant stabilizability of LTI systems

This section is devoted to the concept of invariant stabilizability which is the idea of keeping the invariance of the system as well as stabilizing it.

### 1.3.1 Invariant stability of LTI systems

**Definition 1.3.1** *A LTI homogeneous system (1.18) is said to be* **invariant stable** *if for all $t \geq 0$ and for all $x_0 \geq \bar{x}$, $x(t) \geq \bar{x}$ and $x(t) \to 0$ as $t \to \infty$.*

By using Theorem 1.1.5, we obtain the following characterization for invariant stability :

**Theorem 1.3.1** *A LTI homogeneous system (1.18) is invariant stable if and only if $A$ is a stable Metzler matrix such that $A\,\bar{x} \geq 0$.*

A characterization of invariant stability can be expressed by using the Lyapunov equation (1.37) :

**Theorem 1.3.2** *A LTI homogeneous system (1.18) is invariant stable if and only if $A$ is a Metzler matrix such that $A\,\bar{x} \geq 0$ and if for all symmetric positive definite matrix $Q$, the Lyapunov equation (1.37) has a unique symmetric positive definite solution $P$.*

### 1.3.2 Invariant stabilizability of LTI systems

**Definition 1.3.2** *A LTI system $[A,\, B]$ is said to be* **invariant stabilizable** *if for all $x_0 \geq \bar{x}$, there exists an input $u(t)$ such that for all $t \geq 0$, $x(t) \geq \bar{x}$ and $x(t) \to 0$ as $t \to \infty$.*

The following criteria of invariant stabilizability are inspired by the particular case of positive systems, see e.g. [BNS89, Chapter 7] and [Ava00, pp. 73-75] :

**Theorem 1.3.3** *If there exists a state feedback law $u(t) = K\,x(t)$ such that $A + B\,K$ is a stable Metzler matrix and $(A + B\,K)\,\bar{x} \geq 0$, then the resulting system $[A + B\,K,\, 0]$ is invariant stabilizable.*

A matrix $K$ which verifies the conditions of Theorem 1.3.3 is said to be **invariant stabilizing**. Now a criterion of invariant stabilizability can also be expressed by using the Lyapunov equation (1.37), see e.g. [BEFB94, Section 10.3, p. 144] :

**Theorem 1.3.4** *A LTI system $[A,\ B]$ is invariant stabilizable if there exists a state feedback law $u(t) = K\,x(t)$ such that $A + B\,K$ is a Metzler matrix such that $(A + B\,K)\,\bar{x} \geq 0$ and if for all symmetric positive definite matrix $Q$, the Lyapunov equation*

$$P(A + B\,K)^T + (A + B\,K)\,P = -Q \tag{1.38}$$

*has a unique symmetric positive definite solution $P$.*

**Proof :** The result follows from Theorem 1.3.3 which states that system (1.18) is invariant stabilizable if there exists a state feedback law $u(t) = K\,x(t)$ such that $A + B\,K$ is a stable Metzler matrix such that $(A+B\,K)\,\bar{x} \geq 0$. Then from Theorem 1.2.3 applied to matrix $A+B\,K$ instead of $A$, the property of stability can be translated in terms of the Lyapunov equation (1.38). $\square$

# Chapter 2

# Positive Linear Systems

An important question in system and control theory is the invariance of the nonnegative orthant of the state space for linear systems. When they satisfy that property, such systems are called *positive* linear systems. They encompass controlled dynamical models where all the variables, i.e. the state and output variables, should remain nonnegative for any nonnegative initial conditions and input functions. In comparison with invariant systems which have been studied in the previous chapter, positive systems can be considered as state-invariant systems with respect to the nonnegative orthant $C_{\bar{x}} = \mathbb{R}^n_+$ (i.e. $\bar{x} := 0$ in the previous analysis).

An overview of the state of the art in positive systems theory is given e.g. in [FR00], [Ka 02], [Lue79], [Van07] and [HCH10]. Typical examples of positive systems are economics models, chemical processes or age-structured populations (see e.g. [FR00, God83, Van08, HCH10]).

Numerous system theoretic problems have already been (and are still) investigated for positive systems : for example the realization, controllability and reachability problems (see e.g. [BF04, Van97, BCR$^+$02] and the references therein), the positive stabilization problem (see e.g. [BNS89]), the linear quadratic (LQ) problem (see e.g. [AM90, CD91]) for (general) linear systems with positive controls (see e.g. [HVS98] and the references therein for the general LQ problem with positive controls and [Ka 02] for the minimal energy positive control problem for positive systems).

In this chapter, the main results concerning positive linear systems are described, such as the spectral property, the stability and the stabilizability of such systems. Since the theory of positive linear systems has been widely studied in the literature, the proofs of the results are not provided, but only references where they can be found (see also [Bea06]).

# 2.1 Main concepts and results

## 2.1.1 Positive LTV systems

**Definition 2.1.1**

- *A LTV homogeneous system* $R = [A(\cdot),\, 0]$ *is said to be* **positive** *on* $[t_0,\, t_f]$ *if*

$$\forall\, x_0 \geq 0 \;:\; \forall\, t \in [t_0,\, t_f],\; x(t) = \Phi(t,\, t_0)\, x_0 \geq 0,$$

*where* $\Phi(t,\, t_0)$ *is the fundamental matrix and satisfies the following homogeneous equation :*

$$\begin{cases} \frac{\partial \Phi}{\partial t}(t,\, t_0) = A(t)\, \Phi(t,\, t_0), & \forall\, t \in [t_0,\, t_f], \\ \Phi(t_0,\, t_0) = I_n. \end{cases}$$

- *A LTV system* $R = [A(\cdot),\, B(\cdot)]$ *is said to be* **positive** *on* $[t_0,\, t_f]$ *if*

$$\forall\, x_0 \geq 0,\, \forall\, u(\cdot) \geq 0 \;:\; \forall\, t \in [t_0,\, t_f],\; x(t) = \Phi(t,\, t_0)\, x_0 + \int_{t_0}^{t} \Phi(t, \tau) B(\tau) u(\tau)\, d\tau \geq 0.$$

The following theorem gives a well-known characterization of the positivity of linear time-varying systems in continuous time, see [AS03, Section VIII] (see also [Ka 01, Theorem 2] whose condition turns out to be equivalent to (2.1)).

**Theorem 2.1.1**

- *A LTV homogeneous system* $R = [A(\cdot),\, 0]$ *is positive on* $[t_0,\, t_f]$ *if and only if for all* $t \in [t_0,\, t_f]$,

$$A(t) \text{ is a Metzler matrix.} \tag{2.1}$$

- *A LTV system* $R = [A(\cdot),\, B(\cdot)]$ *is positive on* $[t_0,\, t_f]$ *if and only if for all* $t \in [t_0,\, t_f]$, $A(t)$ *is a Metzler matrix and* $B(t) \geq 0$.

**Remark 2.1.1** *Condition (2.1) corresponds to conditions (1.6)-(1.7) of Theorem 1.1.2 applied with* $\bar{x} = 0$.

## 2.1.2 Positive LTI systems

For LTI systems, we obtain the following well-known characterization of the positivity :

**Theorem 2.1.2**

- *A LTI homogeneous system* $R = [A,\, 0]$ *is positive (on* $\mathbb{R}_+$*) if and only if*

$$A \text{ is a Metzler matrix.} \tag{2.2}$$

- *A LTI system* $R = [A,\, B]$ *is positive (on* $\mathbb{R}_+$*) if and only if* $A$ *is a Metzler matrix and* $B \geq 0$.

**Remark 2.1.2** *Condition (2.2) corresponds to conditions (1.19)-(1.20) of Theorem 1.1.5 applied with* $\bar{x} = 0$.

## 2.2 Stability of positive LTI systems

Thanks to Theorem 2.1.2, we can characterize the stability of a positive LTI system by using the properties of Metzler matrices developed in Section A.1. By using Theorem 1.2.2 on stability of linear systems together with Theorem A.1.5 (Perron-Frobenius Theorem for Metzler matrix), we obtain the following result on the stability of positive systems :

**Theorem 2.2.1** *A positive LTI system $R = [A, 0]$ is (exponentially) stable if and only if the Frobenius eigenvalue $\lambda_F$ of $A$ is negative.*

It is also interesting to study the Lyapunov equation in the case of positive systems. We obtain the following adaptation of Theorem 1.2.3 for positive systems, see [FR00, pp. 41-42] :

**Theorem 2.2.2** *A positive LTI system $R = [A, 0]$ is stable if and only if there exists a diagonal positive definite matrix $P$ such that the matrix $Q$, defined by*

$$-Q = A^T P + P\,A$$

*is positive definite.*

**Remark 2.2.1** *Observe that, by Theorem 1.1.5, a system $[A, 0]$ which is state-invariant w.r.t. $C_{\bar{x}}$ is a positive system. Then Theorem 2.2.2 also holds for state-invariant systems.*

## 2.3 Positive stabilizability of LTI systems

### 2.3.1 Positive stability of LTI systems

**Definition 2.3.1** *A LTI system $R = [A, 0]$ is said to be* **positively stable** *if for all $t \geq 0$ and for all $x_0 \geq 0$, $x(t) \geq 0$ and $x(t) \rightarrow 0$ as $t \rightarrow \infty$.*

By using Theorems 2.1.2 and 1.2.2, we obtain the following characterization for positive stability :

**Theorem 2.3.1** *A LTI system $R = [A, 0]$ is positively stable if and only if $A$ is a stable Metzler matrix.*

**Remark 2.3.1** *Note that the previous theorem corresponds to Theorem 1.3.1 with $\bar{x} = 0$.*

Then a characterization of positive stability can be expressed by using the Lyapunov equation (1.37), see [BEFB94] :

**Theorem 2.3.2** *A LTI system $R = [A, 0]$ is positively stable if and only if $A$ is a Metzler matrix and if there exists a diagonal positive definite matrix $P$ such that $P\,A^T + A\,P$ is negative definite.*

### 2.3.2 Positive stabilizability of LTI systems

**Definition 2.3.2** *A LTI system $R = [A,\, B]$ is said to be* **positively stabilizable** *if for all $x_0 \geq 0$, there exists an input $u(t)$ such that for all $t \geq 0$, the state trajectories $x(t)$ is such that $x(t) \geq 0$ and $x(t) \to 0$ as $t \to \infty$.*

The following criteria of positive stabilizability are inspired by [BNS89, Chapter 7] and [Ava00, pp. 73-75] :

**Theorem 2.3.3** *A LTI system $[A + B\,K,\, 0]$ is positively stabilizable if and only if there exists a state feedback law $u(t) = K\,x(t)$ such that $A + B\,K$ is a stable Metzler matrix.*

**Remark 2.3.2** *Let us notice that the previous result corresponds to Theorem 1.3.3 with $\bar{x} = 0$.*

A characterization of positive stabilizability can also be expressed by using the Lyapunov equation (1.37)

**Theorem 2.3.4** *A LTI system $R = [A,\, B]$ is positively stabilizable if and only if there exist a state feedback law $u(t) = K\,x(t)$ and a diagonal positive definite matrix $P$ such that*

$$P(A + B\,K)^T + (A + B\,K)\,P$$

*is negative definite and $A + B\,K$ is a Metzler matrix.*

Moreover, by applying the change of variables $Y = K\,P$ suggested in [BEFB94, Section 10.3], the previous theorem can be reformulated as follows :

**Theorem 2.3.5** *A LTI system $R = [A,\, B]$ is positively stabilizable if and only if there exist a diagonal positive definite matrix $P$ and a matrix $Y$ such that*

$$P\,A^T + Y^T B^T + A\,P + B\,Y$$

*is negative definite with $A\,P + B\,Y$ a Metzler matrix.*

This change of variables allows us to write the problem of positive stabilization in the form of linear matrix inequalities (LMI) that are used in the resolution of the inverse positive $\boldsymbol{LQ}_+^{\text{inv}}$ problem, see Section 7.2.2.

The problem of positive stabilization is studied in [RD09] where necessary and sufficient conditions are obtained for the stabilization of positive LTI systems using a vertex algorithmic approach.

### 2.3.3 Compartmental systems

Finally a particular class of positive systems is briefly introduced in this subsection, namely the class of compartmental systems, see e.g. [God83], [Van98], [BF02] or [HCH10]. Compartmental models are widely used in e.g. biology, pharmacology and physiology to describe the distribution of a substance (e.g. biomass, drug, ...) among different tissues of an organism.

**Definition 2.3.3**

- *A matrix $A$ is said to be a* **compartmental matrix** *if $A$ is a Metzler matrix and for all* $j = 1, ..., n, \ \ \sum_{i=1}^{n} a_{ij} \leq 0.$
- *A positive LTI system $R = [A,\, B]$ is said to be a* **compartmental system** *if $A$ is a compartmental matrix.*
- *A matrix $A$ is said to be* **reducible** *if there exists a permutation matrix $P$ such that*

$$P\, A\, P^T = \begin{bmatrix} U & 0 \\ Q & V \end{bmatrix},$$

*where $U$ and $V$ are square matrices.*
*A matrix $A$ is said to be* **irreducible** *if $A$ is not reducible.*

In the sequel (see Chapter 6 on the positive $\boldsymbol{LQ}_{+}^{\infty}$ problem in infinite horizon) we consider only a compartmental matrix $A$ such that

$$\sum_{i=1}^{n} a_{ij} \ < \ 0 \ \text{ for all } \ j = 1, ..., n. \tag{2.3}$$

By [Van98, Prop. 3.2, p. 594], (2.3) is a sufficient condition for the stability of an irreducible compartmental matrix $A$. Moreover, in many references, a compartmental system is considered with $B$ equal to the identity matrix (which corresponds to the case where there are external inputs for each compartment). Here we have another result on the stability of such systems, see e.g. [Van98] and [BF02] :

**Proposition 2.3.6** *A compartmental system $R = [A,\, B]$ is stable if and only if 0 is not an eigenvalue of $A$, i.e. $0 \notin \sigma(A)$.*

We conclude this short subsection with a result on the positive stabilizability of compartmental systems.

**Theorem 2.3.7** *A compartmental system $R = [A,\, I]$ is positively stabilizable.*

**Proof :** First assume that $A$ is a nonsingular matrix. Then by Proposition 2.3.6, the system $R = [A,\, B]$ is stable. Therefore by setting $K = 0$ the matrix $A + B\,K$ is a stable Metzler matrix (since $A$ is a compartmental matrix). Hence by Theorem 2.3.3, the system is positively stabilizable.

Now assume that $A$ is a singular matrix. Set $K = -\sigma\, I$ with $\sigma > 0$, then $A + B\, K = A - \sigma\, I$. So the state feedback $K$ will move the unstable eigenvalue $0$ such that $A + B\, K$ will be stable. Moreover, since $A$ is a Metzler matrix, so is the matrix $A + B\, K = A - \sigma\, I$. Hence by Theorem 2.3.3, the system is positively stabilizable. So the property of positive stabilizability is automatically verified for a compartmental system.

$\square$

# Part II

# The Invariant Linear Quadratic Problem

This second part is the main part of this thesis, namely the study of the input/state-invariant linear quadratic (LQ) problem for linear continuous time systems. First of all, the problem is studied for a finite final time and then in infinite horizon. The main objective of the input/state-invariant LQ problem is to ensure constraints of lower bound type on the state and/or on the input trajectories. When optimal control is applied to the system, the resulting state and/or input trajectories satisfy lower bound conditions.

In the first chapter, the input/state-invariant LQ problem is studied, i.e. the finite horizon LQ-optimal control problem with affine inequality constraints on the state and/or the input trajectories. Necessary and sufficient optimality conditions are obtained by using the maximum principle with state and input constraints (see e.g. [HSV95]). In addition, in the case of state constraints or input constraints only, necessary and sufficient conditions are proved for the invariant LQ-optimal control to be given by the standard LQ-optimal state feedback law.

In the second chapter, the positive LQ problem is studied, i.e. the particular LQ problem for nonnegative state constraints. In this case, necessary and sufficient optimality conditions are also established, which are based on the maximum principle and on the admissibility of the solution of the standard LQ problem. In addition, criteria for the positivity of the standard LQ closed-loop system are studied. Sufficient conditions are stated in terms of the matrix solution of the Riccati differential equation. Moreover, the particular problem of minimal energy control with penalization of the final state is studied. The main results are illustrated by numerical examples.

In the next chapters, the infinite horizon input/state-invariant LQ problem is studied by means of a receding horizon approach, see e.g. [WC83]. Criteria for the existence of a solution to the positive LQ problem in infinite horizon are established. These criteria use, respectively, a Newton-like iterative scheme (inspired by [GL00a, GL00b]), an Hamiltonian approach and the study of a diagonal solution of the Algebraic Riccati Equation (ARE). Finally, the last chapter of this part is devoted to the inverse input/state-invariant LQ problem by using linear or bilinear matrix inequalities, see e.g. [BEFB94] and [SW05]. The main results are also illustrated by numerical simulations.

# II.1. Finite Horizon Case

# Chapter 3

# The Input/State-Invariant LQ Problem

This chapter is devoted to the finite horizon input/state-invariant linear quadratic (LQ) problem, i.e. the LQ-optimal control problem with affine inequality constraints on the state and on the input trajectories. Optimality conditions are established by using the maximum principle, see [HSV95]. These conditions characterize the solution of the invariant LQ problem by means of a corresponding Hamiltonian system, both in a vector form and in a matrix form. These Hamiltonian equations depend on the initial condition $x_0$ and that makes them difficult to solve. In the discrete time case, see Chapter 10, the optimality conditions lead to a computational method for the solution of the positive LQ problem ; algorithms are therefore developed to compute the solution. Then the particular problems of LQ-optimal control with either state or input constraints are studied. In these cases, optimality conditions are stated which are based on the admissibility of the solution of the standard LQ problem, see e.g. [AM90] and [CD91].

The linear quadratic (LQ) problem with constraints has already been studied for linear systems with positive controls (see e.g. [HVS98] and the references therein for the general LQ problem with positive controls, see [Goe10] for the infinite horizon LQ problem with conical control constraints ; see also [Ka 02] for the minimal energy positive control problem for positive systems). In [HCPH10], the convergence of a discretization method is established for approximating an optimal solution of LQ problem with mixed linear state-control constraints. The theoretical results developed in this chapter are illustrated numerically in the following chapter for the particular case of the positive LQ problem and also in Chapter 9 on the application of these results to the problem of coexistence in a chemostat model.

## 3.1 Problem statement

Consider the following linear time-invariant system description $R = [A,\, B]$, for $t \in [t_0,\, t_f]$ :

$$\dot{x}(t) = A\,x(t) + B\,u(t), \;\; x(t_0) = x_0, \tag{3.1}$$

where, as previously, the state $x(t)$ and the control $u(t)$ are in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, $A$ and $B$ are real matrices of compatible sizes, $x_0 \in \mathbb{R}^n$ denotes any fixed initial state and $\bar{x}$ is a fixed state.

The finite horizon input/state-invariant linear quadratic problem, which is denoted by $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$, consists of minimizing the quadratic functional :

$$J(x_0,\, u,\, t_f) := \frac{1}{2}\left(\int_{t_0}^{t_f} (\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2)\,\mathrm{dt} + x(t_f)^T S\,x(t_f)\right) \tag{3.2}$$

for a given linear system described by (3.1), where the initial state $x_0$ is fixed such that $W\,x_0 \geq\ \bar{x}$, under the constraint

$$\forall\, t \in [t_0,\, t_f],\ \begin{cases} W\,x(t) \geq \bar{x} \\ Z\,u(t) \geq \bar{u}, \end{cases} \tag{3.3}$$

where $t_f$ is a fixed final time, $u$ is any piecewise-continuous $\mathbb{R}^m$-valued function, $R \in \mathbb{R}^{m\times m}$ is a symmetric positive definite matrix, $C \in \mathbb{R}^{p\times n}$, $S \in \mathbb{R}^{n\times n}$ is a symmetric positive semidefinite matrix, $W \in \mathbb{R}^{n\times n}$, $Z \in \mathbb{R}^{m\times m}$, $\bar{x} \in \mathbb{R}^n$ and $\bar{u} \in \mathbb{R}^m$ are fixed state and input (respectively). The problem can be studied with any matrices $W$ and $Z$ (of full rank, see below). However in the sequel, the particular cases where $W$ and $Z$ are equal to either the identity matrix or the zero matrix are studied. Moreover, when $W$ and $Z$ are equal to the zero matrix, the $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem corresponds to the standard $\boldsymbol{LQ}^{\boldsymbol{t_f}}$ problem, see Section 3.3.

The idea of studying this kind of problem comes notably from an application to the chemostat model (see Chapter 9) where several species are in competition for a single nutrient. This model involves the study of a nonlinear system for which the objective is to ensure the coexistence of species, i.e. to force the concentration variables to be strictly positive. The study is done by means of the linearized system around an equilibrium $(x_e,\, u_e)$ such that $x_e := -\bar{x} \gg 0$ and $u_e := -\bar{u} \gg 0$. Thus the objective of coexistence implies inequality constraints on the state and input trajectories of the linearized system. Then, with the goal of being consistent with this application, $\bar{x}$ and $\bar{u}$ are sometimes imposed to satisfy the inequalities $\bar{x} \ll 0$ and $\bar{u} \ll 0$. Moreover, the next chapter is devoted to the positive $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{+}$ problem, i.e. where $W = I_n$, $Z = 0_m$, and in this case we consider $\bar{x} = 0$. Thus, at the beginning of this chapter, $\bar{x}$ and $\bar{u}$ are fixed in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively but when the results of Chapter 1 are used, in Section 3.5, we assume that $\bar{x} \ll 0$ and $\bar{u} \leq 0$.

## 3.2 Optimality conditions

In this section, optimality conditions for the $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem are established. Applying the maximum principle with state and input constraints (see e.g. [HSV95] recalled in Appendix B) yields a characterization, in vector and matrix forms for an $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$-optimal control. In the sequel, unless otherwise stated, the matrices $W$ and $Z$ are assumed to be of full rank (assumptions which translate the constraint qualifications (B.4) and (B.5) for the maximum principle).

**Theorem 3.2.1 (Optimality conditions based on the maximum principle)**
*Consider the $\boldsymbol{LQ}^{t_f}_{\bar{u},\bar{x}}$ problem with cost (3.2) under the dynamics constraints (3.1) with a fixed initial condition $x_0$ such that $W\,x_0 \geq \bar{x}$, under the inequality constraints (3.3).*

*a) The $\boldsymbol{LQ}^{t_f}_{\bar{u},\bar{x}}$ problem has a solution $u(\cdot)$ if and only if there exist piecewise continuous multiplier functions $\lambda(\cdot)$ mapping $[t_0,\, t_f]$ into $\mathbb{R}^n$ and $\upsilon(\cdot)$ mapping $[t_0,\, t_f]$ into $\mathbb{R}^m$ such that*

$$u(t) = -R^{-1}B^T p(t) + R^{-1}Z^T\upsilon(t),\; t \in [t_0,\, t_f], \tag{3.4}$$

*where $[x(t)^T \quad p(t)^T]^T \in \mathbb{R}^{2n}$ is the solution of the Hamiltonian differential equation*

$$\begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = H \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} + \begin{bmatrix} B\,R^{-1}Z^T\upsilon(t) \\ W^T\lambda(t) \end{bmatrix},\; t \in [t_0,\, t_f] \tag{3.5}$$

*with*

$$\begin{cases} x(t_0) &= x_0, \\ p(t_f) &= S\,x(t_f) - W^T\lambda(t_f) \end{cases} \tag{3.6}$$

*where*

$$H = \begin{bmatrix} A & -B\,R^{-1}B^T \\ -C^TC & -A^T \end{bmatrix} \tag{3.7}$$

*is the Hamiltonian matrix, and for all $t \in [t_0,\, t_f]$,*

$$\begin{cases} W\,x(t) \geq \bar{x}, \\ Z\,u(t) \geq \bar{u}, \\ \lambda(t) \geq 0, \\ \upsilon(t) \geq 0, \\ \lambda(t)^T\,(W\,x(t) - \bar{x}) = 0, \quad \textit{(state complementarity conditions)} \\ \upsilon(t)^T\,(Z\,u(t) - \bar{u}) = 0 \quad \textit{(input complementarity conditions)} \end{cases} \tag{3.8}$$

*b) Assume that $(A,\, B)$ is controllable and that $\bar{x} \leq 0$ and $\bar{u} \leq 0$. By using the matrix form of the Hamiltonian differential equation (3.5), a piecewise-continuous control function $u\,:\,[t_0,\, t_f] \to \mathbb{R}^m$ is solution of the $\boldsymbol{LQ}^{t_f}_{\bar{u},\bar{x}}$ problem if and only if there exist piecewise continuous multiplier matrix functions $\Lambda(\cdot)$ mapping $[t_0,\, t_f]$ into $\mathbb{R}^{n\times n}$ and $\Upsilon(\cdot)$ mapping $[t_0,\, t_f]$ into $\mathbb{R}^{m\times m}$ such that for $t \in [t_0,\, t_f]$,*

$$\begin{aligned} u(t) &= K(t,\, x_0)\,x(t) := (-R^{-1}B^TY(t) + R^{-1}Z^T\Upsilon(t))\,X(t)^{-1}x(t) \\ &= (-R^{-1}B^TY(t) + R^{-1}Z^T\Upsilon(t))\,X(t_0)^{-1}x_0, \end{aligned} \tag{3.9}$$

*where $[X(t)^T \quad Y(t)^T]^T \in \mathbb{R}^{2n\times n}$ is the solution of the matrix Hamiltonian differential equation*

$$\begin{bmatrix} \dot{X}(t) \\ \dot{Y}(t) \end{bmatrix} = H \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} + \begin{bmatrix} B\,R^{-1}Z^T\Upsilon(t) \\ W^T\Lambda(t) \end{bmatrix},\; t \in [t_0,\, t_f] \tag{3.10}$$

*with the final condition*

$$\begin{cases} X(t_f) &= I_n, \\ Y(t_f) &= S - W^T\Lambda(t_f) \end{cases} \tag{3.11}$$

*and for all* $t \in [t_0, t_f]$,

$$W\,X(t)\,X(t_0)^{-1}x_0 \geq \bar{x} \tag{3.12a}$$
$$Z\,(-R^{-1}\,B^TY(t) + R^{-1}Z^T\Upsilon(t))\,X(t_0)^{-1}x_0 \geq \bar{u} \tag{3.12b}$$
$$\Lambda(t)\,X(t_0)^{-1}x_0 \geq 0, \tag{3.12c}$$
$$\Upsilon(t)\,X(t_0)^{-1}x_0 \geq 0, \tag{3.12d}$$
$$x_0^TX(t_0)^{-T}\Lambda(t)^T(W\,X(t)\,X(t_0)^{-1}x_0 - \bar{x}) = 0 \quad \textit{(state complementarity condition)} \tag{3.12e}$$
$$x_0^TX(t_0)^{-T}\Upsilon(t)^T(Z\,u(t) - \bar{u}) = 0 \quad \textit{(input complementarity condition)} \tag{3.12f}$$

**Proof of Theorem 3.2.1 a) Necessity :**
First use the maximum principle with state and input constraints, see [HSV95, Theorem 4.1] which is recalled in Appendix B, with the following identifications :

$$\begin{array}{lcl}
-F(x(t),\,u(t),\,t) & = & \dfrac{1}{2}\left(\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2\right) \\
-S(x(T),\,T) & = & \dfrac{1}{2}x(t_f)^TS\,x(t_f) \\
f(x(t),\,u(t),\,t) & = & A\,x(t) + B\,u(t) \\
g(x(t),\,u(t),\,t) & = & Z\,u(t) - \bar{u} \\
h(x(t),\,t) & = & W\,x(t) - \bar{x} \\
a(x(T),\,T) & = & 0 \\
b(x(T),\,T) & = & 0.
\end{array}$$

Then, with $\lambda(t) := p(t)$, $\lambda_0 := 1$ (normal case), $\mu(t) := \upsilon(t)$, $\upsilon(t) := \lambda(t)$,

$$\begin{array}{lcl}
H(x,\,u,\,p,\,t) & := & -\dfrac{1}{2}\left(\|R^{1/2}u\|^2 + \|C\,x\|^2\right) + p^T(A\,x + B\,u) \\
L(x,\,u,\,p,\,\upsilon,\,\lambda) & := & H(x,\,u,\,p,\,t) + \upsilon^T\,g(x,\,u,\,t) + \lambda^T\,h(x,\,t),
\end{array}$$

conditions (B.8b)-(B.8e) become :

- $-R\,u(t) + B^Tp(t) + Z^T\upsilon(t) = 0 \;\Rightarrow\; u(t) = R^{-1}B^Tp(t) + R^{-1}Z^T\upsilon(t)$
- $\dot{p}(t) = C^TC\,x(t) - A^Tp(t) - W^T\lambda(t)$
- $\upsilon(t) \geq 0,\; \upsilon(t)^T(Z\,u(t) - \bar{u}) = 0$

and at the terminal time $t_f$, transversality conditions (B.9a)-(B.9c) read :

- $p(t_f) = -S\,x(t_f) + W^T\gamma,$
- $\gamma \geq 0$
- $\gamma^T\,(W\,x(t_f) - \bar{x}) = 0.$

Therefore we obtain the following two-point boundary value problem, with $p(t)$ replaced by $-p(t)$ :

$$\begin{cases} \dot{x}(t) &= A\,x(t) + B\,u(t) \\ \dot{p}(t) &= -A^T p(t) - C^T C\,x(t) + W^T \lambda(t) \end{cases}$$

$$\text{with} \quad \begin{cases} x(t_0) &= x_0 \\ p(t_f) &= S\,x(t_f) - W^T\gamma = S\,x(t_f) - W^T\lambda(t_f) \end{cases} \tag{3.13}$$

$$\text{where} \quad u(t) = -R^{-1}B^T p(t) + R^{-1}Z^T \upsilon(t)$$

under the constraints

$$\begin{cases} W\,x(t) &\geq \bar{x} \\ \lambda(t)^T(W\,x(t) - \bar{x}) &= 0 \\ \gamma^T\,(W\,x(t_f) - \bar{x}) &= 0, \text{ with } \gamma = \lambda(t_f), \\ \lambda(t) &\geq 0 \end{cases}$$

and

$$\begin{cases} Z\,u(t) &\geq \bar{u} \\ \upsilon(t)^T(Z\,u(t) - \bar{u}) &= 0 \\ \upsilon(t) &\geq 0 \end{cases}$$

Hence, $[x(t)^T \quad p(t)^T]^T \in \mathbb{R}^{2n}$ is the solution of the Hamiltonian differential equation (3.5). □

In order to prove the sufficiency, the following concepts and lemma are needed :

**Definition 3.2.1**

- *A pair* $(x(t),\, u(t))$ *is said to be* **(dynamically) admissible** *with respect to the* $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ *problem if* $u \in \mathcal{U}$ *and* $\dot{x}(t) = A\,x(t) + B\,u(t)$ *where the initial state* $x_0$ *is fixed.*
- *A pair* $(x^o(t),\, u^o(t))$ *is said to be* **optimal** *with respect to the* $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ *problem if it is admissible and minimizes the cost (3.2), whence* $[x^o(t)^T \quad p^o(t)^T]^T$ *is solution of the Hamiltonian differential equation (3.5) and the control* $u^o(t)$ *is given by* $u^o(t) = -R^{-1}B^T p^o(t) + R^{-1}Z^T\upsilon(t)$*, for* $t \in [t_0,\, t_f]$*.*

**Lemma 3.2.2 (Evaluation Lemma)** *Consider an optimal pair* $(x^o(t),\, u^o(t))$ *with respect to* $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ *problem. Then for any* $\tau \in [t_0,\, t_f)$*, with* $\lambda(t)$ *and* $\upsilon(t)$*, the multipliers associated to* $x^o(t)$ *and* $u^o(t)$ *respectively,*
*a) For all admissible pair* $(x(t),\, u(t))$*,*

$$\begin{aligned} &\int_\tau^{t_f} < x(t),\, C^T C\, x^o(t) > + < u(t),\, R\, u^o(t) > \,dt \\ &= -p^o(t_f)^T x(t_f) + p^o(\tau)^T x(\tau) + \int_\tau^{t_f} \lambda(t)^T W\, x(t)\, dt + \int_\tau^{t_f} \upsilon(t)^T Z\, u(t)\, dt. \end{aligned} \tag{3.14}$$

*where* $< \cdot,\, \cdot >$ *denotes the scalar product defined as follows :*

$$\begin{aligned} < \cdot,\, \cdot >: \quad \mathbb{R}^k \times \mathbb{R}^k &\rightarrow \mathbb{R} \\ (a,\, b) &\rightsquigarrow \; < a,\, b >= b^T a \end{aligned}$$

*for any vectors* $a,\, b \in \mathbb{R}^k$ *where* $k \in \mathbb{N}$*.*

*b) In particular,* $(x^o(t),\, u^o(t))$ *is an admissible pair, whence*

$$\begin{aligned}&\int_\tau^{t_f} (\|R^{1/2}u^o(t)\|^2 + \|C\,x^o(t)\|^2)\,dt \\ &= -p^o(t_f)^T x^o(t_f) + p^o(\tau)^T x^o(\tau) + \int_\tau^{t_f} \lambda(t)^T W\,x^o(t)\,dt + \int_\tau^{t_f} \upsilon(t)^T Z\,u^o(t)\,dt.\end{aligned} \tag{3.15}$$

**Proof :** a) Using (3.5) with $u^o(t) = -R^{-1}B^T p^o(t) + R^{-1}Z^T\upsilon(t)$ gives :

$$\begin{aligned}
& < x(t),\, C^T C\,x^o(t) > + < u(t),\, R\,u^o(t) > \\
= \;& < x(t),\, -A^T p^o(t) + W^T\lambda(t) - \dot{p}^o(t) > + < u(t),\, -B^T p^o(t) + Z^T\upsilon(t) > \\
= \;& - < x(t),\, \dot{p}^o(t) > - < x(t),\, A^T p^o(t) > - < u(t),\, B^T p^o(t) > + < x(t),\, W^T\lambda(t) > \\
& + < u(t),\, Z^T\upsilon(t) > \\
= \;& - < x(t),\, \dot{p}^o(t) > - < \dot{x}(t),\, p^o(t) > + < x(t),\, W^T\lambda(t) > + < u(t),\, Z^T\upsilon(t) >
\end{aligned}$$

Hence

$$\begin{aligned}
& \int_\tau^{t_f} < x(t),\, C^T C\,x^o(t) > + < u(t),\, R\,u^o(t) > \;\mathrm{dt} \\
= \;& -\int_\tau^{t_f} (< x(t),\, \dot{p}^o(t) > + < \dot{x}(t),\, p^o(t) >)\,\mathrm{dt} + \int_\tau^{t_f} < x(t),\, W^T\lambda(t) > \;\mathrm{dt} \\
& + \int_\tau^{t_f} < u(t),\, Z^T\upsilon(t) > \;\mathrm{dt} \\
= \;& -\int_\tau^{t_f} \frac{d}{dt} < x(t),\, p^o(t) > \;\mathrm{dt} + \int_\tau^{t_f} < x(t),\, W^T\lambda(t) > \;\mathrm{dt} + \int_\tau^{t_f} < u(t),\, Z^T\upsilon(t) > \;\mathrm{dt} \\
= \;& -[< x(t),\, p^o(t) >]_\tau^{t_f} + \int_\tau^{t_f} < x(t),\, W^T\lambda(t) > \;\mathrm{dt} + \int_\tau^{t_f} < u(t),\, Z^T\upsilon(t) > \;\mathrm{dt} \\
= \;& -p^o(t_f)^T x(t_f) + p^o(\tau)^T x(\tau) + \int_\tau^{t_f} \lambda(t)^T W\,x(t)\,\mathrm{dt} + \int_\tau^{t_f} \upsilon(t)^T Z\,u(t)\,\mathrm{dt}.
\end{aligned}$$

b) Using the Hamiltonian differential equation (3.5),

$$\begin{aligned}
\dot{x}^o(t) &= A\,x^o(t) - B\,R^{-1}B^T p^o(t) + B\,R^{-1}Z^T\upsilon(t) \\
&= A\,x^o(t) + B\,u^o(t) \;\text{ (by the expression of } u^o(t))
\end{aligned}$$

Then $(x^o(t),\, u^o(t))$ is an admissible pair. Hence (3.14) is verified for $(x(t),\, u(t)) = (x^o(t),\, u^o(t))$, that is (3.15) holds.

□

**Proof of Theorem 3.2.1 a) Sufficiency :**
By the fact that the functional (3.2) is convex and the dynamics and inequality constraints (3.1) and (3.3) are defined by affine functions, see [CD91, pp. 31-32], by a comparison of costs, we obtain the result. Indeed, let us compute the cost for any admissible control $u$ :

$$\begin{aligned}
2\,J(x_0,\, u,\, t_f) &= \int_{t_0}^{t_f} (\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2)\,\mathrm{dt} + \underbrace{x(t_f)^T S\,x(t_f)}_{:=\|x(t_f)\|_S^2} \\
&= \underbrace{\int_{t_0}^{t_f} (\|R^{1/2}u(t) - R^{1/2}u^o(t)\|^2 + \|C\,x(t) - C\,x^o(t)\|^2)\,\mathrm{dt}}_{(I)} + \underbrace{\|x(t_f) - x^o(t_f)\|_S^2}_{(II)} \\
&\quad + \underbrace{2\int_{t_0}^{t_f} (< u(t),\, R\,u^o(t) > + < x(t),\, C^T C\,x^o(t) >)\,\mathrm{dt}}_{(III)}
\end{aligned}$$

$$
\begin{aligned}
& \underbrace{- \int_{t_0}^{t_f} (\|R^{1/2} u^o(t)\|^2 + \|C\,x^o(t)\|^2)\,\mathrm{dt}}_{(IV)} \\
& + \underbrace{2\,x(t_f)^T S\,x^o(t_f)}_{(V)} \underbrace{-x^o(t_f)^T S\,x^o(t_f)}_{(VI)} \\
= \quad & (I) + (II) + (III) + (IV) + (V) + (VI)
\end{aligned}
$$

where

- By Lemma 3.2.2, identity (3.14) with $\tau = t_0$ reads :

$$(III) = -2\,p^o(t_f)^T x(t_f) + 2\,p^o(t_0)^T x_0 + 2\int_{t_0}^{t_f} \lambda(t)^T W\,x(t)\,\mathrm{dt} + 2\int_{t_0}^{t_f} \upsilon(t)^T Z\,u(t)\,\mathrm{dt}$$

- By Lemma 3.2.2, identity (3.15) with $\tau = t_0$ reads :

$$(IV) = p^o(t_f)^T x^o(t_f) - p^o(t_0)^T x_0 - \int_{t_0}^{t_f} \lambda(t)^T W\,x^o(t)\,\mathrm{dt} - \int_{t_0}^{t_f} \upsilon(t)^T Z\,u^o(t)\,\mathrm{dt}$$

- By the final condition (3.6), $S\,x^o(t_f) = p^o(t_f) + W^T\lambda(t_f)$, whence

$$
\begin{aligned}
(V) &= 2\,p^o(t_f)^T x(t_f) + 2\,x(t_f)^T W^T \lambda(t_f) \\
\text{and} \quad (VI) &= -p^o(t_f)^T x^o(t_f) - x^o(t_f)^T W^T \lambda(t_f)
\end{aligned}
$$

Therefore, after simplifications, one gets :

$$
\begin{aligned}
2\,J(x_0,\,u,\,t_f) \;=\; & (I) + (II) + p^o(t_0)^T x_0 \\
& + 2\int_{t_0}^{t_f} \lambda(t)^T W\,x(t)\,\mathrm{dt} + 2\int_{t_0}^{t_f} \upsilon(t)^T Z\,u(t)\,\mathrm{dt} \\
& - \int_{t_0}^{t_f} \lambda(t)^T W\,x^o(t)\,\mathrm{dt} - \int_{t_0}^{t_f} \upsilon(t)^T Z\,u^o(t)\,\mathrm{dt} \\
& + 2\,x(t_f)^T W^T \lambda(t_f) - x^o(t_f)^T W^T \lambda(t_f)
\end{aligned}
$$

On the other hand, by computing the cost for an optimal pair $(x^o(t),\,u^o(t))$, we obtain, by using Lemma 3.2.2, identity (3.15) with $\tau = t_0$ and the final condition (3.6) for the adjoint state :

$$
\begin{aligned}
2\,J(x_0,\,u^o,\,t_f) &= \int_{t_0}^{t_f} (\|R^{1/2} u^o(t)\|^2 + \|C\,x^o(t)\|^2)\,\mathrm{dt} + x^o(t_f)^T S\,x^o(t_f) \\
&= -p^o(t_f)^T x^o(t_f) + p^o(t_0)^T x_0 + \int_{t_0}^{t_f} \lambda(t)^T W\,x^o(t)\,\mathrm{dt} + \int_{t_0}^{t_f} \upsilon(t)^T Z\,u^o(t)\,\mathrm{dt} \\
&\quad + p^o(t_f)^T x^o(t_f) + x^o(t_f)^T W^T \lambda(t_f) \\
&= p^o(t_0)^T x_0 + x^o(t_f)^T W^T \lambda(t_f) + \int_{t_0}^{t_f} \lambda(t)^T W\,x^o(t)\,\mathrm{dt} + \int_{t_0}^{t_f} \upsilon(t)^T Z\,u^o(t)\,\mathrm{dt}
\end{aligned}
$$

Now let us compute the difference between these two costs and use the complementarity conditions of the optimal pair $(x^o(t),\, u^o(t))$ :

$$\begin{array}{rl} & 2\,(J(x_0,\,u,\,t_f) - J(x_0,\,u^o,\,t_f)) \\ = & (I) + (II) + 2\displaystyle\int_{t_0}^{t_f} \lambda(t)^T(W\,x(t)-\bar{x})\,\mathrm{dt} + 2\int_{t_0}^{t_f} \upsilon(t)^T(Z\,u(t)-\bar{u})\,\mathrm{dt} \\ & -2\displaystyle\int_{t_0}^{t_f} \underbrace{\lambda(t)^T(W\,x^o(t)-\bar{x})}_{=0}\,\mathrm{dt} - 2\int_{t_0}^{t_f} \underbrace{\upsilon(t)^T(Z\,u^o(t)-\bar{u})}_{=0}\,\mathrm{dt} \\ & +2\,x(t_f)^T W^T \lambda(t_f) - 2\,x^o(t_f)^T W^T \lambda(t_f) \\ = & (I) + (II) + 2\displaystyle\int_{t_0}^{t_f} \lambda(t)^T(W\,x(t)-\bar{x})\,\mathrm{dt} + 2\int_{t_0}^{t_f} \upsilon(t)^T(Z\,u(t)-\bar{u})\,\mathrm{dt} \\ & +2\,(W\,x(t_f)-\bar{x})^T\lambda(t_f) - 2\,\underbrace{(W\,x^o(t_f)-\bar{x})^T\lambda(t_f)}_{=0} \end{array} \tag{3.16}$$

Hence $2\,(J(x_0,\,u,\,t_f) - J(x_0,\,u^o,\,t_f)) \geq 0$ for all admissible $u$. Indeed, $(I)+(II) \geq 0$, $\lambda(t) \geq 0$, $\upsilon(t) \geq 0$, $W\,x(t)-\bar{x} \geq 0$ and $Z\,u(t)-\bar{u} \geq 0$ for all time $t \in [t_0,\,t_f]$, then each term of equation (3.16) is nonnegative. Therefore $u^o(t)$ given by identity (3.4) is optimal.

□

**Proof of Theorem 3.2.1 b) :**

Necessity :

- Consider an initial condition $x(t_0) = x_0$ such that $W\,x_0 \geq \bar{x}$. Assume that the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ has a solution $u(\cdot)$. Then by Theorem 3.2.1 a), $u(t)$ is given by (3.4) where $[x(t)^T \quad p(t)^T \quad \upsilon(t)^T \quad \lambda(t)^T]^T$ is solution of (3.5) and satisfies (3.6) and (3.8) and where it should be noted that the multipliers are not necessarily unique. In particular, $[x(t)^T \quad p(t)^T]^T$ is solution of the following two-point boundary value problem

$$\left\{ \begin{array}{rcl} \dot{x}(t) & = & A\,x(t) + B\,u(t) \\ \dot{p}(t) & = & -A^T p(t) - C^T C\,x(t) + W^T\lambda(t) \end{array} \right.$$

$$\text{with} \quad \left\{ \begin{array}{rcl} x(t_0) & = & x_0 \\ p(t_f) & = & S\,x(t_f) - W^T\lambda(t_f) \end{array} \right. \tag{3.17}$$

$$\text{where} \quad u(t) = -R^{-1}B^T p(t) + R^{-1}Z^T\upsilon(t)$$

under the constraints

$$\left\{ \begin{array}{rcl} W\,x(t) & \geq & \bar{x} \\ \lambda(t)^T(W\,x(t)-\bar{x}) & = & 0 \\ \lambda(t) & \geq & 0 \end{array} \right.$$

and

$$\left\{ \begin{array}{rcl} Z\,u(t) & \geq & \bar{u} \\ \upsilon(t)^T(Z\,u(t)-\bar{u}) & = & 0 \\ \upsilon(t) & \geq & 0 \end{array} \right.$$

- Observe that one can find piecewise continuous matrix functions $\Upsilon(\cdot)$ and $\Lambda(\cdot)$ such that (3.17) holds with
$$\begin{bmatrix} \upsilon(t) \\ \lambda(t) \end{bmatrix} := \begin{bmatrix} \Upsilon(t) \\ \Lambda(t) \end{bmatrix} \zeta \tag{3.18}$$
where $\zeta := x(t_f) \in \mathbb{R}_0^n$ corresponds to the final state given by (3.17).
Indeed, first consider the pair $(H,\, \mathcal{B})$, where
$$\mathcal{B} := \begin{bmatrix} B\,R^{-1}Z^T\ [\zeta_1\, I_n \ldots \zeta_n\, I_n] & 0_{n\times n^2} \\ 0_{n\times n^2} & W^T\ [\zeta_1\, I_n \ldots \zeta_n\, I_n] \end{bmatrix} \in \mathbb{R}^{2n\times 2n^2}.$$
Observe that the pair $(H, \mathcal{B})$ is controllable, since the matrix $[s\, I_{2n} - H \quad \mathcal{B}]$, $\forall\, s \in \mathbb{C}$ is of full rank, which corresponds to a well-known controllability rank test.
Actually, $\forall\, s \in \mathbb{C}$,
$$[s\, I_{2n} - H \quad \mathcal{B}] = \begin{bmatrix} s\, I_n - A & B\,R^{-1}B^T & B\,R^{-1}Z^T\ [\zeta_1\, I_n \ldots \zeta_n\, I_n] & 0_{n\times n^2} \\ C^TC & s\, I_n + A^T & 0_{n\times n^2} & W^T\ [\zeta_1\, I_n \ldots \zeta_n\, I_n] \end{bmatrix}$$
where
  - ⋆ $[s\, I_n - A \quad B\,R^{-1}Z^T\ [\zeta_1\, I_n \ldots \zeta_n\, I_n]]$ is of full rank $n$ since $\zeta_i \neq 0$ for some $i \in \{1, \ldots, n\}$ and $(A,\, B)$ is assumed to be controllable, that is $[s\, I_n - A \quad B]$ is of full rank $n$, for all $s \in \mathbb{C}$ ;
  - ⋆ $W^T\ [\zeta_1\, I_n \ldots \zeta_n\, I_n]$ is of full rank $n$ since $W$ is assumed to be of full rank $n$.

  Then, there exist two submatrices of full rank $n$ such that $[s\, I_{2n} - H \quad \mathcal{B}]$ is of full rank $2n$. It follows that there exists $\Upsilon(\cdot)$ and $\Lambda(\cdot)$ such that the solution $[x(t)^T \quad p(t)^T]^T$ of the two-point boundary value problem (3.17) is solution of the following controlled system, with $\Upsilon(t) = [\Upsilon^1\ \ldots\ \Upsilon^n]\,(t)$ and $\Lambda(t) = [\Lambda^1\ \ldots\ \Lambda^n]\,(t)$, where $A^i$ denotes here the $i^{\text{th}}$ column of a matrix $A$ :
$$\begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = H \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} + \begin{bmatrix} B\,R^{-1}Z^T\Upsilon(t)\,\zeta \\ W^T\Lambda(t)\,\zeta \end{bmatrix},\ t \in [t_0,\, t_f]$$
where $\begin{cases} x(t_0) &= x_0, \\ p(t_f) &= S\,x(t_f) - W^T\lambda(t_f) \end{cases}$ and $\Lambda(t_f)\,\zeta = \lambda(t_f)$ or equivalently,
$$\begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = H \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} + \mathcal{B} \begin{bmatrix} \Upsilon^1 \\ \vdots \\ \Upsilon^n \\ \Lambda^1 \\ \vdots \\ \Lambda^n \end{bmatrix} (t),\ t \in [t_0,\, t_f].$$

- Then one can solve the matrix Hamiltonian differential equation (3.10) with $X(t_f) = I_n$ and $Y(t_f) = S - W^T\Lambda(t_f)$ and obtain its unique solution $[X(t)^T \quad Y(t)^T]^T$. Consequently, thanks to the choice (3.18) of the multiplier matrix functions $\Upsilon(\cdot)$ and $\Lambda(\cdot)$ in the previous step, $\begin{bmatrix} x(t) \\ p(t) \end{bmatrix} = \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} \zeta$ is the solution of the two-point boundary value problem (3.17). In fact, the matrix Hamiltonian differential equation (3.10) post-multiplied by the vector $\zeta$ reads

$$\begin{bmatrix} \dot{X}(t) \\ \dot{Y}(t) \end{bmatrix} \zeta = H \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} \zeta + \begin{bmatrix} B\,R^{-1}Z^T\Upsilon(t) \\ W^T\Lambda(t) \end{bmatrix} \zeta,$$

i.e.

$$\begin{bmatrix} \dot{X}\,\zeta \\ \dot{Y}\,\zeta \end{bmatrix}(t) = \begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = H \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} + \begin{bmatrix} B\,R^{-1}Z^T\upsilon(t) \\ W^T\lambda(t) \end{bmatrix},$$

with

$$\begin{cases} X(t_f)\,\zeta &= I_n\,\zeta = \zeta, \\ Y(t_f)\,\zeta &= (S - W^T\Lambda(t_f))\,\zeta = S\,\zeta - W^T\Lambda(t_f)\,\zeta \end{cases}$$

i.e.

$$\begin{cases} x(t_f) &= \zeta, \\ p(t_f) &= S\,x(t_f) - W^T\lambda(t_f) \end{cases}$$

- Furthermore, the matrix $X(t)$ is invertible for all $t \in [t_0,\, t_f]$. Indeed, by the evaluation Lemma 3.2.2, with $(x(t),\, u(t))$, an optimal pair, for any $\tau \in [t_0,\, t_f)$,

$$\begin{aligned} \int_\tau^{t_f} (\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2)\,\mathrm{dt} &= -p(t_f)^T x(t_f) + p(\tau)^T x(\tau) + \int_\tau^{t_f} \lambda(t)^T W\,x(t)\,\mathrm{dt} \\ &\quad + \int_\tau^{t_f} \upsilon(t)^T Z\,u(t)\,\mathrm{dt} \end{aligned}$$

$$= \; -(x(t_f)^T S\,x(t_f) - \lambda(t_f)^T W\,x(t_f)) + p(\tau)^T x(\tau) + \int_\tau^{t_f} \lambda(t)^T W\,x(t)\,\mathrm{dt} + \int_\tau^{t_f} \upsilon(t)^T Z\,u(t)\,\mathrm{dt}$$

whence

$$\begin{aligned} \int_\tau^{t_f} (\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2)\,\mathrm{dt} + x(t_f)^T S\,x(t_f) &= p(\tau)^T x(\tau) + \lambda(t_f)^T W\,x(t_f) \\ &\quad + \int_\tau^{t_f} \lambda(t)^T W\,x(t)\,\mathrm{dt} + \int_\tau^{t_f} \upsilon(t)^T Z\,u(t)\,\mathrm{dt} \end{aligned}$$

or equivalently,

$$\begin{aligned} \int_\tau^{t_f} (\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2)\,\mathrm{dt} + x(t_f)^T S\,x(t_f) &= p(\tau)^T x(\tau) + \lambda(t_f)^T \bar{x} \\ &\quad + \int_\tau^{t_f} \lambda(t)^T \bar{x}\,\mathrm{dt} + \int_\tau^{t_f} \upsilon(t)^T \bar{u}\,\mathrm{dt} \end{aligned} \tag{3.19}$$

by using the state and the input complementarity conditions : $\lambda(t)^T(W\,x(t) - \bar{x}) = 0$ and $\upsilon(t)^T(Z\,u(t) - \bar{u}) = 0$, $\forall\, t \in [t_0,\, t_f]$.

Now contradiction is used to prove that $X(t)$ is a nonsingular matrix. Since $X(t_f) = I_n$, assume that there exists a time $\tau \in [t_0,\, t_f)$ such that $\det X(\tau) = 0$. Hence, there exists a nonzero vector $\zeta$ such that $X(\tau)\,\zeta = 0$, i.e. $x(\tau) = 0$. Using (3.19) gives

$$\int_\tau^{t_f} (\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2)\,\mathrm{dt} + x(t_f)^T S\,x(t_f) - \lambda(t_f)^T \bar{x} \\ - \int_\tau^{t_f} \lambda(t)^T \bar{x}\,\mathrm{dt} - \int_\tau^{t_f} \upsilon(t)^T \bar{u}\,\mathrm{dt} = 0.$$

Hence, since each term is nonnegative with $\bar{x} \le 0$ and $\bar{u} \le 0$, each term of the previous sum should be equal to zero, and in particular, one gets :

$$\forall\, t \in [\tau,\, t_f]\ :\ u(t) = 0.$$

Therefore system (3.1) becomes $\dot{x}(t) = A\,x(t)$ on $[\tau,\, t_f]$ with $x(\tau) = 0$. Hence its solution is $x(t) = 0,\ \forall\, t \in [\tau,\, t_f]$. In particular, $x(t_f) = 0$ where $x(t_f) = X(t_f)\,\zeta = \zeta$, then $\zeta = 0$ which is a contradiction.

- So, since $\det X(t) \neq 0$, $\forall\, t \in [t_0,\, t_f]$, $\zeta = X(t_0)^{-1}x_0 = X(t)^{-1}x(t),\ \forall\, t \in [t_0,\, t_f]$. Hence, with

$$\begin{bmatrix} x(t) \\ p(t) \\ \upsilon(t) \\ \lambda(t) \end{bmatrix} = \begin{bmatrix} X(t) \\ Y(t) \\ \Upsilon(t) \\ \Lambda(t) \end{bmatrix} \zeta = \begin{bmatrix} X(t) \\ Y(t) \\ \Upsilon(t) \\ \Lambda(t) \end{bmatrix} X(t_0)^{-1}x_0, \tag{3.20}$$

one can easily verify, as previously, that conditions (3.5) and (3.6) become conditions (3.10) and (3.11) respectively, with $x(t_0) = X(t_0)\,\zeta = X(t_0)\,X(t_0)^{-1}x_0 = x_0$. Furthermore, the optimal control is given by

$$\begin{aligned} u(t) &= -R^{-1}B^T p(t) + R^{-1}Z^T \upsilon(t) \\ &= -R^{-1}B^T Y(t)\,\zeta + R^{-1}Z^T \Upsilon(t)\,\zeta \\ &= (-R^{-1}B^T Y(t) + R^{-1}Z^T \Upsilon(t))\,X(t)^{-1}x(t) \\ &= (-R^{-1}B^T Y(t) + R^{-1}Z^T \Upsilon(t))\,X(t_0)^{-1}x_0. \end{aligned} \tag{3.21}$$

Finally, let us verify that conditions (3.8) become conditions (3.12) :

$$\begin{cases} W\,x(t) \ge \bar{x} \ \Leftrightarrow\ W\,X(t)\,\zeta \ge \bar{x} \ \Leftrightarrow\ W\,X(t)\,X(t_0)^{-1}x_0 \ge \bar{x}, \\ Z\,u(t) \ge \bar{u} \ \Leftrightarrow\ Z\,(-R^{-1}B^T Y(t) + R^{-1}Z^T \Upsilon(t))\,X(t_0)^{-1}x_0 \ge \bar{u}, \text{by using (3.21)} \\ \lambda(t) \ge 0 \ \Leftrightarrow\ \Lambda(t)\,\zeta \ge 0 \ \Leftrightarrow\ \Lambda(t)\,X(t_0)^{-1}x_0 \ge 0, \\ \upsilon(t) \ge 0 \ \Leftrightarrow\ \Upsilon(t)\,\zeta \ge 0 \ \Leftrightarrow\ \Upsilon(t)\,X(t_0)^{-1}x_0 \ge 0, \\ \lambda(t)^T\,(W\,x(t) - \bar{x}) = 0 \ \Leftrightarrow\ (\Lambda(t)\,\zeta)^T (W\,X(t)\,\zeta - \bar{x}) = 0 \ \Leftrightarrow\ \zeta^T \Lambda(t)^T (W\,X(t)\,\zeta - \bar{x}) = 0 \\ \qquad\qquad \Leftrightarrow\ (X(t_0)^{-1}x_0)^T \Lambda(t)^T (W\,X(t)\,X(t_0)^{-1}x_0 - \bar{x}) = 0 \\ \qquad\qquad \Leftrightarrow\ x_0^T X(t_0)^{-T} \Lambda(t)^T (W\,X(t)\,X(t_0)^{-1}x_0 - \bar{x}) = 0 \\ \upsilon(t)^T\,(Z\,u(t) - \bar{u}) = 0 \ \Leftrightarrow\ \zeta^T \Upsilon(t)^T (Z\,u(t) - \bar{x}) = 0 \\ \qquad\qquad \Leftrightarrow\ x_0^T X(t_0)^{-T} \Upsilon(t)^T (Z\,u(t) - \bar{u}) = 0 \end{cases}$$

Sufficiency : Let $[X(t)^T\, Y(t)^T]^T$ be the solution of the matrix Hamiltonian differential equation (3.10) which satisfies the final conditions (3.11) and the constraints (3.12). Since $\det X(t) \neq 0$, $\forall\, t \in [t_0,\, t_f]$, for the given initial condition $x(t_0) = x_0$, there exists a unique vector $\zeta$ such that $x_0 = X(t_0)\,\zeta$. Then by (3.9),

$$\begin{aligned} \dot{x} &= A\,x + B\,u \\ &= (A - B\,R^{-1}B^T Y(t)\,X(t)^{-1} - B\,R^{-1}Z^T\Upsilon(t)\,X(t)^{-1})\,x(t). \end{aligned}$$

It follows that

$$\begin{bmatrix} x(t) \\ p(t) \\ \upsilon(t) \\ \lambda(t) \end{bmatrix} := \begin{bmatrix} X(t) \\ Y(t) \\ \Upsilon(t) \\ \Lambda(t) \end{bmatrix} \zeta$$

is the unique solution of the Hamiltonian differential equation (3.5), where $H$ is given by (3.7), such that (3.6) and (3.8) hold. The result follows by Theorem 3.2.1 a).

$\square$

**Remarks 3.2.1** *a) In the sequel, $(A,\, B)$ will be assumed to be controllable whenever Theorem 3.2.1 b) (or its discrete-time version) is used.*
*b) In Theorem 3.2.1, the solution of the Hamiltonian equation, both in its vector form and its matrix form, clearly depends on the initial condition $x_0$. This solution in this case is difficult to compute whereas, as we can see in the following section, for the standard $\boldsymbol{LQ^{t_f}}$ problem, the solution can be computed a priori independently of the initial condition. Moreover, in the discrete time case, an algorithm to compute this solution is described in Subsection 10.2.5.*

# 3.3 Standard LQ problem

The standard LQ problem, denoted by $\boldsymbol{LQ^{t_f}}$ , consists of minimizing the quadratic functional (3.2) for a given linear system described by (3.1) without any constraint on the state trajectory or input trajectory, see e.g. [CD91]. This problem corresponds to the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ problem with $W = 0_n$ and $Z = 0_m$ where, in Theorem 3.2.1 and its proof, the multipliers associated to the state and input constraints together with the associated equations are no longer present. Its solution is given by $u(t) = K(t)\,x(t) = -R^{-1}B^T Y(t)\,X(t)^{-1}x(t),\ t \in [t_0,\, t_f]$ where $[X(t)^T \quad Y(t)^T]^T \in \mathbb{R}^{2n\times n}$ is the solution of the matrix Hamiltonian differential equation

$$\begin{bmatrix} \dot{X}(t) \\ \dot{Y}(t) \end{bmatrix} = H \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix},\ \begin{bmatrix} X(t_f) \\ Y(t_f) \end{bmatrix} = \begin{bmatrix} I \\ S \end{bmatrix}. \tag{3.22}$$

Moreover, in this case, the solution can be rewritten in terms of the Riccati Differential Equation (RDE) $P(\cdot)$. Indeed the solution of the $\boldsymbol{LQ^{t_f}}$ problem is given, for all $t \in [t_0,\, t_f]$, by $u(t) = -R^{-1}B^T P(t)\,x(t)$, where $P(\cdot) = P(\cdot)^T$ is the positive semidefinite matrix solution of the RDE, (see e.g. [CD91]) :

$$-\dot{P}(t) = A^T P(t) + P(t)\,A - P(t)\,B\,R^{-1}B^T P(t) + C^T C,\ \ P(t_f) = S. \tag{3.23}$$

Therefore, in the sequel, the following closed-loop system is considered for $t \in [t_0,\, t_f]$ :

$$\dot{x}(t) = (A + B\,K(t))\,x(t), \quad x(t_0) = x_0, \tag{3.24}$$

where $K(t) = -R^{-1}B^TP(t)$.

## 3.4 Optimality conditions via admissibility

Consider the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ problem with $W = I_n$ and $Z = I_m$. In view of the analysis above, conditions such that the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ problem has a solution can be obtained. These conditions are based on the admissibility of the solution of the standard $\boldsymbol{LQ^{t_f}}$ problem which was described in the previous section. Clearly such optimality conditions are only sufficient conditions of optimality for the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ problem.

**Corollary 3.4.1 (Optimality conditions based on admissibility)** *The solution of the (standard) $\boldsymbol{LQ^{t_f}}$ problem is solution of the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ problem for $x_0 \geq \bar{x}$ if and only if the $\boldsymbol{LQ^{t_f}}$ -optimal state and input trajectories are admissible, i.e. $x(t) \geq \bar{x}$ and $u(t) \geq \bar{u}$ for all $t \in [t_0,\, t_f]$, or equivalently the matrix solution of the standard matrix Hamiltonian differential equation (3.22) is such that for all $t \in [t_0,\, t_f]$,*

$$X(t)\,X(t_0)^{-1}x_0 \geq \bar{x}$$

*and*

$$-R^{-1}B^TY(t)\,X(t_0)^{-1}x_0 \geq \bar{u}$$

*for $x_0 \geq \bar{x}$.*

**Proof :** This result follows directly from Theorem 3.2.1 b) or equivalently from the fact that, for the $\boldsymbol{LQ^{t_f}}$ problem, we consider the minimization on a larger set.

□

## 3.5 State-invariant LQ problem

In this section the particular case of the state-invariant $\boldsymbol{LQ^{t_f}_{\bar{x}}}$ problem is studied. In this case, admissibility conditions can be obtained by using the standard $\boldsymbol{LQ^{t_f}}$ problem.

### 3.5.1 Problem statement and optimality conditions

The finite horizon state-invariant LQ problem, which is denoted by $\boldsymbol{LQ^{t_f}_{\bar{x}}}$ , consists of minimizing the quadratic functional (3.2) for a given linear system described by (3.1), where the initial state $x_0 \geq \bar{x}$ is fixed, under the constraint

$$\forall\, t \in [t_0,\, t_f],\;\; x(t) \geq \bar{x}. \tag{3.25}$$

This problem corresponds to the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ problem with $W = I_n$ and $Z = 0_m$ where $\lambda$ and $\Lambda$ are the multipliers associated to the state constraint (3.25) and $\bar{x} \ll 0$ (as in Chapter 1). The input constraints and the associated multipliers are no longer needed. As for the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ problem, optimality conditions can be established, using the maximum principle with state constraints (see e.g. [HSV95]). Then we obtain a result which is similar to Theorem 3.2.1 and is therefore omitted.

### 3.5.2 Optimality conditions via admissibility

As previously, conditions such that the $\boldsymbol{LQ^{t_f}_{\bar{x}}}$ problem has a solution can be obtained by using the standard $\boldsymbol{LQ^{t_f}}$ problem.

**Corollary 3.5.1 (Optimality conditions based on admissibility)** *The solution of the (standard) $\boldsymbol{LQ^{t_f}}$ problem is solution of the $\boldsymbol{LQ^{t_f}_{\bar{x}}}$ problem for $x_0 \geq \bar{x}$ if and only if the $\boldsymbol{LQ^{t_f}}$ -optimal state trajectories are admissible, i.e. $x(t) \geq \bar{x}$ for all $t \in [t_0,\, t_f]$, or equivalently one of the following equivalent conditions holds :*
*a) The standard closed-loop matrix $A + B\,K(t) = A - B\,R^{-1}B^T P(t)$, where $P(t)$ is the solution of the RDE, is a Metzler matrix for all $t \in [t_0,\, t_f]$, i.e.*

$$\forall\, i \neq j,\, \forall\, t \in [t_0,\, t_f],\; [B\,R^{-1}B^T P(t)]_{ij} \leq a_{ij}. \tag{3.26}$$

*and*

$$\forall\, t \in [t_0,\, t_f],\; \Phi_K(t,\, t_0)\,\bar{x} \geq \bar{x} \tag{3.27}$$

*where $\Phi_K(t,\, t_0)$ is the fundamental matrix of the closed-loop system (3.24), which satisfies the following homogeneous equation :*

$$\begin{cases} \frac{\partial}{\partial t}\Phi_K(t,\, t_0) = (A + B\,K(t))\,\Phi_K(t,\, t_0), \\ \Phi_K(t_0,\, t_0) = I_n \end{cases} \tag{3.28}$$

*b) The matrix solution of the standard matrix Hamiltonian differential equation (3.22) is such that for all $t \in [t_0,\, t_f]$, $X(t)\,X(t_0)^{-1}x_0 \geq \bar{x}$.*

**Proof :** This result follows directly from Theorem 1.1.2 and Theorem 3.2.1. In addition, the solution of the $\boldsymbol{LQ^{t_f}_{\bar{x}}}$ problem is given as in Theorem 3.2.1 where the multiplier functions $\upsilon(t)$ and $\Upsilon(t)$ are identically equal to zero.

□

## 3.6 Input-invariant LQ problem

### 3.6.1 Problem statement and optimality conditions

The finite horizon input-invariant LQ problem, which is denoted by $\boldsymbol{LQ^{t_f}_{\bar{u}}}$ , consists of minimizing the quadratic functional (3.2) for a given input-invariant linear system described by (3.1), under the constraints

$$\forall\, t \in [t_0,\, t_f],\;\; u(t) \geq \bar{u}. \tag{3.29}$$

This problem corresponds to the $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ problem with $W = 0_n$ and $Z = I_m$ where $\upsilon$ and $\Upsilon$ are the multipliers associated to the input constraint (3.29) and $\bar{u} \leq 0$. In this subsection, optimality conditions for the $\boldsymbol{LQ}_{\bar{u}}^{t_f}$ problem are established using the maximum principle with input constraints (see e.g. [HSV95]). This leads to a result which is similar to Theorem 3.2.1 and is therefore omitted.

### 3.6.2 Optimality conditions via admissibility

As in the previous section, conditions such that the $\boldsymbol{LQ}_{\bar{u}}^{t_f}$ problem has a solution can be obtained by using the standard $\boldsymbol{LQ}^{t_f}$ problem.

**Corollary 3.6.1 (Optimality conditions based on admissibility)** *The solution of the (standard) $\boldsymbol{LQ}^{t_f}$ problem with $m = n$ is solution of the $\boldsymbol{LQ}_{\bar{u}}^{t_f}$ problem if and only if the $\boldsymbol{LQ}^{t_f}$ -optimal control is admissible, i.e. $u(t) \geq \bar{u}$ for all $t \in [t_0,\, t_f]$, or equivalently one of the following conditions holds :*
*a) The state feedback $K(t) = -R^{-1}B^T P(t)$, where $P(t)$ is the solution of the RDE, is of full rank, such that for all $t \in [t_0,\, t_f]$,*

$$K(t)\,\Phi_K(t,\,t_0)\,K(t)^{-1} \geq 0$$

*and*

$$K(t)\,\Phi_K(t,\,t_0)\,K(t)^{-1}\,\bar{u} \geq \bar{u}$$

*where $\Phi_K(t,\,t_0)$ is the fundamental matrix of the closed-loop system (3.24), which satisfies (3.28).*
*b) The matrix solution of the standard matrix Hamiltonian differential equation (3.22) is such that for all $t \in [t_0,\, t_f]$, $\;-R^{-1}B^T Y(t)\,X(t_0)^{-1}x_0 \geq \bar{u}$.*

**Proof :** This result follows directly from Corollary 1.1.4 and Theorem 3.2.1.

□

**Remark 3.6.1** *If $B$ is a full rank matrix, condition a) of Corollary 3.6.1 implies, notably, the invertibility of $P(t)$ (and therefore also of $Y(t)$). The inverse of $P(t)$ is computable as the solution of the following Riccati differential equation, with $V(t) = P(t)^{-1}$ :*

$$-\dot{V}(t) = -V(t)\,A^T - A\,V(t) + B\,R^{-1}B^T - V(t)\,C^T C\,V(t), \;\; V(t_f) = S^{-1}.$$

*Furthermore, let us notice that in the case of the minimal energy control problem ($C = 0$), the inverse of $P(t)$ is the solution of a Lyapunov equation.*

# Chapter 4

# The Positive LQ Problem

In this chapter, the LQ problem is studied for positive systems. This problem corresponds to the $\boldsymbol{LQ}_{\bar{\boldsymbol{x}}}^{\boldsymbol{t_f}}$ problem with $\bar{x} = 0$ and $W = I_n$. The main objective of this problem is to keep the positivity property of the open loop system for the designed controlled system, which is meaningful from the modeling point of view. As we have already seen, many theoretical problems have been studied for positive systems. Here the finite-horizon positive linear quadratic problem is studied for positive linear time systems.

As in the previous chapter, optimality conditions are established, which are based on the maximum principle and on the admissibility of the solution of the standard $\boldsymbol{LQ}^{\boldsymbol{t_f}}$ problem, respectively. In addition, sufficient conditions for the positivity of the standard $\boldsymbol{LQ}^{\boldsymbol{t_f}}$ closed-loop system are stated in terms of the matrix solution of the Riccati differential equation. Moreover, the particular problem of minimal energy control with penalization of the final state is studied. Finally, numerical examples are given in order to illustrate these results.

## 4.1 Problem statement and optimality conditions

In the particular case of the LQ problem for positive systems, the same analysis as in Section 3.5 can be done by considering $\bar{x} = 0$, $W = I_n$. Indeed, the finite horizon positive LQ problem, which is denoted by $\boldsymbol{LQ}_{+}^{\boldsymbol{t_f}}$ , consists of minimizing the quadratic functional (3.2) for a given linear system $R = [A, B]$, where the initial state $x(t_0) = x_0 \geq 0$ is fixed, under the constraint

$$\forall\, t \in [t_0, t_f], \;\; x(t) \geq 0, \tag{4.1}$$

where $t_f$ is a fixed final time, $u$ is any piecewise-continuous $\mathbb{R}^m$-valued function, $R \in \mathbb{R}^{m\times m}$ is a symmetric positive definite matrix, $C \in \mathbb{R}^{p\times n}$ and $S \in \mathbb{R}^{n\times n}$ is a symmetric positive semidefinite matrix.

Therefore we can obtain optimality conditions as previously by using the maximum principle. Applying this principle with state constraints (see e.g. [HSV95]) yields a characterization, in vector and matrix forms for an $\boldsymbol{LQ}_{+}^{\boldsymbol{t_f}}$ -optimal control. Theorem 3.2.1 becomes, with $W = I_n$, $Z = 0_m$ and $\bar{x} = 0$ :

**Theorem 4.1.1 (Optimality conditions based on the maximum principle)**
*a) The* $\boldsymbol{LQ}_+^{\boldsymbol{t_f}}$ *problem has a solution* $u(\cdot)$ *if and only if there exists a piecewise continuous multiplier function* $\lambda(\cdot)$ *mapping* $[t_0,\, t_f]$ *into* $\mathbb{R}^n$ *such that* $u(t) = -R^{-1}B^T p(t),\ t \in [t_0,\, t_f]$*, where* $[x(t)^T \quad p(t)^T]^T \in \mathbb{R}^{2n}$ *is the solution of the Hamiltonian differential equation*

$$\begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = H \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \lambda(t) \end{bmatrix},\ t \in [t_0,\, t_f] \tag{4.2}$$

*with* $x(t_0) = x_0,\ p(t_f) = S\,x(t_f) - \lambda(t_f)$*, where*

$$H = \begin{bmatrix} A & -B\,R^{-1}B^T \\ -C^TC & -A^T \end{bmatrix}$$

*is the Hamiltonian matrix, and for all* $t \in [t_0,\, t_f]$*,* $x(t) \geq 0,\ \lambda(t) \geq 0$ *and* $\lambda(t)^T x(t) = 0$ *(complementarity condition).*
*b) By using the matrix form of the Hamiltonian differential equation (4.2), a piecewise-continuous control function* $u\ :\ [t_0,\, t_f] \to \mathbb{R}^m$ *is solution of the* $\boldsymbol{LQ}_+^{\boldsymbol{t_f}}$ *problem if and only if there exists a piecewise continuous multiplier matrix function* $\Lambda(\cdot)$ *mapping* $[t_0,\, t_f]$ *into* $\mathbb{R}^{n\times n}$ *such that* $u(t) = K(t,\, x_0)\,x(t) := -R^{-1}B^TY(t)\,X(t)^{-1}x(t),\ t \in [t_0,\, t_f]$*, where* $[X(t)^T \quad Y(t)^T]^T \in \mathbb{R}^{2n\times n}$ *is the solution of the matrix Hamiltonian differential equation*

$$\begin{bmatrix} \dot{X}(t) \\ \dot{Y}(t) \end{bmatrix} = H \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \Lambda(t) \end{bmatrix},\ t \in [t_0,\, t_f]$$

*with the final condition* $X(t_f) = I$ *and* $Y(t_f) = S - \Lambda(t_f)$*, and for all* $t \in [t_0,\, t_f]$*,*

$$\Lambda(t)\,X(t_0)^{-1}x_0 \geq 0, \tag{4.3}$$

$$x_0^T X(t_0)^{-T}\Lambda(t)^T X(t)\,X(t_0)^{-1}x_0 = 0 \quad \textit{(complementarity condition)} \tag{4.4}$$

*and*

$$X(t)\,X(t_0)^{-1}x_0 \geq 0. \tag{4.5}$$

**Remark 4.1.1** *A priori, in view of conditions (4.3)-(4.5), the function* $K(t,\, x_0)$ *in Theorem 4.1.1 (b) clearly depends upon the choice of the initial state* $x_0$*. Stronger conditions are needed in order to make it independent of the initial state, i.e. such that the optimal control law be of state feedback type* $u(t) = K(t)\,x(t)$*. Such conditions are stated next.*

**Proposition 4.1.2** *Conditions (4.3)-(4.5) are satisfied for all initial states* $x_0 \geq 0$ *if and only if the following conditions hold for all* $t \in [t_0,\, t_f]$ *:*

$$\Lambda(t)\,X(t_0)^{-1} \geq 0, \tag{4.6}$$

$$\Lambda(t)^T X(t) + X(t)^T \Lambda(t) = 0 \tag{4.7}$$

*and*

$$X(t)\,X(t_0)^{-1} \geq 0. \tag{4.8}$$

The proof of this result is based on the following lemma.

**Lemma 4.1.3** *A matrix $M \in \mathbb{R}^{n\times n}$ is a skew-symmetric matrix, i.e. $M = -M^T$, if and only if*

$$\text{for all } x \geq 0,\ x^T M\, x = 0. \tag{4.9}$$

**Proof :** First recall that a matrix $M$ is skew-symmetric if and only if, for all $x \in \mathbb{R}^n$, $x^T M\, x = 0$. Then the necessity of condition (4.9) is obvious. Conversely, observe that condition (4.9) implies that

$$\forall\, x,\, y \geq 0, \quad y^T M\, x + x^T M\, y = (x+y)^T M\,(x+y) - x^T M\, x - y^T M\, y = 0. \tag{4.10}$$

Then observe that any $x \in \mathbb{R}^n$ can be written as $x = x_+ - x_-$ where $x_+ := \max\{x,\, 0\} = \frac{1}{2}(|x| + x) \geq 0$ and $x_- := \max\{-x,\, 0\} = \frac{1}{2}(|x| - x) \geq 0$. By using this decomposition of $x$ in (4.9) and the identity (4.10), it follows that, for all $x \in \mathbb{R}^n$, $x^T M\, x = 0$.
$\square$

**Proof of Proposition 4.1.2 :** The fact that conditions (4.3) and (4.5) hold for all $x_0 \geq 0$ is obviously equivalent to conditions (4.6) and (4.8). By Lemma 4.1.3, condition (4.4) holds for all $x_0 \geq 0$ if and only if the matrix $X(t_0)^{-T}\Lambda(t)^T X(t)\, X(t_0)^{-1}$ is skew-symmetric, or equivalently $\Lambda(t)^T X(t)$ is a skew-symmetric matrix, i.e. (4.7) holds.
$\square$

**Remark 4.1.2** *a) Conditions (4.6)-(4.8) can be hard to check in general. However they obviously hold with $\Lambda(t) = 0$ in an important particular case. See Corollary 4.1.4 below.*
*b) The optimality conditions in Theorem 4.1.1 and Proposition 4.1.2 also hold for linear systems $R = [A,\, B]$ that are not positive. However the positivity assumption plays a crucial role for obtaining the criteria established in Section 4.2.*

In view of the analysis above, conditions such that the $\boldsymbol{LQ_+^{t_f}}$ problem has a solution can be obtained by using the standard $\boldsymbol{LQ^{t_f}}$ problem as in the previous chapter, see Sections 3.3 and 3.5 and especially Corollary 3.5.1 applied with $\bar{x} = 0$.

**Corollary 4.1.4 (Optimality conditions based on admissibility)** *The solution of the $\boldsymbol{LQ^{t_f}}$ problem is solution of the $\boldsymbol{LQ_+^{t_f}}$ problem for all $x_0 \geq 0$ if and only if the $\boldsymbol{LQ^{t_f}}$ -optimal state trajectories are admissible, i.e. nonnegative for all $t \in [t_0,\, t_f]$ and for all $x_0 \geq 0$, or equivalently one of the following equivalent conditions holds :*
*a) The standard closed-loop matrix $A + B\,K(t) = A - B\,R^{-1}B^T P(t)$, where $P(t)$ is the solution of the RDE, is a Metzler matrix for all $t \in [t_0,\, t_f]$, i.e.*

$$\forall\, i \neq j,\, \forall\, t \in [t_0,\, t_f],\ [B\,R^{-1}B^T P(t)]_{ij} \leq a_{ij}. \tag{4.11}$$

*b) The matrix solution of the matrix Hamiltonian differential equation (3.22) is such that for all $t \in [t_0,\, t_f],\ X(t)\, X(t_0)^{-1} \geq 0$.*

**Remark 4.1.3** *The analysis and results of this section are readily extendable to the case where the final state penalty term in the cost (3.2) is of the form $(x(t_f) - x_f)^T S\,(x(t_f) - x_f)$ where $x_f \in \mathbb{R}_+^n$ is a fixed reference state for the final state $x(t_f)$.*

## 4.2 Positivity criteria

In this section, the $\boldsymbol{LQ_+^{t_f}}$ problem is studied with an additional assumption, namely the positivity of the open-loop system. In other words, the $\boldsymbol{LQ_+^{t_f}}$ problem consists of minimizing a quadratic functional for a given positive system while requiring that the state trajectories be nonnegative for any fixed nonnegative initial state, whence the positivity property should be kept for the optimal state trajectories. It is important to observe that, in this framework, it is not required that the input function $u(t)$ be nonnegative. The system $R = [A,\, B]$ being assumed to be positive, the latter constraint, i.e. $\forall\, t \in [t_0,\, t_f],\ u(t) \geq 0$, is clearly stronger than the constraint (4.1). Hence, here the $\boldsymbol{LQ^{t_f}}$ problem is studied with the aim of finding conditions on the problem data such that the standard closed-loop system is positive, i.e. such that the conditions of Corollary 4.1.4 hold. This can be interpreted as solving an inverse $\boldsymbol{LQ_+^{t_f}}$ problem.

### 4.2.1 Upper bound for the solution of the RDE

First, sufficient conditions are established for the positivity of the $\boldsymbol{LQ^{t_f}}$ closed-loop system in terms of an upper bound of the solution $P(t)$ of the RDE. We use an approach similar to the one developed in [MPS90].

**Theorem 4.2.1** *Consider the $\boldsymbol{LQ^{t_f}}$ problem (3.1)-(3.2). If $B\,R^{-1}B^T \geq 0$, if the solution of the RDE is nonnegative, i.e.*

$$\forall\, t \in [t_0,\, t_f], \quad P(t) \geq 0, \tag{4.12}$$

*and if*

$$\forall\, i \neq j, \quad \forall\, t \in [t_0,\, t_f], \quad \left[B\,R^{-1}B^T F(t)\right]_{ij} \leq a_{ij}, \tag{4.13}$$

*where $F(t)$ is the solution of the matrix Lyapunov differential equation,*

$$\dot F(t) = -A^T F(t) - F(t)\,A - C^T C,\ F(t_f) = S, \tag{4.14}$$

*then the $\boldsymbol{LQ^{t_f}}$ closed-loop system is positive and therefore the solution of the $\boldsymbol{LQ^{t_f}}$ problem is solution of the $\boldsymbol{LQ_+^{t_f}}$ problem.*

**Proof :** Since condition (4.13) is equivalent to the fact that $A - B\,R^{-1}B^T F(t)$ is a Metzler matrix for all $t \in [t_0,\, t_f]$, and in view of Corollary 4.1.4, it suffices to show that, for all $t \in [t_0,\, t_f]$, $P(t) \leq F(t)$. Now, thanks to the assumption (4.12), the matrix $V(t) := P(t)\,B\,R^{-1}B^T P(t)$ is nonnegative for all $t \in [t_0,\, t_f]$. In addition, the derivative of $Z(t) := P(t) - F(t)$ is given by $\dot Z(t) = -A^T Z(t) - Z(t)\,A + V(t)$. Hence $Z(t) = -\displaystyle\int_t^{t_f} e^{A^T(\tau-t)} V(\tau) e^{A\,(\tau-t)}\,\mathrm{d}\tau$ is a nonpositive matrix for all $t \in [t_0,\, t_f]$.

$\square$

**Remarks 4.2.1** *a) If $(A,\, B)$ is stabilizable and $(C,\, A)$ is detectable, then for $t_f \to \infty$, the solution $P(t) = P(t,\, t_f,\, S)$ of the RDE tends to the unique stabilizing positive semidefinite solution $P_+$ of the corresponding algebraic Riccati equation (ARE),*

$$A^T P_+ + P_+\,A - P_+\,B\,R^{-1}B^T P_+ + C^T C = 0. \tag{4.15}$$

*Moreover by Proposition 6.1.1, if $A$ is a stable Metzler matrix and $A - B\,R^{-1}B^T P_+$ is a (stable) Metzler matrix, then $P_+ \geq 0$ whenever $Q = C^T C \geq 0$ ; hence if in addition $P_+ \gg 0$, condition (4.12) must hold for $t_f$ sufficiently large.*
*b) Condition (4.13) implies that the weighting matrices $R$, $C^T C$ and $S$ have to be chosen such that $(B\,R^{-1}B^T S)_{ij} \leq a_{ij}$ and $(B\,R^{-1}B^T F(t_0))_{ij} \leq a_{ij}$, for all $i \neq j$, where $F(t)$ is the solution of (4.14) which is given by $F(t) = e^{A^T(t_f - t)} S\, e^{A(t_f - t)} + \int_0^{t_f - t} e^{A^T s} C^T C e^{A\,s}\, ds$. Moreover, condition (4.12) implies that $S$ has to be a nonnegative matrix.*
*c) In view of the analysis above, assumption (4.12) can be replaced by a weaker one, viz. $P(t)\,B\,R^{-1}B^T P(t) \geq 0$ for all $t \in [t_0,\, t_f]$.*

**Example 4.2.1** *Consider an unstable positive system $R = [A,\, B]$ where*

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix},\ B = I_2 \tag{4.16}$$

*and the cost (3.2) where*

$$C = \rho \begin{bmatrix} 1 & 0 \end{bmatrix} \ \text{ and } \ R = r\, I_2. \tag{4.17}$$

For all results and figures presented here, unless otherwise stated, the initial time $t_0$ is equal to $0$, the final time $t_f$ is equal to $20$ and the sampling step is $0.5$. We used MATLAB with the solver `ode23s` notably to integrate the RDE (3.23) and the Lyapunov equation (4.14).
Let $\rho = 1$ and $r = 6$. The eigenvalues of $A$ are $-2$ and $0$, thus the matrix $A$ is obviously unstable. Computing $S$ as the stabilizing positive semidefinite solution $P_+$ of the ARE gives

$$S = \begin{bmatrix} 1.1372 & 0.7980 \\ 0.7980 & 0.7037 \end{bmatrix}. \tag{4.18}$$

Condition (4.12) is clearly satisfied since $P(t)$ is equal to $S$ for all $t \in [0,\, t_f]$ and $S \geq 0$. Moreover, condition (4.13) is also numerically verified. See Figure 4.1.

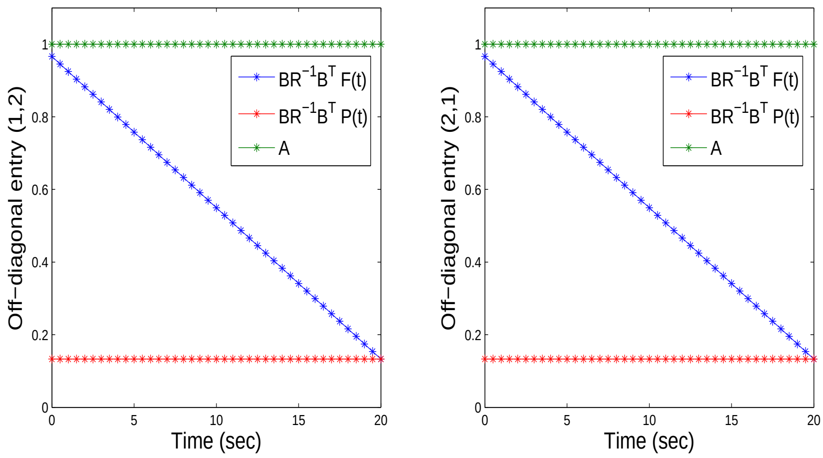



Figure 4.1: Off-diagonal entries of $B\,R^{-1}B^T F(t)$, $B\,R^{-1}B^T P(t)$ and $A$ for system (4.16)-(4.18).

The off-diagonal entries of $B\,R^{-1}B^T F(t)$ are clearly less than those of $A$. Now as one can expect, the closed-loop matrix $A + B\,K(t) = A - B\,R^{-1}B^T S = \begin{bmatrix} -1.18953 & 0.86700 \\ 0.86700 & -1.11728 \end{bmatrix}$ is clearly a Metzler matrix. See Figure 4.2 which represents the optimal state trajectories at the sampling times, for the initial states $x_0 = [1 \quad 0]^T$ (graphs on the left) and $x_0 = [0 \quad 1]^T$ (graphs on the right) respectively, i.e. the columns of $e^{A\,t}$ at the sampling times. One can numerically verify that the closed-loop system is positive.

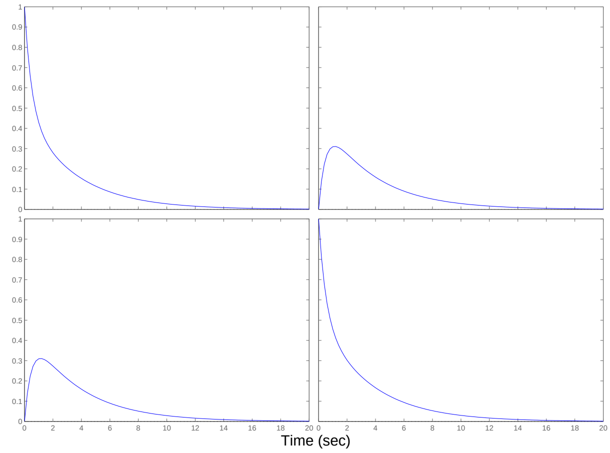



Figure 4.2: Optimal state trajectories $x(t)$ for system (4.16)-(4.18).

Notice that the closed-loop system is stable since the eigenvalues of the (constant) closed-loop matrix are $-2.0212$ and $-0.2857$. This observation is not really surprising, since the matrix $S$ was selected to be the unique stabilizing solution $P_+$ of ARE. Moreover, it could also be interesting to observe the behavior of the optimal control $u(t)$, which is represented in Figure 4.3 with the same initial state as above. We observe that $u(t) \le 0$ for all $t$ since $u(t) = -R^{-1}B^T P(t)\,x(t)$ with $P(t) \ge 0$ and $B \ge 0$. Actually, whenever $R^{-1} \ge 0$ the optimal control is always nonpositive.

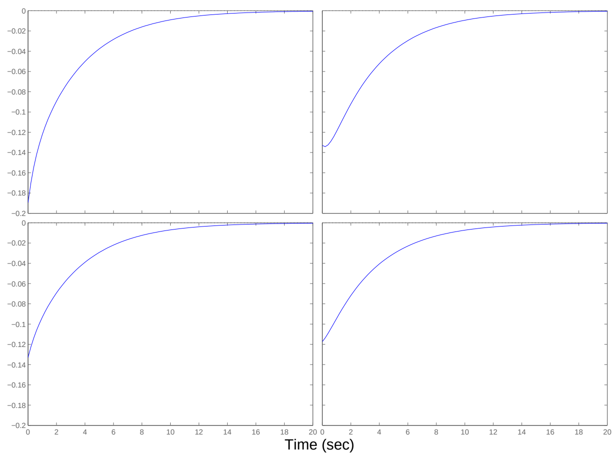



Figure 4.3: Optimal control $u(t)$ for system (4.16)-(4.18).

Furthermore, the matrix $R$ plays a paramount role here : for a fixed horizon $t_f$, condition (4.13) holds if $r$ is sufficiently large ; if the horizon $t_f$ is increased, $r$ has to be increased accordingly. See Figure 4.4 which compares the results for $r = 6$ and $r = 10$ where $t_f = 30$.

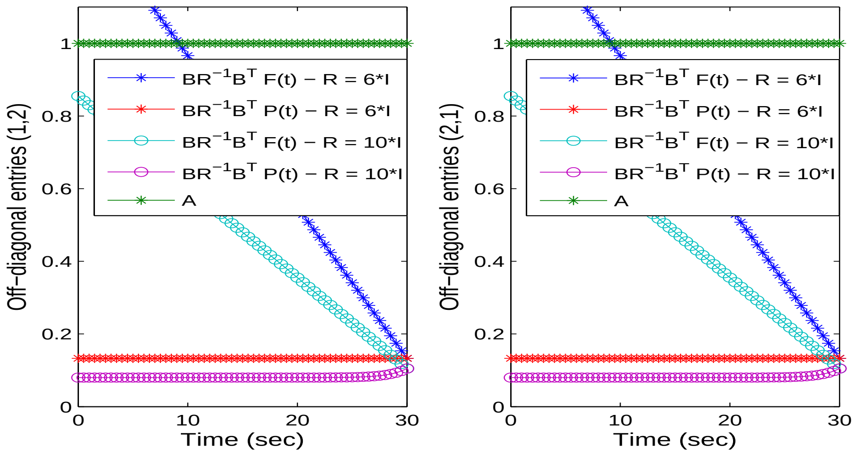



Figure 4.4: Off-diagonal entries of $B\,R^{-1}B^T F(t)$, $B\,R^{-1}B^T P(t)$ and $A$ for system (4.16)-(4.18) with $r = 6$ and $r = 10$ where $t_f = 30$.

**Remark 4.2.2** *The nonnegativity condition (4.12) in Theorem 4.2.1, appears to be a drawback of that result : it is indeed not clear how to check this condition without having to integrate the RDE. The following subsection is an attempt to avoid this assumption.*

### 4.2.2 Minimal energy control

In this subsection, the particular problem of minimal energy control with penalization of the final state is studied. An interesting feature of this approach is the fact that one can force the final state to be approximately close to zero by using a penalization term, for systems which are not necessarily reachable. A zero final state can be (exactly) reached by means of a minimal energy control (and therefore the state) trajectories nonnegative for reachable systems with a monomial gramian on a finite time interval, see [Ka 02, Subsection 3.4.2].

Here, sufficient conditions are established for the minimal energy control problem in terms of the spectral radius of the penalty matrix $S$. In the sequel, $\sigma(A)$ and $\rho(A)$ denote the spectrum and the spectral radius of a matrix $A$, respectively. The matrix norm that is used here is the one induced by the euclidean vector norm.

**Theorem 4.2.2** *Consider the minimal energy* $\boldsymbol{LQ^{t_f}}$ *problem (3.1)-(3.2), i.e. with* $C = 0$*. Let us denote* $\lambda_{\min}(R) := \min\{\lambda \ :\ \lambda \in \sigma(R)\}$*. Assume that* $a_{ij} > 0$ *for all* $i \neq j$*. If the spectral radius* $\rho(S)$ *of the final state penalty matrix is sufficiently small such that*

$$\rho(S) = \max_{\mu_i \in \sigma(S)} \mu_i \quad < \quad \gamma := \begin{cases} \dfrac{\lambda_{\min}(R)}{\alpha^2 \|B\|^2\, t_f}, & \textit{if } \lambda_F < 0 \\ \dfrac{\lambda_{\min}(R)}{\alpha^2 \|B\|^2\, t_f}\, e^{-2\lambda_F t_f}, & \textit{if } \lambda_F \geq 0, \end{cases} \tag{4.19}$$

*where $\lambda_F$ denotes the Frobenius eigenvalue (see Theorem A.1.5) and the constant $\alpha \geq 0$ is such that, for all $t \in [t_0,\, t_f]$, $\|e^{A\,t}\| \leq \alpha\, e^{\lambda_F\, t}$, then the $\boldsymbol{LQ^{t_f}}$ closed-loop system is positive and therefore the solution of the $\boldsymbol{LQ^{t_f}}$ problem is solution of the $\boldsymbol{LQ_+^{t_f}}$ problem.*

**Proof :** The positivity constraint on the closed-loop matrix can be written in terms of the solution $P(t)$ of the RDE (see condition (4.11)), where $B \geq 0$ since $[A,\, B]$ is a positive system. In addition, $P(t) = Y(t)\, X(t)^{-1} = e^{A^T(t_f-t)} S\, [I + G(t_f - t)\, S]^{-1}\, e^{A\,(t_f-t)}$, where $G(t_f - t) = \displaystyle\int_0^{t_f-t} e^{A\,\sigma} B\, R^{-1} B^T e^{A^T\sigma}\, \mathrm{d}\sigma$, and

$$\begin{aligned}
\|G(t_f - t)\, S\| &\leq \frac{\alpha^2\, \|B\|^2}{\lambda_{\min}(R)} \int_0^{t_f-t} e^{2\lambda_F \sigma}\, \mathrm{d}\sigma\ \rho(S) \\
&= \frac{\alpha^2\, \|B\|^2}{\lambda_{\min}(R)}\, e^{2\lambda_F\, \tau}\, (t_f - t)\, \rho(S) \qquad \text{(by the mean value theorem)} \\
&\leq \frac{\alpha^2 \|B\|^2 t_f}{\lambda_{\min}(R)}\, e^{2\lambda_F\, \tau}\, \rho(S)
\end{aligned}$$

for some $\tau \in [0,\, t_f - t]$ where $e^{2\lambda_F\, \tau}$ depends on the stability of $A$. If $A$ is stable, $\lambda_F$ is negative (see Theorem 2.2.1) and therefore $e^{2\lambda_F\, \tau} < 1$. If $A$ is unstable, $\lambda_F$ is nonnegative, then $e^{2\lambda_F\, \tau} < e^{2\lambda_F\, t_f}$. Thus, if (4.19) holds, then $\|G(t_f - t)\, S\| \leq \dfrac{\rho(S)}{\gamma} < 1$, whence $\|S\, [I + G(t_f - t)\, S]^{-1}\,\| \leq \dfrac{\rho(S)}{1 - \frac{\rho(S)}{\gamma}}$ (by applying Neumann's Lemma). Hence, by choosing $\rho(S)$ sufficiently small, condition (4.11) holds, since $\forall\, i \neq j$, $a_{ij} > 0$ and $e^{A'(t_f-t)}$ and $e^{A\,(t_f-t)}$ are bounded on $[t_0,\, t_f]$.

□

**Remarks 4.2.3** *a) The constant $\alpha$ can be interpreted as a condition number. Indeed, $\alpha$ can be chosen to be given by $\alpha = \kappa(V) = \|V\|\|V^{-1}\|$, where $V$ is the (generalized) eigenvector matrix of $A$. In addition, if $A$ is a symmetric matrix, one can choose $\alpha = 1$.*
*b) For a fixed final time $t_f$, if the entries of the penalization matrix $S$ of the final state are increased, one has to increase the entries of the control penalization matrix $R$ accordingly. On the other hand, condition (4.19) can also be written as*

$$\begin{cases} t_f &< E\, \dfrac{\lambda_{\min}(R)}{\lambda_{\max}(S)}, \quad \textit{if } \lambda_F < 0 \\ t_f\, e^{2\lambda_F\, t_f} &< E\, \dfrac{\lambda_{\min}(R)}{\lambda_{\max}(S)}, \quad \textit{if } \lambda_F \geq 0, \end{cases}$$

*where $E = \frac{1}{\alpha^2 \|B\|^2}$ is a constant depending only on the system data. Hence, if the time horizon $t_f$ is increased, the fraction $\frac{\lambda_{\min}(R)}{\lambda_{\max}(S)}$ has to be increased accordingly for condition (4.19) to be satisfied. This reveals a tradeoff between positivity and stability of the closed-loop system in a receding horizon approach, see e.g. [CW96]. The following example is an illustration of this tradeoff.*

**Example 4.2.2** *Consider the unstable positive system $R = [A,\ B]$ described by (4.16) and the cost (3.2) where*

$$C = 0_{1\times 2} \ \textit{and} \ R = I_2. \tag{4.20}$$

We compute the different parameters which play a role in condition (4.19). First observe that the right-hand side of this condition is equal to $0.05$, since $\lambda_F = 0$, $t_f = 20$ and $\alpha = 1$ and that condition (4.19) reads $\rho(S) < 1/t_f$. Therefore in order to guarantee this condition, the matrix $S$ has to be chosen such that its spectral radius is less than $0.05$. We choose the other eigenvalue less than $\rho(S)$ and we compute random associated eigenvectors. In this way we obtain the following values : $\rho(S) = 0.0495$, $\sigma(S) = \{0.0495,\ 0.0300\}$ and

$$S = \begin{bmatrix} 0.0397 & -0.0098 \\ -0.0098 & 0.0397 \end{bmatrix}. \tag{4.21}$$

We obtain numerically that the closed-loop system is positive, since the closed-loop matrix is a Metzler matrix for all sampling times : see Figure 4.5, representing the off-diagonal entries of $A + B\,K(t)$ at the sampling times and Figure 4.6, representing the optimal state trajectories for the initial states $x_0 = [1 \quad 0]^T$ (graphs on the left) and $x_0 = [0 \quad 1]^T$ (graphs on the right) respectively.

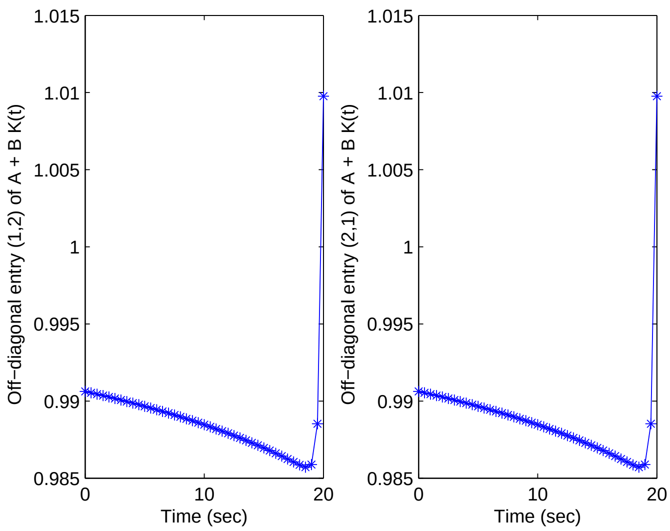



Figure 4.5: Off-diagonal entries of $A + B\,K(t)$ for system (4.16), (4.20)-(4.21).

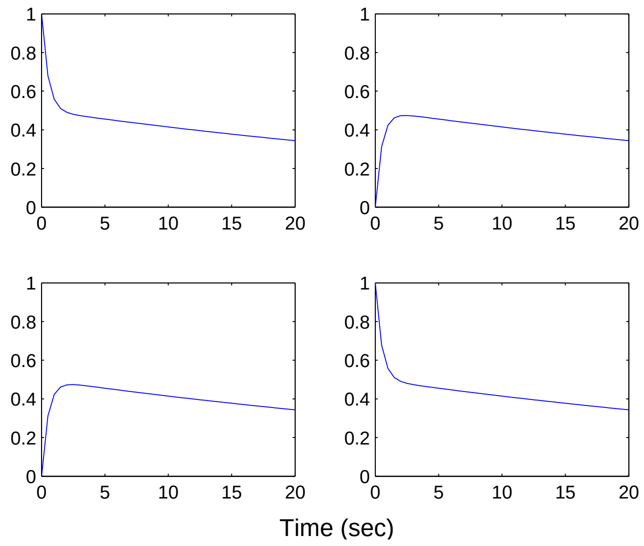



Figure 4.6: Optimal state trajectories $x(t)$ for system (4.16), (4.20)-(4.21).

Observe that one of the eigenvalues of the closed-loop matrix stays near zero, e.g. $\sigma(A + BK(20)) = \{-2,\ -0.0187\}$. Since condition (4.19) is given here by $\rho(S) < 1/t_f$, as long as $t_f$ increases, $S$ has to be decreased and therefore the final state $x(t_f)$ is less penalized (in the cost functional).

Now, as we have seen in Remark 4.2.3 b), the matrix $R$ also plays an important role. For a fixed final time $t_f = 20$, if $\lambda_{\min}(R)$ is increased, for example $\lambda_{\min}(R) = 10$, then $\rho(S)$ can be increased without violating condition (4.19). For example, the latter condition holds for $\rho(S) = 0.3276$ and $S = \begin{bmatrix} 0.2451 & -0.0825 \\ -0.0825 & 0.2451 \end{bmatrix}$. However, in this case, one of the eigenvalues of the closed-loop matrix is even closer to 0 than previously, e.g. $\sigma(A+BK(20)) = \{-2,\ -0.0123\}$. In addition, increasing the final time $t_f$ will emphasize this fact since the matrix $S$ should be modified such that condition (4.19) holds with $\rho(S)$ sufficiently small. For example, with $t_f = 100$, $\lambda_{\min}(R) = 10$, one can choose

$$\rho(S) = 0.0623,\ S = \begin{bmatrix} 0.0611 & -0.0012 \\ -0.0012 & 0.0611 \end{bmatrix} \text{ whence } \sigma(A + BK(20)) = \{-2,\ -0.0037\}.$$

So this example reveals a tradeoff between positivity and stability of the closed-loop system in a receding horizon approach. However, for $S$ given by (4.21), the solution of the infinite horizon problem is not the limit, when $t_f$ tends to infinity, of the solution of the finite horizon one, see e.g. [CW95] or [WC83]. Indeed, the last assumption (5.9) of Theorem 5.2.3 is not verified. Solving the $\boldsymbol{LQ}_+^{\text{inv}}$ problem (with techniques developed in Subsection 7.2.2) for system (4.16), (4.20)-(4.21) reveals that the closed-loop matrix is stable.

**Remark 4.2.4** *In Example 4.2.2, the solution $P(t)$ of the RDE is not nonnegative for all time, since $P(t_f) = S$ where $S$ is not nonnegative (see (4.21)). Hence Theorem 4.2.1 can not be applied in this case.*

**Remarks 4.2.5** *a) Obviously the Metzler property of a given matrix $A$ is kept under any diagonal perturbation. By considering the $\boldsymbol{LQ^{t_f}}$ problem (3.1)-(3.2) where the system $[A,\, B]$ is a positive system, if the solution $P(t)$ of the RDE is such that for all $t \in [t_0,\, t_f]$, $-B\, R^{-1}B^T P(t)$ is a diagonal matrix, then the $\boldsymbol{LQ^{t_f}}$ closed-loop system is positive and therefore the solution of the $\boldsymbol{LQ^{t_f}}$ problem is solution of the $\boldsymbol{LQ_+^{t_f}}$ problem.*

*b) One easy way to get this condition is to impose notably that $P(t)$ is a diagonal matrix, provided that the matrix $B\, R^{-1}B^T$ be also diagonal. Sufficient conditions for achieving this goal are stated as follows : consider the $\boldsymbol{LQ^{t_f}}$ problem (3.1)-(3.2) where the system $[A,\, B]$ is a positive system, with $B$ equal to $I_n$. Choose a constant $\alpha$ such that $\alpha > \max\{0,\, \lambda_F\}$. Define*

$$\mathcal{A}_\alpha := A\, A^T - (\alpha\, I_n + A)\, (\alpha\, I_n + A)^T.$$

*Assume that*

$$\begin{cases} S &= \alpha\, I_n, \\ R &= r\, I_m, \\ C^T C &= \alpha^2\, (\frac{1}{r} + 1)\, I_n + \mathcal{A}_\alpha, \end{cases}$$

*with $r > 0$ such that*

$$\frac{1}{r} > -\frac{\lambda_{\min}(\mathcal{A}_\alpha)}{\alpha^2} - 1, \tag{4.22}$$

*where $\lambda_{\min} := \min\{\lambda \;:\; \lambda \in \sigma(\mathcal{A}_\alpha)\}$.*
*Then for all $t \in [t_0,\, t_f]$, $P(t) = \alpha\, I_n$ and the $\boldsymbol{LQ^{t_f}}$ closed-loop system is positive and therefore the solution of the $\boldsymbol{LQ^{t_f}}$ problem is solution of the $\boldsymbol{LQ_+^{t_f}}$ problem.*
*Indeed, first observe that $R = R^T$ is positive definite since $r > 0$. Now $Q = C^T C$ is also positive definite because it can be shown that $\forall\, \mu \in \sigma(Q),\; \mu > 0$. Indeed, $\mu = \alpha^2\, (\frac{1}{r} + 1) + \lambda$ where $\lambda \in \sigma(\mathcal{A}_\alpha)$. Then condition (4.22) gives*

$$\alpha^2 (\frac{1}{r} + 1) > -\lambda_{\min}(\mathcal{A}_\alpha) \geq -\lambda, \quad \forall\, \lambda \in \sigma(\mathcal{A}_\alpha).$$

*Then $\forall\, \lambda \in \sigma(\mathcal{A}_\alpha),\; \mu = \alpha^2(\frac{1}{r} + 1) + \lambda > 0$ and $Q$ is positive definite. Hence $(Q,\, A)$ is observable and $(A,\, B)$ is controllable since $B = I_n$. Now the stabilizing positive semidefinite solution $P_+$ of the ARE, see (4.15), is given by $P_+ = \alpha\, I_n$. Indeed,*

$$\begin{aligned} & A^T P_+ + P_+\, A - P_+\, B\, R^{-1} B^T P_+ + C^T C \\ = \;& \alpha\, A^T + \alpha\, A - \frac{\alpha^2}{r}\, I_n + \alpha^2 (\frac{1}{r} + 1)\, I_n + \mathcal{A}_\alpha \\ = \;& \alpha\, A^T + \alpha\, A + \alpha^2\, I_n + A\, A^T - \alpha^2 I_n - \alpha\, A^T - \alpha\, A + A\, A^T \\ = \;& 0 \end{aligned}$$

*Then the solution $P(t)$ of the RDE such that $P(t_f) = S := \alpha\, I_n$ is constant and is given by $P(t) = \alpha\, I_n$. Hence the matrix $-B\, R^{-1} B^T P(t)$ is diagonal and is given by $-\frac{\alpha}{r}\, I_n$.* □

*c) In the previous remark, the control matrix $B$ is assumed to be the identity matrix. This assumption is verified for compartmental systems, which are an important subclass of positive systems, see e.g. [Van98, p. 593] and Subsection 2.3.3. Moreover, a stable positive system is* *equivalent to a compartmental system, modulo a positive diagonal transformation matrix, see e.g. [BF02] ; hence, in the statement of the $\boldsymbol{LQ^{t_f}}$ problem, a stable matrix $A$ can be assumed without loss of generality to be a compartmental matrix.*

**Example 4.2.3** *Consider the unstable positive system $R = [A,\ B]$ described by (4.16) and the cost (3.2). Select the following parameters $\alpha = 0.5$, $r = 2$, such that*

$$Q = C^T C = \begin{bmatrix} 1.1250 & -1.0000 \\ -1.0000 & 1.1250 \end{bmatrix} ;\ R = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} ;\ S = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}. \tag{4.23}$$

Then $P(t) = S$ for all $t$ and $A + BK(t)$ is a constant Metzler matrix, given by :

$$A + B\,K(t) = \begin{bmatrix} -6.2500 & 1.0000 \\ 1.0000 & -6.2500 \end{bmatrix}.$$

As one can expect, we can numerically verify in Figure 4.7 that the closed-loop system is positive and also stable. In this figure, the optimal state trajectories are depicted for several values of $\alpha$. Observe that the stability of the closed-loop system is improved by increasing the parameter $\alpha$.

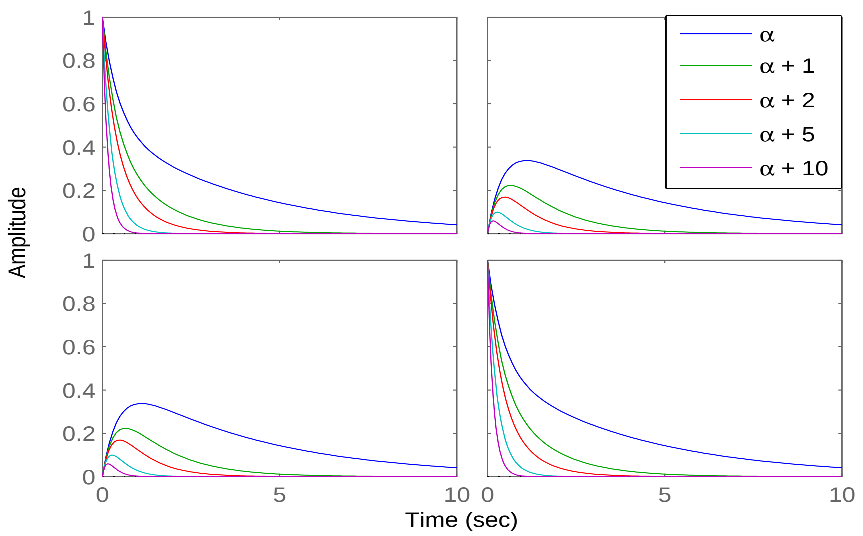



Figure 4.7: Optimal state trajectories $x(t)$ for system (4.16),(4.23).

## 4.3 Numerical examples

This section is devoted to numerical examples. First of all, the solution of the standard $\boldsymbol{LQ^{t_f}}$ problem is studied. The sufficient conditions of Theorem 4.2.2 are checked numerically in order to obtain an admissible solution for the positive LQ problem. Then the analytical solution of the standard $\boldsymbol{LQ^{t_f}}$ problem is computed. Finally the positive $\boldsymbol{LQ_+^{t_f}}$ problem is solved, by applying the optimality conditions of Theorem 4.1.1.

### 4.3.1 Standard LQ problem

Consider the unstable positive system described by

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \tag{4.24}$$

and the minimization of the functional

$$J(x_0,\, u,\, t_f) = \frac{1}{2}\left(\int_0^{t_f} \|u(t)\|^2\,\mathrm{dt} + a\,x_1^2(t_f) + b\,x_2^2(t_f)\right) \tag{4.25}$$

where

$$C = 0_{1\times 2},\ R = I_1,\ S = \operatorname{diag}(a,\, b).$$

First we compute the different parameters of condition (4.19) such that this condition reads $\rho(S) < \dfrac{\lambda_{\min}(R)}{\alpha^2\|B\|^2\,t_f}\,e^{-2\lambda_F t_f} = \dfrac{e^{-2\,t_f}}{t_f} = 0.20612\ 10^{-9}$ with $t_f = 10$. Then we choose

$$S = \operatorname{diag}(a,\, b) = 10^{-9}\begin{bmatrix} 0.1063 & 0 \\ 0 & 0.0736 \end{bmatrix} \tag{4.26}$$

with $\rho(S) = 0.1063\ 10^{-9}$. We obtain numerically that the closed-loop system is positive, since the closed-loop matrix is a Metzler matrix for all sampling times : see Figure 4.8, which shows the off-diagonal entries of $A + B\,K(t)$ at the sampling times. As already mentioned in subsection 4.2.2, the matrix $S$ should be chosen sufficiently small to guarantee the positivity at the expense of the final state penalization.

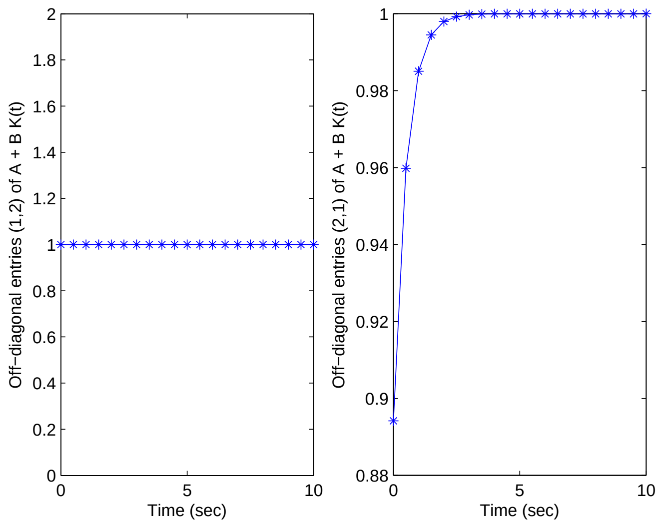



Figure 4.8: Off-diagonal entries of $A + B\,K(t)$ for system (4.24)-(4.26).

Now for a fixed $S = I_2$, the analytical solution of the $\boldsymbol{LQ^{t_f}}$ problem is computed by solving the Hamiltonian differential equation (3.22). The analytical expressions of the optimal control and the state trajectories are given, for all $t \in [0,\, t_f]$ and for $x_0 = [1 \quad 0]^T$, by

$$\begin{cases} u(t) &= -2\,e^{-t} \\ x_1(t) &= e^{-t} + t\,e^{-t} \\ x_2(t) &= -t\,e^{-t} \end{cases}$$

They are drawn in Figures 4.9 and 4.10. One can observe that $x_2(t)$ is nonpositive for all time. Now the analytical expressions of the optimal control and the state trajectories for $x_0 = [0 \quad 1]^T$ are given, for all $t \in [0,\, t_f]$, by

$$\begin{cases} u(t) &= -2\,e^{-t} \\ x_1(t) &= e^{-t} + t\,e^{-t} \\ x_2(t) &= -t\,e^{-t} \end{cases}$$

They are drawn in Figures 4.11 and 4.12. One can observe that here $x_2(t)$ is also not nonnegative for all time. So an additional nonnegativity constraint on the state trajectories is needed. The following subsection is devoted to the positive $\boldsymbol{LQ_+^{t_f}}$ problem in order to obtain nonnegative optimal state trajectories.

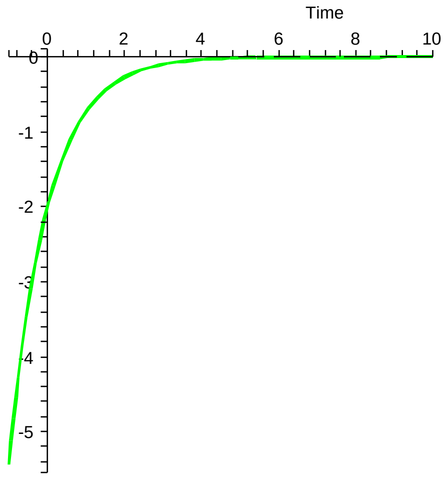



Figure 4.9: Optimal control for system (4.24) for $x_0 = [1 \quad 0]^T$.

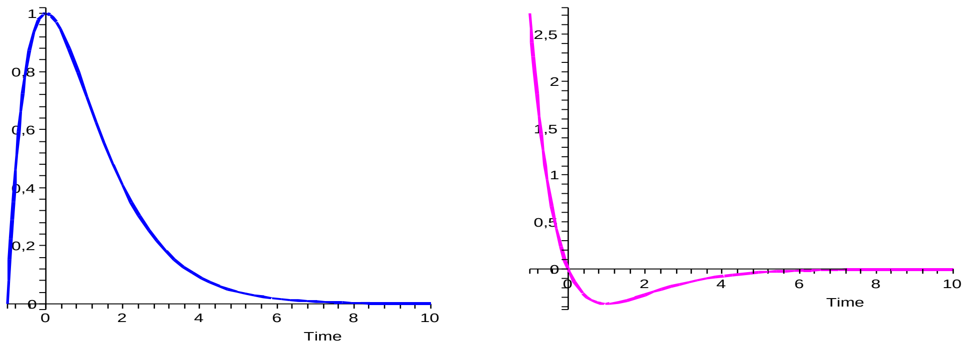



Figure 4.10: State trajectories for system (4.24) for $x_0 = [1 \quad 0]^T$.

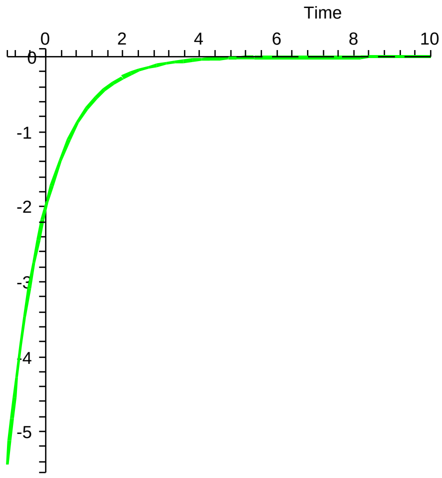



Figure 4.11: Optimal control for system (4.24) for $x_0 = [0 \quad 1]^T$.

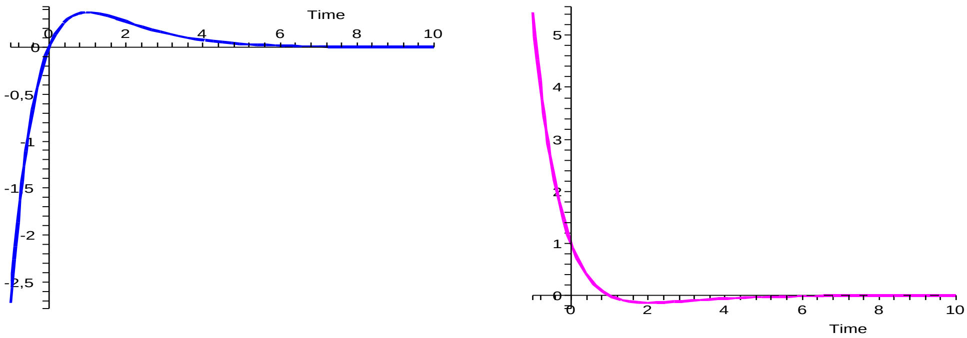



Figure 4.12: State trajectories for system (4.24) for $x_0 = [0 \quad 1]^T$.

## 4.3.2 Positive LQ problem

In this subsection, the $\boldsymbol{LQ_+^{t_f}}$ problem is solved, by applying the optimality conditions of Theorem 4.1.1, for which we first compute the analytical solution and then the numerical solution. Consider the positive system described by (4.24) and the minimization of the functional (4.25) under the constraints $\forall t \in [0,\, t_f],\ x_1(t) \geq 0$ and $x_2(t) \geq 0$. By applying the optimality conditions of Theorem 4.1.1, we obtain the following two-point boundary value problem, as found in the proof of Theorem 3.2.1 a) (see (3.13)) adapted with $W = I_n$, $Z = 0_m$ and $\bar{x} = 0$) :

$$\begin{cases} \dot{x}(t) &= A\,x(t) + B\,u(t) \\ \dot{p}(t) &= -A^T p(t) - C^T C\,x(t) + \lambda(t) \end{cases}$$

$$\text{with} \quad \begin{cases} x(t_0) &= x_0 \\ p(t_f) &= S\,x(t_f) - \lambda(t_f) \end{cases}$$

$$\text{where} \quad u(t) = -R^{-1} B^T p(t)$$

under the constraints

$$\begin{cases} x(t) &\geq 0 \\ \lambda(t)^T x(t) &= 0 \\ \lambda(t) &\geq 0 \end{cases}$$

or equivalently, with

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ C = 0_{1\times 2},\ R = I_1,\ S = I_2,$$

$$\begin{cases} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= x_1(t) + u(t) \end{cases}$$

and

$$\begin{cases} \dot{p}_1(t) &= -p_2(t) + \lambda_1(t) \\ \dot{p}_2(t) &= -p_1(t) + \lambda_2(t) \end{cases}$$

where $\lambda_1,\ \lambda_2 \geq 0$ and $\lambda_i^T x_i = 0,\ i = 1, 2,$ on $[0,\ t_f]$, with the following boundary conditions :

$$\begin{cases} x_1(0) &= x_{01} \\ x_2(0) &= x_{02} \end{cases} \quad \text{and} \quad \begin{cases} p_1(t_f) &= x_1(t_f) - \lambda_1(t_f) \\ p_2(t_f) &= x_2(t_f) - \lambda_2(t_f). \end{cases} \tag{4.27}$$

Moreover, the optimal control is given by $u(t) = -B\,R^{-1} B^T p(t) = -p_2(t),\ t \in [0,\ t_f]$.

The analytical expressions can be computed by solving this boundary value problem. There exist two similar ways to compute the state and adjoint state trajectories : let us consider two arbitrary times $t_1$ and $t_2$ in the interval $[0,\ t_f]$, then

$$\begin{cases} x(t_2) &= e^{A\,(t_2 - t_1)} x(t_1) + \displaystyle\int_{t_1}^{t_2} e^{A(t_2-\sigma)} B\,u(\sigma)\,\mathrm{d}\sigma \\ p(t_2) &= e^{-A^T\,(t_2-t_1)} p(t_1) + \displaystyle\int_{t_1}^{t_2} e^{-A^T(t_2-\sigma)} \lambda(\sigma)\,\mathrm{d}\sigma \end{cases} \tag{4.28}$$

where $e^{A\,t} = \begin{bmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{bmatrix}$ for $t \in [0,\ t_f]$ ;
or equivalently,

$$\begin{bmatrix} x(t_2) \\ p(t_2) \end{bmatrix} = e^{H\,(t_2-t_1)} \begin{bmatrix} x(t_1) \\ p(t_1) \end{bmatrix} + \int_{t_1}^{t_2} e^{H\,(t_2-\sigma)} \begin{bmatrix} 0 \\ \lambda(\sigma) \end{bmatrix} \mathrm{d}\sigma \tag{4.29}$$

where $e^{H\,t} = \begin{bmatrix} \cosh(t) & \sinh(t) & \dfrac{1}{2}\,(t\,\cosh(t) - \sinh(t)) & -\dfrac{1}{2}\,t\,\sinh(t) \\ \sinh(t) & \cosh(t) & \dfrac{1}{2}\,t\,\sinh(t) & -\dfrac{1}{2}\,(t\,\cosh(t) + \sinh(t)) \\ 0 & 0 & \cosh(t) & -\sinh(t) \\ 0 & 0 & -\sinh(t) & \cosh(t) \end{bmatrix}$

for $t \in [0,\, t_f]$.

Now let us fix the initial state to be $x_0 = [1 \quad 0]^T$. The numerical simulations given below (see Figures 4.16 and 4.17) will serve as a guide for further analytical calculations. In these figures, one can observe that there is one switching time $\tau$ after which the behavior of the state trajectories and the multipliers changes. Therefore, in the sequel, we analyze in details two intervals $[0,\, \tau[$ and $[\tau,\, t_f]$.

1. <u>For $t \in [0,\, \tau]$</u>

   - $x(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ;
   - $u(t) = -1 = -p_2(t)$, i.e. $p_2(t) = 1$ ;
   - $\lambda_1(t) = 0$ ;
   - $\dot{p}_1(t) = -p_2(t) + \lambda_1(t) = -1$ such that $p_1(t) = -t + p_1(0)$ by integration ;
   - $\dot{p}_2(t) = -p_1(t) + \lambda_2(t) = 0$ such that $\lambda_2(t) = p_1(t) = -t + p_1(0)$.

   It follows that

   $x(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ; $p(t) = \begin{bmatrix} -t + p_1(0) \\ 1 \end{bmatrix}$ ; $\lambda(t) = \begin{bmatrix} 0 \\ -t + p_1(0) \end{bmatrix}$ ; $u(t) = -1$ for $t \in [0,\, \tau[$

2. <u>For $t \in ]\tau,\, t_f]$</u>

   - $\lambda(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
   - In order to compute the expressions of $x(t)$ and $p(t)$, one can similarly use the identity (4.28) or (4.29) applied with $t_2 = t$ and $t_1 = \tau$ and where $x(\tau)$, $p(\tau)$ and $\lambda(\tau)$ are given as in the previous case. These two different ways of computing give the following analytical expressions :

$$\begin{aligned} x(t) &= \begin{bmatrix} \frac{1}{2}\,(p_1(0) - \tau)\,((t-\tau)\,\cosh(t-\tau) - \sinh(t-\tau)) \\ \quad + \cosh(t-\tau) - \frac{1}{2}\,(t-\tau)\,\sinh(t-\tau) \\ \\ \frac{1}{2}\,(p_1(0) - \tau)\,(t-\tau)\,\sinh(t-\tau) \\ \quad + \sinh(t-\tau) - \frac{1}{2}\,((t-\tau)\,\cosh(t-\tau) + \sinh(t-\tau)) \end{bmatrix} \\ p(t) &= \begin{bmatrix} (p_1(0) - \tau)\,\cosh(t-\tau) - \sinh(t-\tau) \\ \\ -(p_1(0) - \tau)\,\sinh(t-\tau) + \cosh(t-\tau) \end{bmatrix} \end{aligned}$$

$$
\begin{array}{rrcl}
 & \lambda(t) & = & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\text{and} & u(t) & = & \sinh(t-\tau)\,(p_1(0)-\tau) - \cosh(t-\tau)
\end{array}
$$

We can now summarize the analytical computations as follows. The optimal control is given by

$$
u(t) = \begin{cases} -1 & \text{if } t \in [0,\,\tau[ \\ \sinh(t-\tau)\,(p_1(0)-\tau) - \cosh(t-\tau) & \text{if } t \in [\tau,\,t_f] \end{cases}
$$

and the state trajectories and the adjoint states are given by

$$
x(t) = \begin{cases} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \text{if } t \in [0,\,\tau[ \\ \begin{bmatrix} \frac{1}{2}(p_1(0)-\tau)\,((t-\tau)\cosh(t-\tau) - \sinh(t-\tau)) \\ \quad + \cosh(t-\tau) - \frac{1}{2}(t-\tau)\sinh(t-\tau) \\ \\ \frac{1}{2}(p_1(0)-\tau)\,(t-\tau)\sinh(t-\tau) \\ \quad + \sinh(t-\tau) - \frac{1}{2}((t-\tau)\cosh(t-\tau) + \sinh(t-\tau)) \end{bmatrix} & \text{if } t \in [\tau,\,t_f] \end{cases}
$$

$$
p(t) = \begin{cases} \begin{bmatrix} p_1(0) - t \\ 1 \end{bmatrix} & \text{if } t \in [0,\,\tau[ \\ \begin{bmatrix} (p_1(0)-\tau)\cosh(t-\tau) - \sinh(t-\tau) \\ \\ -(p_1(0)-\tau)\sinh(t-\tau) + \cosh(t-\tau) \end{bmatrix} & \text{if } t \in [\tau,\,t_f] \end{cases}
$$

The multipliers associated to the nonnegativity constraints are given by

$$
\lambda(t) = \begin{cases} \begin{bmatrix} 0 \\ p_1(0) - t \end{bmatrix} & \text{if } t \in [0,\,\tau[ \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{if } t \in [\tau,\,t_f] \end{cases}
$$

These functions depend upon several parameters. The parameters $\tau$ and $p_1(0)$ are obtained to ensure the final condition (4.27) of the adjoint state, which is given, with $\lambda(t_f) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, by

$$
p_1(t_f) = x_1(t_f) \tag{4.30}
$$

$$
p_2(t_f) = x_2(t_f) \tag{4.31}
$$

By using the expressions of $x(t)$ and $p(t)$ with $t = t_f$, equation (4.30) reads :

$$
\begin{array}{rrl}
 & & (p_1(0)-\tau)\cosh(t_f-\tau) - \sinh(t_f-\tau) \\
 & = & \frac{1}{2}(p_1(0)-\tau)\,((t_f-\tau)\cosh(t_f-\tau) - \sinh(t_f-\tau)) \\
 & & + \cosh(t_f-\tau) - \frac{1}{2}(t_f-\tau)\sinh(t_f-\tau) \\
 & & \\
\Leftrightarrow & & p_1(0)\,(\cosh(t_f-\tau) - \frac{1}{2}(t_f-\tau)\cosh(t_f-\tau) + \frac{1}{2}\sinh(t_f-\tau)) \\
 & = & \tau\cosh(t_f-\tau) + \sinh(t_f-\tau) - \frac{1}{2}\tau\,(t_f-\tau)\cosh(t_f-\tau) \\
 & & + \frac{1}{2}\tau\sinh(t_f-\tau) + \cosh(t_f-\tau) - \frac{1}{2}(t_f-\tau)\sinh(t_f-\tau)
\end{array}
$$

Isolating $p_1(0)$ in this equation gives :

$$p_1(0) = \frac{2\cosh(t_f-\tau) - 8\tau\,\cosh(t_f-\tau) + \tau^2\,\cosh(t_f-\tau) - 8\sinh(t_f-\tau) + 2\tau\,\sinh(t_f-\tau)}{8\cosh(t_f-\tau) + \tau\,\cosh(t_f-\tau) + \sinh(t_f-\tau)} \tag{4.32}$$

Substituting $p_1(0)$ in (4.31), with $t_f = 10$, yields the following value of the switching time $\tau = 9.378729150$ by using MAPLE for solving the intersection of two curves ($p_2(t_f) = x_2(t_f)$). Finally, replacing $\tau$ in (4.32) gives $p_1(0) = 10.80878883$. By using these analytical forms, the control, state, multiplier and adjoint state trajectories are drawn in Figures 4.13, 4.14 and 4.15.

Now the numerical solution of this problem is computed by using `Matlab` and the function `quadprog`. This function uses an active set method which is also a projection method, similar to the one described in [Bix92]. First, the continuous time problem is converted into a discrete time one by sampling : for $i = 0, \ldots, N-1$, with $t_f = N\,h$, $u(t) = u(i\,h) =: u_i$, where $h$ is the sampling time. The resulting discrete time system is given by :

$$x_{i+1} = \begin{bmatrix} \cosh(h) & \sinh(h) \\ \sinh(h) & \cosh(h) \end{bmatrix} x_i + \begin{bmatrix} \cosh(h) - 1 \\ \sinh(h) \end{bmatrix} u_i, \quad i = 0, \ldots, N-1 \tag{4.33}$$

with the following discrete time cost $\dfrac{1}{2}\displaystyle\sum_{i=0}^{N-1} h\,\|u_i\|^2 + x_N^T S\,x_N$, (see Appendix C for details on discretization). Consider the final time $t_f = 10$ with the sampling time $h = 0.05$ and $a = b = 1$. The optimization algorithm mentioned above leads to the optimal control depicted in Figure 4.16. The corresponding state $x_i(t)$ and multiplier $\lambda_i(t)$ trajectories are depicted in Figure 4.17.

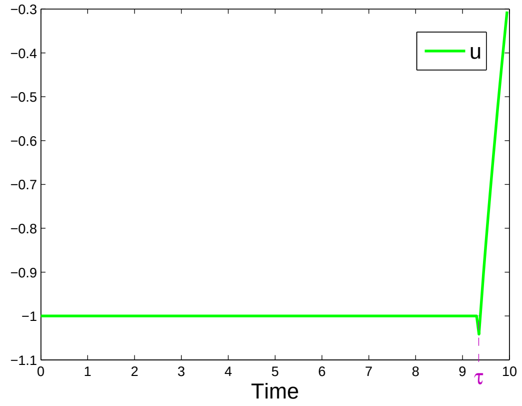



Figure 4.13: Optimal control for system (4.24).

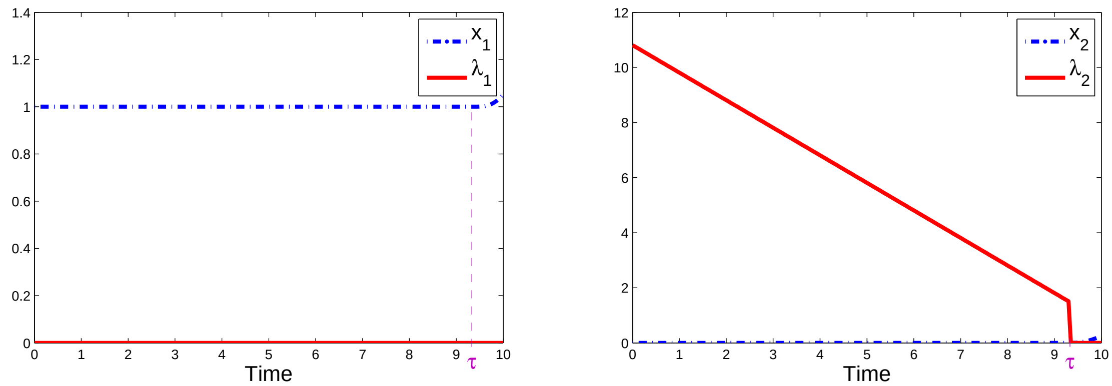



Figure 4.14: State trajectories and associated multipliers for system (4.24).

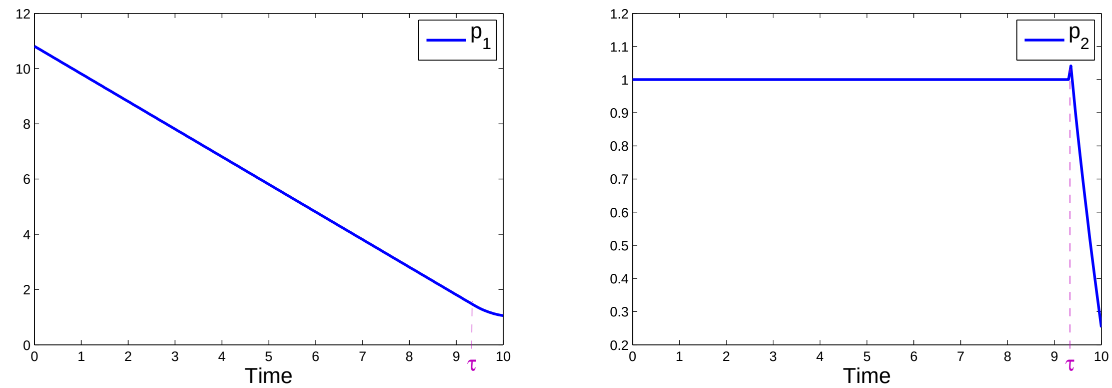



Figure 4.15: Adjoint states for system (4.24).

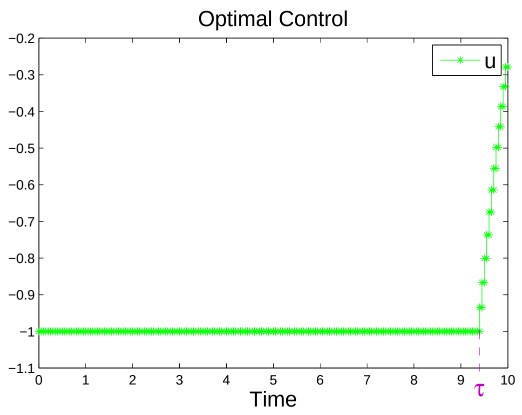



Figure 4.16: Optimal control for sampled data system (4.33).

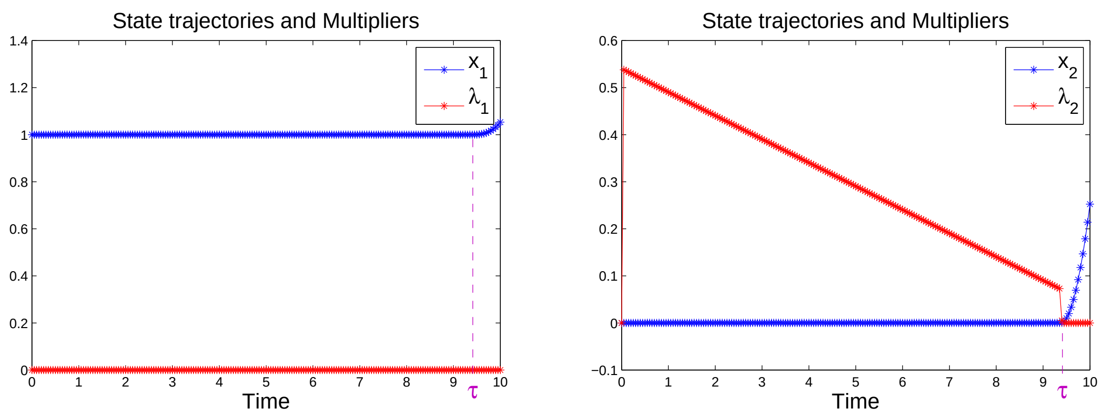



Figure 4.17: State trajectories and associated multipliers for sampled data system (4.33).

In these figures, one can observe that the constraints are always satisfied, as well as the complementarity conditions : whenever one of the state trajectories is strictly positive, the corresponding multiplier is equal to zero. For example, the multiplier $\lambda_1(t)$ is equal to zero as long as $x_1(t)$ is strictly positive.

Now fix the initial state to be $x_0 = [0 \quad 1]^T$ and perform the same analysis. The analytical expressions of the optimal control is given by

$$u(t) = \begin{cases} p_1(0)\,\sinh(t) + u_0\,\cosh(t) & \text{if } t \in [0,\, \tau_1[ \\ u_1 & \text{if } t \in [\tau_1,\, \tau_2[ \\ p_1(\tau_2)\,\sinh(t-\tau_2) + u_1\,\cosh(t-\tau_2) & \text{if } t \in [\tau_2,\, t_f] \end{cases}$$

and the state trajectories and the adjoint states are given by

$$x(t) = \begin{cases} \begin{bmatrix} \frac{1}{2}\,p_1(0)\,(t\,\cosh(t) - \sinh(t)) + \sinh(t) + \frac{1}{2}\,u_0\,t\,\sinh(t) \\ \frac{1}{2}\,p_1(0)\,t\,\sinh(t) + \cosh(t) + \frac{1}{2}\,u_0(t\,\cosh(t) + \sinh(t)) \end{bmatrix} & \text{if } t \in [0,\,\tau_1[ \\ \begin{bmatrix} -u_1 \\ 0 \end{bmatrix} & \text{if } t \in [\tau_1,\,\tau_2[ \\ \begin{bmatrix} \frac{1}{2}\,p_1(\tau_2)\,((t-\tau_2)\,\cosh(t-\tau_2) - \sinh(t-\tau_2)) \\ \quad -u_1\,\cosh(t-\tau_2) + \frac{1}{2}\,u_1\,(t-\tau_2)\,\sinh(t-\tau_2) \\ \\ \frac{1}{2}\,p_1(\tau_2)\,(t-\tau_2)\,\sinh(t-\tau_2) \\ \quad -u_1\,\sinh(t-\tau_2) + \frac{1}{2}\,u_1((t-\tau_2)\,\cosh(t-\tau_2) + \sinh(t-\tau_2)) \end{bmatrix} & \text{if } t \in [\tau_2,\,t_f] \end{cases}$$

$$p(t) = \begin{cases} \begin{bmatrix} p_1(0)\,\cosh(t) + u_0\,\sinh(t) \\ -p_1(0)\,\sinh(t) - \cosh(t)\,u_0 \end{bmatrix} & \text{if } t \in [0,\,\tau_1[ \\ \begin{bmatrix} p_1(\tau_2) + u_1\,(t-\tau_2) \\ -u_1 \end{bmatrix} & \text{if } t \in [\tau_1,\,\tau_2[ \\ \begin{bmatrix} p_1(\tau_2)\,\cosh(t-\tau_2) + u_1\,\sinh(t-\tau_2) \\ -p_1(\tau_2)\,\sinh(t-\tau_2) - u_1\,\cosh(t-\tau_2) \end{bmatrix} & \text{if } t \in [\tau_2,\,t_f] \end{cases}$$

The multipliers associated to the nonnegativity constraints are given by

$$\lambda(t) = \begin{cases} \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{if } t \in [0,\,\tau_1[ \\ \begin{bmatrix} 0 \\ p_1(\tau_2) + u_1\,(t-\tau_2) \end{bmatrix} & \text{if } t \in [\tau_1,\,\tau_2[ \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{if } t \in [\tau_2,\,t_f] \end{cases}$$

where the constants $u_0$ and $u_1$ are given by $u_0 = -2.6386$ and $u_1 = -0.26040$. These functions depend upon several parameters. The first parameters $\tau_2$ and $p_1(\tau_2)$ are obtained to ensure the final condition (4.27) of the adjoint state, which is given, with $\lambda(t_f) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, by

$$\begin{cases} p_1(t_f) & = & x_1(t_f) \\ p_2(t_f) & = & x_2(t_f) \end{cases} \tag{4.34}$$

By using the expressions of $x(t)$ and $p(t)$ given previously with $t = t_f$, system (4.34) becomes :

$$\begin{cases} & p_1(\tau_2)\,\cosh(t_f-\tau_2) + u_1\,\sinh(t_f-\tau_2) \\ = & \frac{1}{2}\,p_1(\tau_2)\,((t-\tau_2)\,\cosh(t-\tau_2) - \sinh(t-\tau_2)) - u_1\,\cosh(t-\tau_2) + \frac{1}{2}\,u_1\,(t-\tau_2)\,\sinh(t-\tau_2) \\ \\ & -p_1(\tau_2)\,\sinh(t-\tau_2) - u_1\,\cosh(t-\tau_2) \\ = & \frac{1}{2}\,p_1(\tau_2)\,(t-\tau_2)\,\sinh(t-\tau_2) - u_1\,\sinh(t-\tau_2) + \frac{1}{2}\,u_1((t-\tau_2)\,\cosh(t-\tau_2) + \sinh(t-\tau_2)) \end{cases}$$

with $u_0 = -2.6386$ and $u_1 = -0.26040$. That is a system of two equations with two parameters $p_1(\tau_2)$ and $\tau_2$. These equations are linear in $p_1(\tau_2)$. Solving the system by using MAPLE, gives $\tau_2 = 9.378729146$ and $p_1(\tau_2) = 0.3723875346$.

Now the values of $\tau_1$ and $p_1(0)$ are obtained in order to ensure the continuity of the state (and adjoint state) trajectories : the expressions of $x(t)$ and $p(t)$ on the intervals $[0,\ \tau_1]$ and $[\tau_1,\ \tau_2]$ should coincide at $t = \tau_1$. By using the expressions given in the previous subsection for $x(t)$, one gets :

$$\begin{bmatrix} \frac{1}{2}\, p_1(0)\, (\tau_1\, \cosh(\tau_1) - \sinh(\tau_1)) + \sinh(\tau_1) - \frac{2.63861}{2}\, \tau_1\, \sinh(\tau_1) \\ \frac{1}{2}\, p_1(0)\, \tau_1\, \sinh(\tau_1) + \cosh(\tau_1) - \frac{2.63861}{2} (\tau_1\, \cosh(\tau_1) + \sinh(\tau_1)) \end{bmatrix} = \begin{bmatrix} 0.26040 \\ 0 \end{bmatrix}$$

and for $p(t)$,

$$\begin{bmatrix} p_1(0)\, \cosh(\tau_1) - 2.6386\, \sinh(\tau_1) \\ -p_1(0)\, \sinh(\tau_1) + 2.6386 \cosh(\tau_1) \end{bmatrix} = \begin{bmatrix} 0.3723875346 - 0.26040\, (\tau_1 - 9.378729146) \\ 0.26040 \end{bmatrix}$$

Solving these four equations with MAPLE gives $\tau_1 = 0.8614720784$ and $p_1(0) = 3.517682318$. By using these analytical forms, the control, state, multiplier and adjoint state trajectories are drawn in Figures 4.18, 4.19 and 4.20.

Now the numerical solution of this problem is computed as previously. The optimization algorithm mentioned above leads to the optimal control depicted in Figure 4.21. The corresponding state $x_i(t)$ and multiplier $\lambda_i(t)$ trajectories are depicted in Figure 4.22.

In these figures, as in the previous case, one can observe that the constraints are always satisfied, as well as the complementarity conditions. For example, the multiplier $\lambda_2(t)$ is equal to zero as long as $x_2(t)$ is strictly positive. Whenever $x_2(t)$ is equal to zero (at time $\tau_1$), the multiplier $\lambda_2(t)$ becomes instantaneously strictly positive. According to the terminology used in [HSV95, p. 183], the time $\tau_1$ is called an *entry time* for $x_2(t)$ (with respect to the boundary of the nonnegative orthant of the state space) and an *exit time* for $\lambda_2(t)$. Conversely, the time $\tau_2$ is called an *exit time* for $x_2(t)$ and an *entry time* for $\lambda_2(t)$. In addition, the numerical study of initial conditions of the form $x_0 = [\alpha \quad \beta]^T$ where $\alpha \geq 0$ and $\beta > 0$, reveals that, in general, the solution has the same structure as the solution for the initial condition $x_0 = [0 \quad 1]^T$ with two switching times. Unfortunately, it is not possible to obtain the solution for general initial condition $x_0 = [\alpha \quad \beta]^T$ from the study of the solutions for initial conditions $x_0 = [0 \quad 1]^T$ and $x_0 = [0 \quad 1]^T$. Furthermore, since the numerical solution comes from a discretization, there is a scaling factor between the numerical and the analytical expressions of the multipliers which depends on the sampling time $h$, see equation (C.6) in Appendix C. By comparing the previous figures (Figures 4.13 with 4.16 ; 4.14 with 4.17 ; 4.18 with 4.21 and 4.19 with 4.22), one can observe that the two different approaches give similar results.

**Remark 4.3.1** *Here, the minimal energy control problem with final state constraints by using nonnegative input can not be solved by the method of [Ka 02, Subsection 3.4.2] since the reachability gramian is not monomial. However, by increasing the values of some entries of the penalization matrix $S$, the final state $x(t_f)$ can be made closer to 0.*

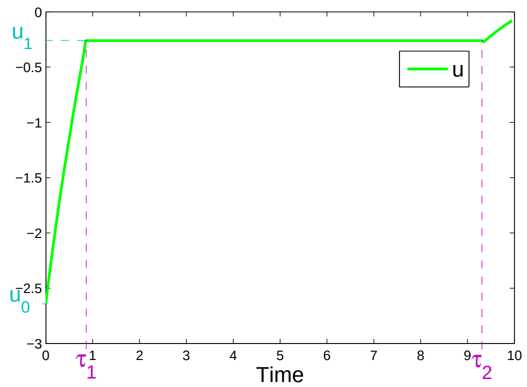
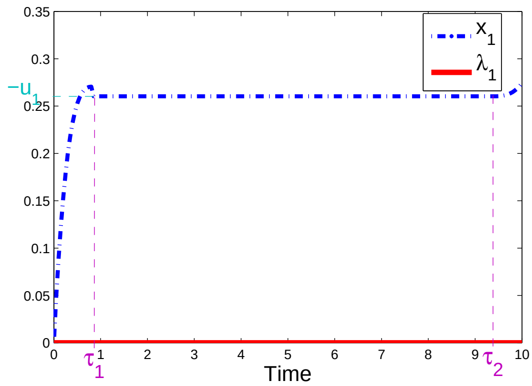



Figure 4.18: Optimal control for system (4.24).



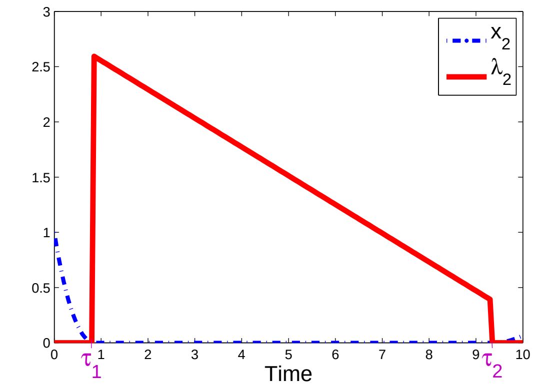
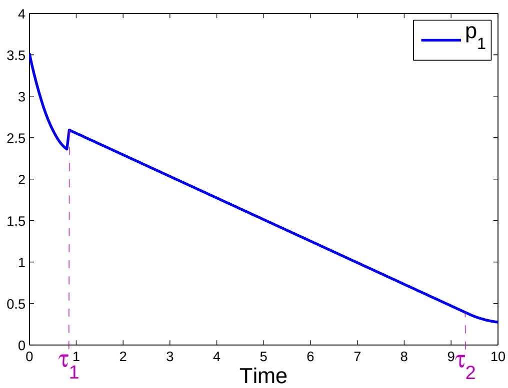




Figure 4.19: State trajectories and associated multipliers for system (4.24).



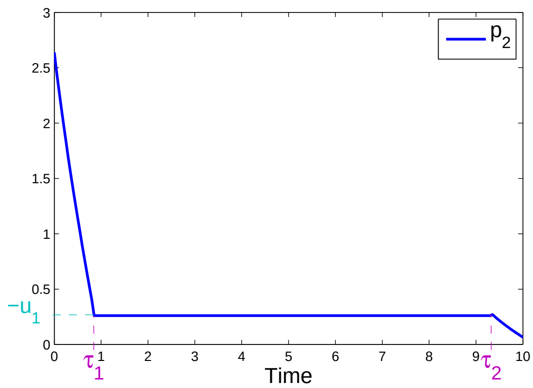




Figure 4.20: Adjoint states for system (4.24).

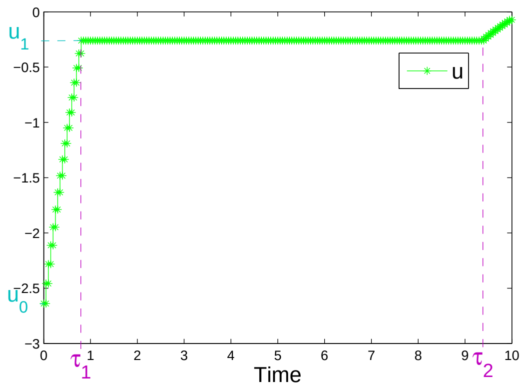



Figure 4.21: Optimal control for sampled data system (4.33).

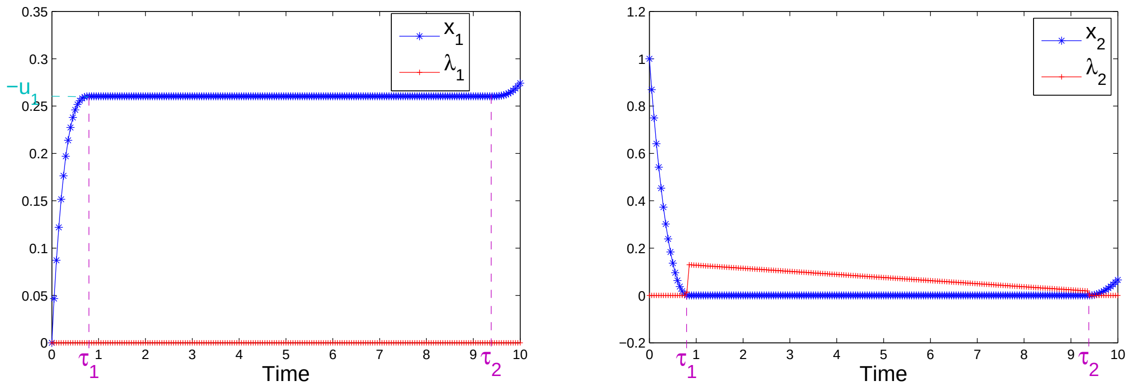



Figure 4.22: State trajectories and associated multipliers for sampled data system (4.33).

# II.2. Infinite Horizon Case

# Chapter 5

# The Input/State-Invariant LQ Problem

This chapter is devoted to the input/state-invariant linear quadratic problem in infinite horizon. Since under appropriate assumptions, the solution of the infinite horizon standard LQ problem is the limit of the finite horizon one, see [WC83], the results of Chapter 3 can be extended to the receding horizon case. These infinite horizon results are briefly described and analyzed.

The LQ problem with input constraints in infinite horizon has already been studied. In the recent paper [Goe10], the infinite horizon LQ problem with conical constraints on the input is studied by means, notably, of a stationary Hamilton-Jacobi equation and by the study of the dual problem. In [HVS98], the infinite horizon LQ problem with nonnegative controls has been studied by means of a receding horizon approach.

## 5.1 Problem statement

Consider the linear time-invariant system description $R = [A,\, B]$ :

$$\dot{x}(t) = A\,x(t) + B\,u(t), \;\; x(0) = x_0, \tag{5.1}$$

where, as previously, the state $x(t)$ and the control $u(t)$ are in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, $A$ and $B$ are real matrices of compatible sizes, $x_0 \in \mathbb{R}^n$ denotes any fixed initial state and $\bar{x}$ is a fixed state.

The infinite horizon input/state-invariant linear quadratic problem, which is denoted by $\boldsymbol{LQ^{\infty}_{\bar{u},\bar{x}}}$, consists of minimizing the quadratic functional :

$$J(x_0,\, u) := \frac{1}{2}\left(\int_0^{\infty} (\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2)\,\mathrm{dt}\right) \tag{5.2}$$

for a given linear system described by (5.1), where the initial state $x_0 \geq \bar{x}$ is fixed, under the constraint

$$\forall\, t \in [0,\, \infty), \;\; \begin{cases} W\,x(t) \geq \bar{x} \\ Z\,u(t) \geq \bar{u}, \end{cases} \tag{5.3}$$

where $u$ is any piecewise-continuous $\mathbb{R}^m$-valued function, $R \in \mathbb{R}^{m\times m}$ is a symmetric positive definite matrix, $C \in \mathbb{R}^{p\times n}$, $W \in \mathbb{R}^{n\times n}$, $Z \in \mathbb{R}^{m\times m}$, and $\bar{x} \in \mathbb{R}^n$ ($\bar{u} \in \mathbb{R}^m$ respectively) is a fixed state (input, respectively). The problem can be studied with any matrices $W$ and $Z$ of full rank. Recall that when $W$ and $Z$ are equal to the zero matrix, the input/state-invariant $\boldsymbol{LQ^{\infty}_{\bar{u},\bar{x}}}$ problem corresponds to the standard $\boldsymbol{LQ^{\infty}}$ problem. For this standard problem, the following results are well-known, see [AM90], [CD91, pp. 333-339], [KS72, Theorem 3.7] and [WC83].

**Theorem 5.1.1** *Assume that the pair $(A,\, B)$ is stabilizable and that the pair $(C,\, A)$ is detectable. Under these conditions, the Algebraic Riccati Equation (ARE)*

$$A^T P + P\,A - P\,B\,R^{-1}B^T P + C^T C = 0 \tag{5.4}$$

*has a unique stabilizing positive semidefinite solution $P_+$, i.e. the ARE has a unique solution $P_+$ such that $P_+$ is positive semidefinite and that the corresponding closed-loop matrix $A - B\,R^{-1}B^T P_+$ is stable. In addition, under these conditions, the optimal control for the $\boldsymbol{LQ^{\infty}}$ problem is of state feedback type and is given by*

$$u(t) = K\,x(t) = -R^{-1}B^T P_+\,x(t) \tag{5.5}$$

*and the optimal cost is $J(x_0,\, u) = x_0^T P_+\, x_0$. Furthermore, the matrix $P_+$ is the limit solution of the Riccati Differential Equation (RDE), more precisely, for any fixed $t \in \mathbb{R}$,*

$$P_+ = \lim_{t_f \to \infty} P(t,\, 0,\, t_f) \tag{5.6}$$

*where $P(t,\, 0,\, t_f)$ denotes the unique solution of the RDE, given by (3.23).*

## 5.2 Receding horizon approach

First, let us consider the following notations and concepts. For a square matrix $A$, $V$ is an $A$-invariant subspace if $A\,V \subset V$. In particular, $\mathcal{L}^-(A)$, $\mathcal{L}^0(A)$, $\mathcal{L}^+(A)$, $\mathcal{L}^{0+}(A)$ denote the $A$-invariant subspaces spanned by the (generalized) eigenvectors corresponding to eigenvalues with negative, zero, positive and nonnegative real parts, respectively. In the sequel, $\mathcal{N}(S)$ denotes the null space of $S$, $NO(C,\, A)$ the unobservable subspace and $ND(C,\, A) := NO(C,\, A) \cap \mathcal{L}^{0+}(A)$ the undetectable subspace.

Using Theorem 5.1.1 on the standard $\boldsymbol{LQ^{\infty}}$ problem together with the following lemmas on the receding horizon approach, leads to appropriate conditions to obtain a solution of the input/state-invariant $\boldsymbol{LQ^{\infty}_{\bar{u},\bar{x}}}$ problem in infinite horizon. A first useful result is developed in [WC83] :

**Lemma 5.2.1** *Assume that $(A,\, B)$ is stabilizable and that the Hamiltonian matrix $H$ (described by (3.7)) has no eigenvalues on the imaginary axis (i.e. $\mathcal{L}^0(H) = \{0\}$). Moreover, assume that $\mathcal{N}(S) \cap ND(C,\, A)$ is $A$-invariant. Let us define the two following problems :*

- $(P_1)$ *: Find $\eta_1^o := \inf\limits_{u(\cdot)} \lim\limits_{t_f \to \infty} \eta(t_f)$ and the corresponding optimal control $u_1^o(\cdot)$ achieving a minimum cost ;*

- $(P_2)$ *: Let* $\eta^*(t_f) := \inf_{u(\cdot)} \eta(t_f)$ *and let* $u^*(\cdot)$ *denotes the corresponding optimal control. Determine* $\eta_2^o(\cdot) := \lim_{t_f \to \infty} \eta^*(t_f)$ *and, if the limit exists, find the limiting behavior* $u_2^o(\cdot)$ *of the optimal control* $u^*(\cdot)$ *for* $t_f \to \infty$ *;*

*where* $\eta(t_f) := J(x_0, u, t_f)$ *denotes the linear quadratic cost in finite horizon (3.2). Then problems* $(P_1)$ *and* $(P_2)$ *yield identical solutions given by*

$$\eta_1^o = \eta_2^o = x_0^T P_+ x_0$$

*and*

$$u_1^o(t) = u_2^o(t) = -R^{-1} B^T P_+ x^o(t).$$

The second result, see [CWW94], concerns the uniform convergence between the optimal state and control trajectories of the finite and infinite horizon LQ problems :

**Lemma 5.2.2** *Assume that* $(A, B)$ *is stabilizable,* $\mathcal{L}^0(H) = \{0\}$ *and* $\mathcal{N}(S) \cap ND(C, A) = \{0\}$. *Let* $x^*(\cdot)$ *and* $u^*(\cdot)$ *be the optimal state and control trajectories on* $[0, t_f]$ *of* $\boldsymbol{LQ^{t_f}}$ *. Let* $e^{(A+B\,K)(\cdot)} x_0$ *and* $-B\,R^{-1} P_+ e^{(A+B\,K)(\cdot)} x_0$ *be the optimal state and control trajectories of the* $\boldsymbol{LQ^\infty}$ *problem. Then*

*a)* $\|x(\cdot) - e^{(A+B\,K)(\cdot)} x_0\|_\infty$ *tends to zero exponentially fast when* $t_f \to \infty$ *;*

*b)* $\|u(\cdot) + B\,R^{-1} P_+ e^{(A+B\,K)(\cdot)} x_0\|_\infty$ *tends to zero exponentially fast when* $t_f \to \infty$ *;*

*where* $\|\cdot\|_\infty$ *denotes the uniform vector norm.*

**Remark 5.2.1** *For a discussion of the assumptions (i.e.* $(A, B)$ *is stabilizable and* $\mathcal{L}^0(H) = \{0\}$*), see e.g. [CW81], where it is stated that the requirement* $\mathcal{L}^0(H) = \{0\}$ *is equivalent to* $\mathcal{L}^0(A) \subset C(A, B)$ *and* $NO(C, A) \cap \mathcal{L}^0(A) = ND(C, A) \cap \mathcal{L}^0(A) = \{0\}$*, i.e. all modes corresponding to the eigenvalues of* $A$ *with zero real part are required to be controllable and observable.*

Using these two lemmas, we obtain the following result for the $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ problem :

**Theorem 5.2.3** *If the solution* $(x^*, u^*)$ *of the standard* $\boldsymbol{LQ^{t_f}}$ *problem is admissible for the* $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ *problem for all sufficiently large horizons* $t_f$*, i.e.*

$$\exists\, T \geq 0 \ \text{such that} \ \forall\, t_f \geq T, \forall\, t \in [0, t_f], \ \begin{cases} W\,x^*(t, t_f) & \geq & \bar{x} \\ Z\,u^*(t, t_f) & \geq & \bar{u}, \end{cases}$$

*and if*

$$(A, B) \ \text{is stabilizable,} \tag{5.7}$$

$$H \ \text{has no eigenvalues on the imaginary axis,} \tag{5.8}$$

$$\text{and} \qquad \mathcal{N}(S) \cap ND(C, A) = \{0\} \tag{5.9}$$

*then the solution* $(x^o, u^o)$ *of the* $\boldsymbol{LQ^\infty}$ *problem is admissible for the* $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ *problem and is therefore solution of the* $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ *problem.*

**Proof :** By Lemmas 5.2.1 and 5.2.2,

$$x^0(t) = \lim_{t_f \to \infty} x^*(t,\, t_f) \;\text{ and }\; u^0(t) = \lim_{t_f \to \infty} u^*(t,\, t_f)$$

such that, with $W\,x^*(t,\, t_f) \geq \bar{x}$ and $Z\,u^*(t,\, t_f) \geq \bar{u}$ and the uniform convergence of the state and control trajectories, the inequalities are preserved by taking the limit :

$$W\,x^0(t) \geq \bar{x} \;\text{ and }\; Z\,u^0(t) \geq \bar{u}.$$

Hence the solution $(x^o,\, u^o)$ of the $\boldsymbol{LQ^\infty}$ problem, is admissible for the $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ and is therefore a solution of the $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ problem. □

**Remark 5.2.2** *Note that the conditions* $(A,\, B)$ *is stabilizable and* $\mathcal{L}^0(H) = \{0\}$ *hold if* $(A,\, B)$ *is stabilizable and* $(C,\, A)$ *is detectable. This yields the following corollary :*

**Corollary 5.2.4** *Assume that* $(A,\, B)$ *is stabilizable and* $(C,\, A)$ *is detectable. If the solution* $(x^*,\, u^*)$ *of the standard* $\boldsymbol{LQ^{t_f}}$ *problem is admissible for the* $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ *problem for all sufficiently large horizons* $t_f$*, then the solution* $(x^o,\, u^o)$ *of the* $\boldsymbol{LQ^\infty}$ *problem is admissible for the* $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ *problem and is therefore solution of the* $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ *problem.*

Therefore the optimality conditions via admissibility developed in Sections 3.4, 3.5.2 and 3.6.2, can be applied here. The following result is stated for the input/state-invariant $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ problem. However, observe that it is readily extendable to the $\boldsymbol{LQ^{t_f}_{\bar{x}}}$ and $\boldsymbol{LQ^{t_f}_{\bar{u}}}$ problems in infinite horizon.

**Theorem 5.2.5** *Assume that* $(A,\, B)$ *is stabilizable and* $(C,\, A)$ *is detectable. Consider* $(x^*,\, u^*)$ *the optimal solution of the* $\boldsymbol{LQ^{t_f}}$ *problem such that* $u^*(t) = K\,x^*(t)$*. Then the solution of the* $\boldsymbol{LQ^\infty}$ *problem is admissible for the* $\boldsymbol{LQ^\infty_{\bar{u},\bar{x}}}$ *problem if, for all sufficiently large horizons* $t_f$*, the matrix solution of the standard matrix Hamiltonian differential equation (3.22) is such that for all* $t \in [0,\, t_f]$*,*

$$W\,X(t)\,X(0)^{-1}x_0 \geq \bar{x}$$

*and*

$$-Z\,R^{-1}B^TY(t)\,X(0)^{-1}x_0 \geq \bar{u}.$$

# Chapter 6

# The Positive LQ Problem

This chapter is devoted to the positive LQ problem in infinite horizon. This problem corresponds to the $\boldsymbol{LQ}^{\boldsymbol{\infty}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem with $\bar{x} = 0$, $W = I_n$ and $\bar{u} = 0$, $V = 0_m$. Criteria for the existence of a solution to this problem are established, in terms of the weighting matrices defining the quadratic cost criterion to be minimized. These criteria are obtained by using a Newton-type iterative scheme (known in the control literature as the Kleinman method) converging to the unique stabilizing positive semi-definite solution of the algebraic Riccati equation, for LTI positive systems. The approach which is used here is inspired by the one developed in [GL00a, GL00b]. This method was recently extended to positive game theory, see [JK04]. Also, the Kronecker product is often used in this part and the properties of M-matrices, see Sections A.4 and A.2.

Positivity criteria are also established in terms of the solution of the algebraic Riccati equation (ARE) and in terms of the Hamiltonian matrix $H$ by using the characterization of positive systems with scalar products, see [AS03] and [DL04]. Finally, as in the finite horizon case, a diagonal solution of the algebraic Riccati equation is studied.

The LQ problem for positive systems is studied in [HCH10, Chapter 13] by optimizing the cost within a class of fixed-structure controllers satisfying internal controller constraints that guarantee the positivity of the closed-loop system.

## 6.1 Newton iterative scheme

### 6.1.1 Problem statement

Consider the following LTI system description :

$$\dot{x}(t) = A\,x(t) + B\,u(t), \;\; x(0) = x_0, \tag{6.1}$$

$$y(t) = C\,x(t), \tag{6.2}$$

where the state $x(t)$, the control $u(t)$ and the output $y(t)$ are in $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^p$, respectively, $A$, $B$ and $C$ are real matrices of compatible sizes, and $x_0 \in \mathbb{R}^n$ denotes any initial state.

The infinite horizon positive LQ problem, which is denoted by $\boldsymbol{LQ_+^\infty}$, consists of minimizing the quadratic functional (5.2),

$$J(x_0,\, u) = \frac{1}{2}\left(\int_0^\infty \|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2\,\mathrm{dt}\right),$$

for a given positive system described by (6.1), where the initial state $x(0) = x_0 \geq 0$ is fixed, under the constraint

$$\forall\, t \in [0,\, \infty),\;\; x(t) \geq 0,$$

and $R \in \mathbb{R}^{m\times m}$ and $Q = C^T C \in \mathbb{R}^{n\times n}$ are assumed to be positive-definite (pd) and positive-semidefinite (psd), respectively, symmetric matrices.

As already studied in the previous chapters, conditions such that the $\boldsymbol{LQ_{\bar{u},\bar{x}}^\infty}$ problem has a solution can be obtained by using the standard $\boldsymbol{LQ^\infty}$ problem. Therefore, our main objective is to find necessary and/or sufficient conditions on the weighting matrices $Q$ and $R$ in the cost $J$ defined by (5.2), such that there exists a state feedback $K$ such that the $\boldsymbol{LQ^\infty}$-optimal closed-loop system $[A + B\,K = A - B\,R^{-1}B^T P_+, 0]$ is positive, i.e. such that the closed-loop matrix $A + B\,K = A - B\,R^{-1}B^T P_+$ is a Metzler matrix.
Therefore, by Theorem 2.3.3, the assumption of positive stabilizability is a necessary condition for the existence of a solution of the $\boldsymbol{LQ_+^\infty}$ problem and by Theorem 5.1.1, stabilizability and detectability are needed assumptions for the $\boldsymbol{LQ^\infty}$ problem. So, from now on, conditions

$$(H_0) \qquad \begin{cases} [A,\, B] & \text{is a positive system} \\ (A,\, B) & \text{is positively stabilizable} \\ (C,\, A) & \text{is detectable} \end{cases} \tag{6.3}$$

will be assumed to hold throughout unless otherwise stated.
The following proposition highlights an important property of the solution $P_+$ of the ARE :

**Proposition 6.1.1** *If $A$ is a stable Metzler matrix and the solution $P_+$ of the ARE, given by (5.4), i.e.*

$$A^T P_+ + P_+\,A - P_+\,B\,R^{-1}B^T P_+ + C^T C = 0,$$

*is such that $A - B\,R^{-1}B^T P_+$ is a (stable) Metzler matrix, then $P_+ \geq 0$ whenever $Q \geq 0$.*

**Proof :** Indeed, using the Kronecker product, recalled in Appendix A.4, the Algebraic Riccati equation (ARE) can be rewritten as

$$\left[I_n \otimes (-A^T) + (-A + B\,R^{-1}B^T P_+)^T \otimes I_n\right]\;\mathrm{vect}(P_+) = \mathrm{vect}(Q). \tag{6.4}$$

Observe that, by Theorem A.2.2, $-A$ and $-A + B\,R^{-1}B^T P_+$ are nonsingular $M$-matrices. Indeed, since $A$ is a Metzler matrix, $-A$ is a $Z$-matrix by definition. Then, since $A$ stable means that its eigenvalues have negative real parts by Theorem 1.2.2, the eigenvalues of $-A^T$ and $(-A + B\,R^{-1}B^T P_+)^T$ have positive real parts. Therefore, these matrices are nonsingular $M$-matrices. It follows that $[I_n \otimes (-A^T) + (-A + B\,R^{-1}B^T P_+)^T \otimes I_n]$ is also a nonsingular $M$-matrix (by Theorem A.4.2), such that its inverse is nonnegative (by Theorem A.2.2). Hence, if the matrix $Q$ is nonnegative, then so is $P_+$, by equation (6.4).

$\square$

### 6.1.2 Positivity criteria for the stable case

Here sufficient conditions which guarantee the positivity of the $\boldsymbol{LQ^\infty}$-optimal closed-loop system are established. First let us introduce the following assumptions :

$(H_1)$ The matrix $A$ is stable.
$(H_2)$ The weighting matrices $Q$ and $R$ are such that

$$Q \gg 0 \quad \text{and} \quad B\,R^{-1}B^T \geq 0.$$

$(H_3)$ $-A + B\,R^{-1}B^T X_1$ is a $Z$-matrix, where $X_1$ is the solution of the Lyapunov equation

$$A^T X_1 + X_1\,A + Q = 0. \tag{6.5}$$

**Remark 6.1.1** *Assumption $(H_3)$ can be read $(B\,R^{-1}B^T X_1)_{ij} \leq a_{ij}$, for $i \neq j$, where $-A$ is a $Z$-matrix, $B\,R^{-1}B^T X_1 \geq 0$ by assumption $(H_2)$ and by the fact that $X_1 \geq 0$, see Lemma 6.1.2 below. Then the matrix $R$ has to be chosen sufficiently large to guarantee this assumption. Also $Q$ can be chosen such that $X_1$ is sufficiently small for this assumption, see the methodology developed below to find the weighting matrices $Q$ and $R$.*

Observe that the matrix $D := -A$ is a nonsingular $M$-matrix, by Theorem A.2.2. Then the algebraic Riccati equation (5.4) can be written equivalently as :

$$D^T P_+ + P_+ D + P_+ B\,R^{-1}B^T P_+ \quad = \quad Q. \tag{6.6}$$

Now let us consider the following iterative scheme :

$$\begin{cases} X_0 = 0 \\ (D^T + X_k\,B\,R^{-1}B^T)\,X_{k+1} + X_{k+1}\,(D + B\,R^{-1}B^T X_k) \\ \qquad = X_k\,B\,R^{-1}B^T X_k + Q, \qquad\qquad \forall\, k \geq 0 \end{cases} \tag{6.7}$$

Observe that for $k = 0$, equation (6.7) is equivalent to the Lyapunov equation (6.5). The following auxiliary result can be proved by induction.

**Lemma 6.1.2** *Consider a LTI system $[A,\, B]$, described by (6.1), such that conditions $(H_0)$–$(H_3)$ hold. Then $(X_k)_{k\geq 1}$, defined by (6.7), is a psd matrix sequence, which is (elementwise) decreasing such that, for all $k \geq 1$,*

$$0 \leq X_{k+1} \leq X_k \leq X_1$$

*and*

$$\left[I_n \otimes (D^T + X_k B\,R^{-1}B^T) + (D + B\,R^{-1}B^T X_k)^T \otimes I_n\right] \text{ is a nonsingular M-matrix.}$$

**Proof :** Define for $k \geq 1$,

$$\mathcal{R}(X_k) := D^T X_k + X_k\,D + X_k\,B\,R^{-1}B^T X_k - Q.$$

Let $k = 0$ in (6.7), that gives, with $X_0 = 0$,

$$D^T X_1 + X_1 D = Q,$$

or equivalently

$$[I_n \otimes D^T + D^T \otimes I_n] \ \mathrm{vect}(X_1) = \mathrm{vect}(Q).$$

Then, since $D$ is a $M$-matrix, it is the same for $[I_n \otimes D^T + D^T \otimes I_n]$ by Theorem A.4.2. Hence by Proposition A.2.3, with $Q \gg 0$, $\mathrm{vect}(X_1) \geq 0$, that is $X_1 \geq 0$. Moreover, $X_1$ is a psd matrix by Theorem 1.2.3. Now, with $D$ a nonsingular $M$-matrix, $X_1 \geq 0$, a psd matrix, $Q \gg 0$ and $D^T X_1 + X_1 D = Q$, one has

$$(D^T + X_1 B R^{-1} B^T) X_1 + X_1 (D + B R^{-1} B^T X_1) = Q + 2 X_1 B R^{-1} B^T X_1 \gg 0,$$

or equivalently

$$[I_n \otimes (D^T + X_1 B R^{-1} B^T) + (D + B R^{-1} B^T X_1)^T \otimes I_n] \ \mathrm{vect}(X_1)$$

$$= \mathrm{vect}(Q + 2 X_1 B R^{-1} B^T X_1) \gg 0.$$

By $(H_3)$, $D + B R^{-1} B^T X_1$ is a $Z$-matrix with $X_1$ solution of $D^T X_1 + X_1 D = Q$. Then $[I_n \otimes (D^T + X_1 B R^{-1} B^T) + (D + B R^{-1} B^T X_1)^T \otimes I_n]$ is a $Z$-matrix and a nonsingular $M$-matrix by Theorem A.2.2.
Moreover, with $D^T X_1 + X_1 D = Q$, $X_1 \geq 0$ and $B R^{-1} B^T \geq 0$,

$$\mathcal{R}(X_1) = X_1 B R^{-1} B^T X_1 \geq 0.$$

By calculation, one has

$$(D^T + X_1 B R^{-1} B^T) (X_1 - X_2) + (X_1 - X_2) (D + B R^{-1} B^T X_1) = \mathcal{R}(X_1) \geq 0,$$

which can be written equivalently as

$$[I_n \otimes (D^T + X_1 B R^{-1} B^T) + (D + B R^{-1} B^T X_1)^T \otimes I_n] \ \mathrm{vect}(X_1 - X_2) = \mathrm{vect}(\mathcal{R}(X_1)).$$

Then since $[I_n \otimes (D^T + X_1 B R^{-1} B^T) + (D + B R^{-1} B^T X_1)^T \otimes I_n]$ is a nonsingular matrix, by Theorem A.2.2, its inverse is nonnegative, and, with $\mathcal{R}(X_1) \geq 0$, we have $\mathrm{vect}(X_1 - X_2) \geq 0$, i.e. $X_2 \leq X_1$, and the recurrency is verified for $k = 1$.

Now assume that for a fixed $k \geq 1$, the following assumptions hold :

$$\begin{gathered} X_k \text{ is a psd matrix} \\ 0 \leq X_{k+1} \leq X_k \leq X_1 \\ \text{and} \\ \left[I_n \otimes (D^T + X_k B R^{-1} B^T) + (D + B R^{-1} B^T X_k)^T \otimes I_n\right] \\ \text{is a nonsingular } M\text{-matrix.} \end{gathered} \tag{6.8}$$

Let us show that these assertions hold for $X_{k+1}$ et $X_{k+2}$. As done previously for $X_1$, one can show that $X_{k+1}$ is a psd matrix. Then since $B\,R^{-1}B^T \geq 0,\; X_{k+1} \geq 0$ and $X_1 \geq 0$, one has

$$(D^T + X_{k+1}\,B\,R^{-1}B^T)\,X_1 + X_1\,(D + B\,R^{-1}B^T X_{k+1}) \geq D^T X_1 + X_1\,D = Q \gg 0.$$

Consequently, $[I_n \otimes (D^T + X_{k+1}\,B\,R^{-1}B^T) + (D + B\,R^{-1}B^T X_{k+1})^T \otimes I_n]$ is a nonsingular $M$-matrix, by Theorem A.2.2. Hence, using the iterative scheme (6.7) :

$$[I_n \otimes (D^T + X_{k+1}\,B\,R^{-1}B^T) + (D + B\,R^{-1}B^T X_{k+1})^T \otimes I_n]\;\mathrm{vect}(X_{k+2})$$

$$= \mathrm{vect}(X_{k+1}\,B\,R^{-1}B^T X_{k+1} + Q) \gg 0.$$

Then, by Proposition A.2.3, $X_{k+2} \geq 0$.
Moreover, by calculation, one has

$$\mathcal{R}(X_{k+1}) = (X_{k+1} - X_k)\,B\,R^{-1}B^T(X_{k+1} - X_k) \geq 0,$$

since by $(H_2)$, $B\,R^{-1}B^T \geq 0$ and since $(X_{k+1} - X_k) \leq 0$. Therefore, by developing the iterative scheme (6.7), we have :

$$(D^T + X_{k+1}\,B\,R^{-1}B^T)\,(X_{k+2} - X_{k+1}) + (X_{k+2} - X_{k+1})\,(D + B\,R^{-1}B^T X_{k+1})$$

$$= -\mathcal{R}(X_{k+1}) \leq 0.$$

Since $[I_n \otimes (D^T + X_{k+1}\,B\,R^{-1}B^T) + (D + B\,R^{-1}B^T\,X_{k+1})^T \otimes I_n]$ is a nonsingular $M$-matrix of nonnegative inverse, the matrix $(X_{k+2} - X_{k+1})$ is nonpositive and then $X_{k+2} \leq X_{k+1}$.

□

It follows from this lemma that one can take the limit in (6.7) and that the following theorem can then be established easily.

**Theorem 6.1.3** *Consider a LTI system $[A,\,B]$, described by (6.1), such that conditions $(H_0)$–$(H_3)$ hold. Then the stabilizing psd solution $P_+$ of the algebraic Riccati equation (5.4) is such that the corresponding* ***LQ***$^{\boldsymbol{\infty}}$ *-optimal closed-loop system is positive, i.e. the closed-loop matrix $(A - B\,R^{-1}B^T P_+)$ is a Metzler matrix.*

**Proof :** By Lemma 6.1.2, the sequence $(X_k)_{k\geq 1}$ of scheme (6.7) is a decreasing sequence of psd matrices such that $X_k \geq 0$ for $k \geq 1$. Hence $\lim\limits_{k\to+\infty} X_k = P_+ \geq 0$ exists and is equal to a psd matrix $P_+$. Therefore, taking the limit with $k \to +\infty$ in (6.7) gives the algebraic Riccati equation (6.6). Hence, $P_+ \geq 0$ is the psd solution of the ARE. Thanks to the assumption of stabilizability and detectability, classical theory says that this solution is unique and stabilizing, see [CD91, Theorem 38, p. 348] and Theorem 5.1.1. Moreover, by assumption $(H_3)$, using the fact that the sequence $(X_k)_{k\geq 1}$ is decreasing, $-A + B\,R^{-1}B^T X_k$ is a $Z$-matrix and so is $-A + B\,R^{-1}B^T P_+$ by taking the limit. Then $A - B\,R^{-1}B^T P_+$ is a Metzler matrix.

□

**Remarks 6.1.2** *a) By Proposition A.2.3, it can be shown that $X_k \gg 0$. However, with $X_k \geq 0$ or $X_k \gg 0$, the limit $P_+$ is nonnegative in both cases.*
*b) By Theorem A.2.2, since $D$ is a nonsingular $M$-matrix, there exists a diagonal pd matrix $P$ such that $D^T P + P\,D$ is a pd matrix. Therefore, by assumption $(H_3)$, with $X_1$ solution of the Lyapunov equation $D^T X_1 + X_1\,D = Q$, it is possible to take $X_1$ a diagonal matrix with diagonal elements strictly positiv,e as initial iterate of scheme (6.7). In this case, the choice of the matrix $Q$ is dictated by $Q = D^T X_1 + X_1\,D$.*

It turns out that the assumption $(H_3)$ can be replaced by the following one :

$(H_4)$ There exists a symmetric matrix $Y \gg 0$ such that

$$Q + A^T Y + Y A \leq 0 \quad \text{and} \quad -A + BR^{-1}B^T Y \ \text{ is a Z-matrix.}$$

**Theorem 6.1.4** *The conclusion of Theorem 6.1.3 remains valid if the assumption $(H_3)$ is replaced by $(H_4)$.*

**Remark 6.1.3** *With the assumptions $(H_1)$, $(H_2)$ and $(H_4)$, the iterative scheme (6.7) yields a sequence $(X_k)$ such that for all $k \geq 1$,*

$$\begin{cases} X_k \text{ is symmetric positive definite, } 0 \leq X_{k+1} \leq X_k \leq Y, \text{ and} \\ \left[I_n \otimes (D + BR^{-1}B^T X_k)^T + (D + BR^{-1}B^T X_k)^T \otimes I_n\right] \text{ is a } M\text{-matrix.} \end{cases}$$

**Proof :** It is clear that $(H_3)$ implies $(H_4)$. Show that $(H_4)$ implies $(H_3)$. Indeed, consider the matrix $Y$ such that $-A + B\,R^{-1}B^T Y$ is a $Z$-matrix, then

$$0 \leq \left(B\,R^{-1}B^T Y\right)_{ij} \leq a_{ij}, \ \text{ for } i \neq j. \tag{6.9}$$

Now, with $D = -A$,

$$\begin{aligned} D^T(X_1 - Y) + (X_1 - Y)\,D &= D^T X_1 + X_1\,D - D^T Y - Y\,D \\ &= Q - D^T Y - Y\,D \leq 0 \end{aligned}$$

or equivalently

$$\left[I_n \otimes D^T + D^T \otimes I_n\right] \ \mathrm{vect}(X_1 - Y) = \mathrm{vect}(Q - D^T Y - Y\,D) \leq 0$$

Hence

$$\mathrm{vect}(X_1 - Y) = \underbrace{\left[I_n \otimes D^T + D^T \otimes I_n\right]^{-1}}_{\geq 0} \underbrace{\mathrm{vect}(Q - D^T Y - Y\,D)}_{\leq 0} \leq 0,$$

since $\left[I_n \otimes D^T + D^T \otimes I_n\right]$ is a nonsingular $M$-matrix. Therefore $X_1 - Y \leq 0$, so $X_1 \leq Y$. Hence, with (6.9) :

$$0 \leq \left(B\,R^{-1}B^T X_1\right)_{ij} \leq \left(B\,R^{-1}B^T Y\right)_{ij} \leq a_{ij}, \ \text{ for } i \neq j.$$

Then, $-A + B\,R^{-1}B^T X_1$ is a $Z$-matrix. Now, assume that $A^T Y + Y\,A \leq -Q$. Consider a matrix $P$ such that $P \geq Q$ and take $Y$ solution of the following Lyapunov equation

$$A^T Y + Y\,A = -P \leq -Q$$

and we obtain the condition of assumption $(H_3)$. $\square$

**– Application to compartmental systems**

Consider a compartmental system, (see Definition 2.3.3), where the matrix $A$ satisfies condition (2.3), i.e.

$$\sum_{i=1}^{n} a_{ij} < 0 \ \text{ for all } \ j = 1, \ldots, n.$$

This condition implies that $Z\,A \ll 0$, i.e.

$$A^T Z + Z\,A \ll 0$$

where $Z = (z_{ij}) \in \mathbb{R}^{n\times n}$ is defined as follows

$$\forall\, i,\ j = 1, ..., n, \quad z_{ij} = 1.$$

In this case, it turns out that the assumption $(H_2)$ can be replaced by the following one :

$(H_5)$ The weighting matrices $Q$ and $R$ are such that

$$Q \geq 0 \text{ and } B\,R^{-1}B^T \geq 0.$$

**Theorem 6.1.5** *Assume that $[A,\ B]$ is a compartmental system such that (2.3) holds. Then the conclusion of Theorem 6.1.3 remains valid if the assumption $(H_2)$ is replaced by $(H_5)$.*

**Proof :** The proof is similar to the one of Theorem 6.1.3 and follows the lines of Lemma 6.1.2. First recall the iterative scheme (6.7) used in these proofs :

$$\begin{cases} X_0 = 0 \\ (D^T + X_k\,B\,R^{-1}B^T)\,X_{k+1} + X_{k+1}\,(D + B\,R^{-1}B^T X_k) \\ \quad = X_k\,B\,R^{-1}B^T X_k + Q, \qquad\qquad \forall\, k \geq 0 \end{cases}$$

First show that $(X_k)_{k\geq 1}$ is a decreasing sequence of psd matrices such that

$$\forall\, k \geq 1\ :\ 0 \leq X_{k+1} \leq X_k \leq X_1$$

and

$$\left[I_n \otimes (D^T + X_k\,B\,R^{-1}B^T) + (D + B\,R^{-1}B^T X_k)^T \otimes I_n\right]$$

is a nonsingular $M$-matrix.

Let $k = 0$ in (6.7), that gives, with $X_0 = 0$,

$$D^T X_1 + X_1\,D = Q,$$

or equivalently

$$[I_n \otimes D^T + D^T \otimes I_n]\ \mathrm{vect}(X_1) = \mathrm{vect}(Q).$$

Then, since $D$ is a $M$-matrix, it is the same for $[I_n \otimes D^T + D^T \otimes_n]$ by Theorem A.4.2. Hence its inverse is nonnegative by Theorem A.2.2 and then

$$\mathrm{vect}(X_1) = \underbrace{[I_n \otimes D^T + D^T \otimes I_n]^{-1}}_{\geq 0} \underbrace{\mathrm{vect}(Q)}_{\geq 0},$$

one has $X_1 \geq 0$. Moreover, $X_1$ is a psd matrix by Theorem 1.2.3. Now, with $D$ a nonsingular $M$-matrix, $X_1 \geq 0$, a psd matrix, $Q \geq 0$ and $D^T Z + Z\,D \gg 0$, one has

$$(D^T + X_1\, B\, R^{-1} B^T)\, Z + Z\,(D + B\, R^{-1} B^T X_1) \geq D^T Z + Z\, D \gg 0,$$

or equivalently

$$[I_n \otimes (D^T + X_1\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_1)^T \otimes I_n]\ \mathrm{vect}(Z)$$

$$= \mathrm{vect}(D^T Z + Z\, D) \gg 0.$$

By $(H_3)$, $D + B\,R^{-1}B^T X_1$ is a $Z$-matrix with $X_1$ solution of $D^T X_1 + X_1\, D = Q$. Then $[I_n \otimes (D^T + X_1\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_1)^T \otimes I_n]$ is a $Z$-matrix and a nonsingular $M$-matrix by Theorem A.2.2. By using the same arguments as in the proof of Lemme 6.1.2, it can be shown that $X_2 \leq X_1$ and the recurrency is verified for $k = 1$.

Now assume that for a fixed $k \geq 1$, the following assumptions hold :

$$\begin{gathered} X_k \text{ is a psd matrix} \\ 0 \leq X_{k+1} \leq X_k \leq X_1 \\ \text{and} \\ \left[I_n \otimes (D^T + X_k\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_k)^T \otimes I_n\right] \\ \text{is a nonsingular } M\text{-matrix.} \end{gathered} \tag{6.10}$$

As done previously for $X_1$, one can show that $X_{k+1}$ is a psd matrix. Then since $B\,R^{-1}B^T \geq 0$, $X_{k+1} \geq 0$ and $Z \gg 0$, one has

$$(D^T + X_{k+1}\, B\, R^{-1} B^T)\, Z + Z\,(D + B\, R^{-1} B^T X_{k+1}) \geq D^T Z + Z\, D \gg 0.$$

Consequently, $[I_n \otimes (D^T + X_{k+1}\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_{k+1})^T \otimes I_n]$ is a nonsingular $M$-matrix, by Theorem A.2.2. Hence, using the iterative scheme (6.7) :

$$[I_n \otimes (D^T + X_{k+1}\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_{k+1})^T \otimes I_n]\ \mathrm{vect}(X_{k+2})$$

$$= \mathrm{vect}(X_{k+1}\, B\, R^{-1} B^T X_{k+1} + Q) \geq 0.$$

Then, $X_{k+2} \geq 0$ since by Theorem A.2.2, $[I_n \otimes (D^T + X_{k+1}\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_{k+1})^T \otimes I_n]$ is a nonsingular $M$-matrix of nonnegative inverse. Moreover, by calculation, using the scheme (6.7) as in the proof of Lemma 6.1.2, one has $X_{k+2} \leq X_{k+1}$. The rest of the proof is similar to the one of Theorem 6.1.3 by taking the limit in (6.7).

□

– **Design methodology to find $Q$ and $R$**

An important question is to find suitable matrices $Q$ and $R$ such that the assumptions $(H_2)$ and $(H_4)$ (or $(H_3)$) are satisfied. Concerning the matrix $Q$, it is not hard to choose $Q = C^TC$ such that $C \gg 0$ and $(C, A)$ is detectable according to the assumption $(H_0)$. If the system $[A,\ B]$ is compartmental, the matrix $C$ can be selected such that $C \geq 0$ as stated in the assumption $(H_5)$. For the matrix $R$, it is possible to choose

$$R: \ = \ s\,I_m - \tilde{R} \ \text{ such that } \ \tilde{R}^T = \tilde{R} \geq 0 \text{ and } s > \rho(\tilde{R}). \tag{6.11}$$

It is clear that $R$ is a nonsingular pd matrix such that $R^{-1} \geq 0$. Consequently, $B\,R^{-1}B^T \geq 0$. Now let $P$ be a symmetric psd matrix such that $P \geq Q$ and consider the solution $Y$ of the following Lyapunov equation :

$$A^TY + Y\,A = -P. \tag{6.12}$$

Since $A$ is a stable matrix, the solution of (6.12) is given by :

$$Y = \int_0^{+\infty} e^{A^Tt}P\,e^{At}dt.$$

Obviously $Y = Y^T$ is a psd matrix and $Y \geq 0$ , since $Q \gg 0$. Moreover

$$A^TY + Y\,A + Q \leq 0.$$

In order to check the feasibility of the second part of assumption $(H_4)$, i.e. $A - B\,R^{-1}B^TY$ is a Metzler matrix, we assume that the following condition holds :

$$\alpha := \min\{a_{ij} : i, j = 1, \cdots, n \text{ such that } i \neq j\} > 0. \tag{6.13}$$

In view of (6.11) and since $\lim\limits_{s \to +\infty}(s\,I_m - \tilde{R})^{-1} = 0$, there is some sufficiently large $s > \rho(\tilde{R})$ such that $-A + B\,R^{-1}B^TY$ is a $Z$-matrix. Therefore one gets the following result :

**Proposition 6.1.6** *If $[A,\ B]$ is a positive system such that $A$ is stable and condition (6.13) holds, then there exist weighting matrices $Q$ and $R$ such that $(H_2)$ and $(H_4)$ are satisfied.*

**Remark 6.1.4** *The numerical example below reveals that it is not easy to find such $s$. In fact, it shows that $R$ has to be chosen "sufficiently large" (i.e. $s$ sufficiently large such that $s > \rho(\tilde{R})$) and $Q$ "sufficiently small" in order to satisfy assumption $(H_4)$ which depends on $Q$ and $Y$. Indeed, the condition $(H_4)$ can be read as $B\,R^{-1}B^TY$ be a sufficiently small perturbation to keep the positivity property of $A$. Then the parameter $\alpha > 0$ provides a degree of freedom to disturb $A$ in order to keep its Metzler property. Therefore, in the cost (5.2), the penalization coefficient of the state is less than the weight of the control.*

By Theorem 6.1.4, the latter result together with the paragraph above indicate a design methodology for choosing the weighting matrices $Q$ and $R$ in order to get the positivity of the resulting $\boldsymbol{LQ^\infty}$-optimal closed-loop system, see Table 6.1.

1. Choose $Q = C^T C$ with $C \gg 0$ ;
2. Choose $R = s\,I_m - \tilde{R}$ such that $\tilde{R} \geq 0$ symmetric matrix and $s > \rho(\tilde{R})$, sufficiently large ;
3. Choose a symmetric psd matrix $P$ such that $P \geq Q$ ;
4. Compute the solution $Y$ of the Lyapunov equation $A^T Y + Y\,A = -P$ ;
5. Check that $\alpha := \min\{a_{ij} : i \neq j\} > 0$ ;
6. Choose $s$ sufficiently large ($s > \rho(\tilde{R})$) such that $A - B\,R^{-1} B^T Y$ is a Metzler matrix.

Table 6.1: Design methodology for $Q$ and $R$.

**– Numerical example**

Consider the following LTI stable positive system in order to illustrate the design methodology to find $Q$ and $R$,

$$A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \quad ; \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{6.14}$$

The design methodology of Table 6.1 applying to this system yields the following matrices $Q$ and $R$, (detailed calculations are described here after) :

$$Q = \begin{bmatrix} 4 & 0.01 \\ 0.01 & 6 \end{bmatrix} \quad ; \quad R = \begin{bmatrix} 6 & -2 \\ -2 & 5 \end{bmatrix}. \tag{6.15}$$

Consequently, with these weighting matrices, we obtain that the resulting closed-loop system is positive, since the closed-loop matrix is given by

$$A + B\,K = A - B\,R^{-1} B^T P_+ = \begin{bmatrix} -2.2692 & 0.7688 \\ 0.7789 & -2.4100 \end{bmatrix}.$$

These results are obtained by applying the methodology described in Table 6.1 as follows :

1. Choose $Q = C^T C$ with $C \gg 0$

   Take $C = Q^{1/2} = \begin{bmatrix} 2 & 0.022 \\ 0.022 & 2.4495 \end{bmatrix}$ such that

$$Q = C^T C = \begin{bmatrix} 4 & 0.01 \\ 0.01 & 6 \end{bmatrix} \gg 0.$$

2. Choose $R = s\, I_m - \tilde{R}$ such that $\tilde{R} \geq 0$ symmetric matrix and $s > \rho(\tilde{R})$, sufficiently large

   For example, let $\tilde{R} = \begin{bmatrix} 4 & 2 \\ 2 & 5 \end{bmatrix}$ with $\rho(\tilde{R}) = 6.5616$.

   First, we take $s$ as the smallest integer larger than $\rho(\tilde{R})$, then $s = 7$. Unfortunately, as we can seen in Table 6.2, this choice of $s$ is not larger enough in order to obtain that $-A + B\, R^{-1} B^T Y$ is a $Z$-matrix. Therefore, we have to choose at least $s = 10$ to achieve this condition, see below. Then

   $$R = \begin{bmatrix} 6 & -2 \\ -2 & 5 \end{bmatrix}$$

   is a pd matrix such that

   $$B\, R^{-1} B^T = \begin{bmatrix} 0.1923 & 0.0769 \\ 0.0769 & 0.2308 \end{bmatrix} \geq 0.$$

3. Choose $P$ symmetric psd matrix such that $P \geq Q$

   Next, we choose $P \geq Q$ with, for example,

   $$P = Q + \beta\, I_n, \quad \beta = 10,$$

   that gives

   $$P = \begin{bmatrix} 14 & 0.01 \\ 0.01 & 10 \end{bmatrix}.$$

4. Compute $Y$ the solution of the Lyapunov equation $A^T Y + Y\, A = -P$

   The solution $Y$ of the Lyapunov equation is given by

   $$Y = \begin{bmatrix} 4.7517 & 2.5033 \\ 2.5033 & 5.2517 \end{bmatrix}$$

   which is a psd symmetric and nonnegative matrix.

5. Check $\alpha := \min\{a_{ij} : i \neq j\} > 0$

   Since $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$, $\alpha = 1 > 0$.

6. Choose $s$ sufficiently large ($s > \rho(\tilde{R})$) such that $A - B\, R^{-1} B^T Y$ is a Metzler matrix

   As we have seen previously, the parameter $s$ has to be chosen sufficiently large to obtain this last condition. Indeed, see Table 6.2 where several values of $s$ and the resulting matrices $-A + B\, R^{-1} B^T Y$ are given. In this table, the positivity of the closed-loop system is also checked by verifying if the closed-loop matrix is a Metzler matrix.

| $s$ | $-A + B\,R^{-1}B^T Y$ <br> $Y$ Lyapunov solution | **Z-matrix** | $A - B\,R^{-1}B^T P_+$ <br> $P_+$ Riccati solution | **Metzler** |
|---|---|---|---|---|
| 7 | $\begin{bmatrix} 9.2550 & 6.7550 \\ 7.5067 & 12.3808 \end{bmatrix}$ | $\times$ | $\begin{bmatrix} -2.9864 & -0.2193 \\ -0.0639 & -3.7515 \end{bmatrix}$ | $\times$ |
| 8 | $\begin{bmatrix} 4.4077 & 1.2517 \\ 1.4396 & 5.2517 \end{bmatrix}$ | $\times$ | $\begin{bmatrix} -2.4855 & 0.5060 \\ 0.5435 & -2.7888 \end{bmatrix}$ | $\checkmark$ |
| 9 | $\begin{bmatrix} 3.5008 & 0.2823 \\ 0.3762 & 3.9541 \end{bmatrix}$ | $\times$ | $\begin{bmatrix} -2.3427 & 0.6851 \\ 0.7026 & -2.5350 \end{bmatrix}$ | $\checkmark$ |
| 10 | $\begin{bmatrix} 3.1063 & -0.1146 \\ -0.0568 & 3.4045 \end{bmatrix}$ | $\checkmark$ | $\begin{bmatrix} -2.2692 & 0.7688 \\ 0.7789 & -2.4100 \end{bmatrix}$ | $\checkmark$ |

Table 6.2: Table of different values of parameter $s$.

> We can observe that this methodology is only a sufficient condition to guarantee the positivity of the closed-loop system. Since for $s = 8$, the last condition of the methodology is not verified while the closed-loop system is positive.

Now Figure 6.1 represents the optimal state trajectories, for the initial states $x_0 = [1 \quad 0]^T$ (graphs on the left) and $x_0 = [0 \quad 1]^T$ (graphs on the right) respectively, i.e. the columns of $e^{A\,t}$ at the sampling times. One can numerically verify that the closed-loop system is positive.

Notice that the closed-loop system is stable since the eigenvalues of the (constant) closed-loop matrix are $-1.5626$ and $-3.1166$. Moreover, it could also be interesting to observe the behavior of the optimal control $u(t)$, which is represented in Figure 6.2 with the same initial states as above. We can observe that $u(t) \leq 0$ for all $t$ since $u(t) = -R^{-1}B^T P_+\,x(t)$ with $P_+ \geq 0$, $B \geq 0$ and $R^{-1} \geq 0$ by construction. Then the optimal control is always nonpositive.

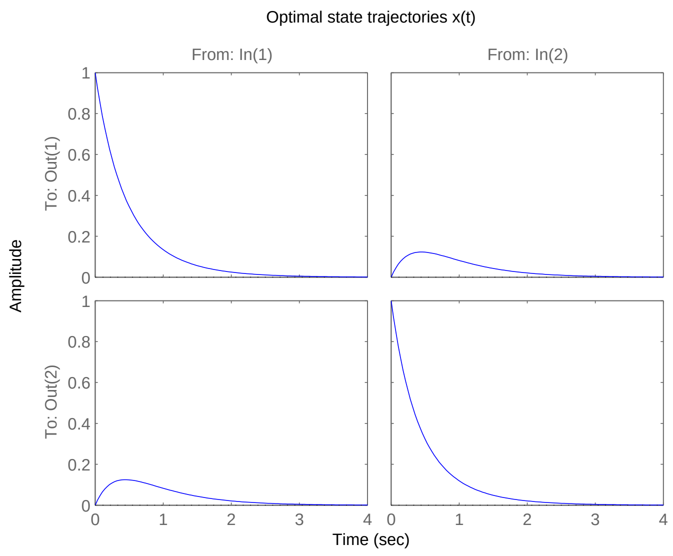



Figure 6.1: Optimal state trajectories $x(t)$ for system (6.14)-(6.15).

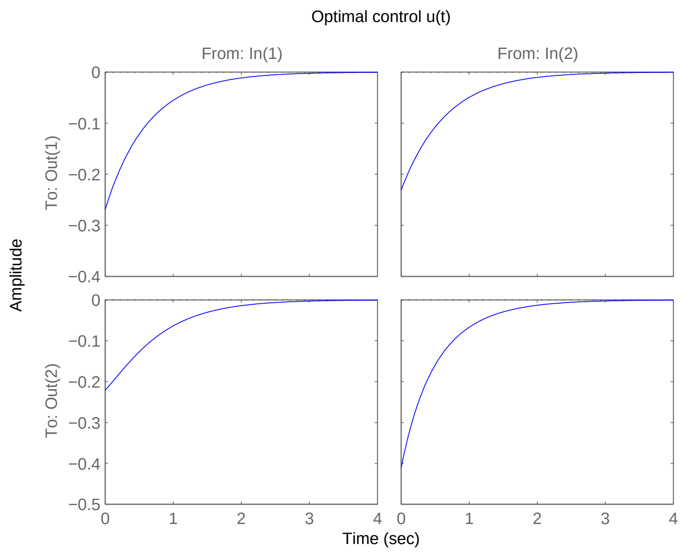



Figure 6.2: Optimal control $u(t)$ for system (6.14)-(6.15).

### 6.1.3 Positivity criteria for the unstable case

In this subsection, we consider the case of a positive system $[A,\ B]$ where $A$ is unstable, i.e.

$$s(A) = \sup\{\mathcal{R}e(\lambda)\ :\ \lambda \in \sigma(A)\} \geq 0.$$

Consequently, $-A$ is no longer a $Z$-matrix, by Theorem A.2.2. Let $D = (s\,I_n - A)$ with $s > s(A)$ such that $-D$ is a stable matrix. This gives a nonsingular $M$-matrix such that the characterizations of such matrices given in Theorem A.2.2 are applicable. Indeed, since $A$ a Metzler matrix, for all $t \geq 0$, $e^{A\,t} \geq 0$, so by the Laplace transform, see [Nag86] or [CD91],

$$\forall\, s > s(A) \quad (s\, I_n - A)^{-1} = \int_0^\infty e^{-s\,t}\, e^{A\,t}\, \mathrm{dt} \geq 0.$$

Therefore, with $s > s(A)$, $D = (s\, I_n - A)$ is a nonsingular $M$-matrix by Theorem A.2.2. As in the stable case, the algebraic Riccati equation is written in terms of $D$ instead of $A$, that gives :

$$D^T P_+ + P_+\, D + P_+\, B\, R^{-1} B^T P_+ = Q + 2\, s\, P_+. \tag{6.16}$$

The following assumption is assumed to hold, similarly to the stable case :

$(H_3')$ There exists $s > s(A)$ such that $-A + B\, R^{-1} B^T X_1$ is a $Z$-matrix, and $X_1\, B\, R^{-1} B^T - 2\, s\, I_n \geq 0$, where $X_1$ is the solution of the following Lyapunov equation :

$$(A - s\, I_n)^T X_1 + X_1\, (A - s\, I_n) = -Q. \tag{6.17}$$

**Remark 6.1.5** *Note that assumption $(H_3')$ is assumption $(H_3)$ of the stable case with $s = 0$.*

**Theorem 6.1.7** *Consider a positive system $[A,\, B]$ where the assumptions $(H_2) - (H_3')$ hold, then the algebraic Riccati equation (5.4) has a psd solution $P_+ \geq 0$ such that $A - B\, R^{-1} B^T P_+$ is a Metzler matrix.*

**Proof :** Consider $s > s(A) \geq 0$ given by $(H_3')$ and define $D = (s\, I_n - A)$ which is a nonsingular $M$-matrix. Introduce the following iterative scheme :

$$\begin{cases} X_0 = 0 \\ (D^T + X_k\, B\, R^{-1} B^T)\, X_{k+1} + X_{k+1}\, (D + B\, R^{-1} B^T X_k) \\ \qquad = Q + 2sX_k + X_k\, B\, R^{-1} B^T X_k, \qquad\qquad \forall\, k \geq 1. \end{cases} \tag{6.18}$$

The proof is similar to the one of Theorem 6.1.3 and follows the lines of Lemma 6.1.2. First show that $(X_k)_{k \geq 1}$ is a decreasing sequence of psd matrices such that

$$\forall\, k \geq 1\ :\ 0 \leq X_{k+1} \leq X_k \leq X_1$$

and

$$\left[I_n \otimes (D^T + X_k\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_k)^T \otimes I_n\right]$$

is a nonsingular $M$-matrix.

Let $k = 0$ in (6.18), that gives, $D^T X_1 + X_1\, D = Q$. As in the stable case, it can be shown that $X_1 \geq 0$, psd matrix such that $[I_n \otimes (D^T + X_1\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_1)^T \otimes I_n]$ is a nonsingular $M$-matrix. Now, by computation, it follows that :

$$(D^T + X_1\, B\, R^{-1} B^T)\, (X_2 - X_1) + (X_2 - X_1)\, (D + B\, R^{-1} B^T X_1)$$

$$= (2s\, I_n - X_1\, B\, R^{-1} B^T)\, X_1 \leq 0$$

by assumption $(H_3')$. Hence, $X_2 \leq X_1$ since $[I_n \otimes (D^T + X_1 B R^{-1} B^T) + (D + B R^{-1} B^T X_1)^T \otimes I_n]$ is a nonsingular $M$-matrix with a nonnegative inverse. Then the recurrency is verified for $k = 1$.

Now assume that for a fixed $k \geq 1$, the following assumptions hold :

$$\begin{gathered} X_k \text{ is a psd matrix} \\ 0 \leq X_{k+1} \leq X_k \leq X_1 \\ \text{and} \\ \left[I_n \otimes (D^T + X_k\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_k)^T \otimes I_n\right] \\ \text{is a nonsingular } M\text{-matrix.} \end{gathered} \tag{6.19}$$

As done previously for $X_1$, one can show that $X_{k+1}$ is a psd matrix. Then since

$$(D^T + X_{k+1}\, B\, R^{-1} B^T)\, X_1 + X_1\, (D + B\, R^{-1} B^T X_{k+1}) \geq D^T X_1 + X_1\, D = Q \gg 0,$$

one has that $\left[I_n \otimes (D^T + X_{k+1}\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T X_{k+1})^T \otimes I_n\right]$ is a nonsingular $M$-matrix by Theorem A.2.2. Thus, by using the iterative scheme (6.18) and by inverting, we have $X_{k+2} \geq 0$. Now let us show that $X_{k+2} \leq X_{k+1}$. Using the scheme (6.18), by computation, it follows that :

$$\begin{aligned} &(D^T + X_{k+1}\, B\, R^{-1} B^T)\, (X_{k+2} - X_{k+1}) + (X_{k+2} - X_{k+1})\, (D + B\, R^{-1} B^T X_{k+1}) \\ &= (D^T + X_{k+1}\, B\, R^{-1} B^T)\, X_{k+2} + X_{k+2}\, (D + B\, R^{-1} B^T X_{k+1}) \\ &- \big[(D^T + X_k\, B\, R^{-1} B^T)\, X_{k+1} + X_{k+1}\, (D + B\, R^{-1} B^T X_k) \\ &+ (X_{k+1} - X_k)\, B\, R^{-1} B^T X_{k+1} + X_{k+1}\, B\, R^{-1} B^T\, (X_{k+1} - X_k)\big] \\ &= Q + 2s X_{k+1} + X_{k+1}\, B\, R^{-1} B^T X_{k+1} \\ &- \big[Q + 2s X_k + X_k\, B\, R^{-1} B^T X_k \\ &+ (X_{k+1} - X_k)\, B\, R^{-1} B^T X_{k+1} + X_{k+1}\, B\, R^{-1} B^T\, (X_{k+1} - X_k)\big] \end{aligned}$$

$$\begin{aligned} &= 2s\, (X_{k+1} - X_k) + X_{k+1}\, B\, R^{-1} B^T X_{k+1} \\ &- X_k\, B\, R^{-1} B^T X_k \\ &- X_{k+1}\, B\, R^{-1} B^T X_{k+1} + X_k\, B\, R^{-1} B^T X_{k+1} \\ &- X_{k+1}\, B\, R^{-1} B^T\, X_{k+1} + X_{k+1}\, B\, R^{-1} B^T\, X_k \\ &= 2s\, (X_{k+1} - X_k) + (X_k - X_{k+1})\, B\, R^{-1} B^T (X_{k+1} - X_k) \\ &\leq 0 \text{ since } X_{k+1} \leq X_k. \end{aligned}$$

Hence,

$$(D^T + X_{k+1}\, B\, R^{-1} B^T)\, (X_{k+2} - X_{k+1}) + (X_{k+2} - X_{k+1})\, (D + B\, R^{-1} B^T X_{k+1}) \leq 0,$$

that gives $(X_{k+2}-X_{k+1}) \leq 0$ since $[I_n \otimes (D^T + X_{k+1}\, B\, R^{-1} B^T) + (D + B\, R^{-1} B^T\, X_{k+1})^T \otimes I_n]$ is a nonsingular $M$-matrix.

Therefore the sequence $(X_k)_{k\geq 1}$ defined in (6.18) is a decreasing sequence of psd nonnegative matrices. Hence

$$\lim_{k\to +\infty} X_k = P_+ \geq 0 \text{ and } P_+ \text{ is a psd matrix.}$$

Thus, taking the limit in (6.18) gives (5.4). Hence the matrix $P_+ \geq 0$, psd matrix, is the solution of the ARE. Moreover, by assumption $(H_3')$, there exists $s > s(A)$ such that $-A + B\, R^{-1} B^T X_1$ is a $Z$-matrix and so it is for $-A + B\, R^{-1} B^T X_k$ by the decreasing of the sequence $(X_k)_{k\geq 1}$. It is the same for $-A + B\, R^{-1} B^T P_+$ by taking the limit. That is equivalently $A - B\, R^{-1} B^T P_+$ is a Metzler matrix and the closed-loop system is positive.

$\square$

**Remark 6.1.6** *It could be interesting to consider $A + B\, K$ instead of $D = s\, I_n - A$ in the iterative scheme (6.18) and to use a matrix $K$ such that $A + B\, K$ is a stable Metzler matrix. Such matrix $K$ exists by the assumption of positive stabilizability. Therefore, using the iterative scheme, we can obtain, as previously, a solution to the positive $\boldsymbol{LQ}_+^\infty$ problem.*

# 6.2 Hamiltonian approach

## 6.2.1 Using scalar products

In this section, the positivity condition on the closed-loop matrix is reinterpreted, first in terms of the solution of the ARE (given by (5.4)) and then in terms of the Hamiltonian matrix $H$ (defined by (3.7)). In [AS03] and [DL04], characterizations of monotone systems with scalar products are described, where LTI positive systems are in particular monotone systems. Using these characterizations, we obtain the following result :

**Theorem 6.2.1** *Consider the $\boldsymbol{LQ}^\infty$ problem (5.1)-(5.2). The $\boldsymbol{LQ}^\infty$ closed-loop system is positive if and only if*

$$\forall\, x,\, \tilde{x} \in I\!R^n \text{ such that } \begin{cases} x \geq 0 \\ P_+\, \tilde{x} \geq 0 \\ (P_+\, \tilde{x})^T x = 0, \end{cases} \qquad \tilde{x}^T (A^T P_+ + C^T C)\, x \leq 0$$

$$\text{i.e. } \left( \begin{bmatrix} A \\ C \end{bmatrix} \tilde{x} \right)^T \begin{bmatrix} P_+ \\ C \end{bmatrix} x \leq 0.$$

*Therefore the solution of the $\boldsymbol{LQ}^\infty$ problem is solution of the $\boldsymbol{LQ}_+^\infty$ problem.*

**Proof :** Using the characterization of positive (monotone) systems with scalar products developed in [AS03] and [DL04], the positivity of the closed-loop system can be rewritten as :

$$\begin{array}{ll}
 & A - B\,R^{-1}B^T P_+ \ \text{Metzler matrix} \\
\Leftrightarrow & \forall\, x,\, y \geq 0 \ \text{such that}\ y^T x = 0, \quad y^T(A - B\,R^{-1}B^T P_+)\,x \geq 0 \\
\Leftrightarrow & \forall\, x,\, \tilde{x} \in \mathbb{R}^n \ \text{such that} \left\{ \begin{array}{l} x \geq 0 \\ P_+\,\tilde{x} \geq 0 \\ (P_+\,\tilde{x})^T x = 0, \end{array} \right. \\
 & (P_+\,\tilde{x})^T A\,x - (P_+\,\tilde{x})^T B\,R^{-1}B^T P_+\,x \geq 0 \\
\Leftrightarrow & \forall\, x,\, \tilde{x} \in \mathbb{R}^n \ \text{such that} \left\{ \begin{array}{l} x \geq 0 \\ P_+\,\tilde{x} \geq 0 \\ (P_+\,\tilde{x})^T x = 0, \end{array} \right. \\
 & \tilde{x}^T(A^T P_+ + C^T C)\,x \leq 0 \quad \text{(by applying the ARE)} \\
 & \text{i.e.} \ \left( \begin{bmatrix} A \\ C \end{bmatrix} \tilde{x} \right)^T \begin{bmatrix} P_+ \\ C \end{bmatrix} x \leq 0.
\end{array}$$

□

Then by using the Hamiltonian matrix

$$H = \begin{bmatrix} A & -B\,R^{-1}B^T \\ -C^T C & -A^T \end{bmatrix}$$

we obtain the following result :

**Theorem 6.2.2** *Consider the $\boldsymbol{LQ^\infty}$ problem (5.1)-(5.2). The $\boldsymbol{LQ^\infty}$ closed-loop system is positive if and only if*

$$\forall\, x,\, \tilde{x} \in \mathbb{R}^n \ \textit{such that} \left\{ \begin{array}{l} x \geq 0 \\ P_+\,\tilde{x} \geq 0 \\ (P_+\,\tilde{x})^T x = 0, \end{array} \right. \quad \tilde{x}^T[P_+ \quad I_n]\,H \begin{bmatrix} I_n \\ P_+ \end{bmatrix} x \geq 0. \tag{6.20}$$

*Therefore the solution of the $\boldsymbol{LQ^\infty}$ problem is solution of the $\boldsymbol{LQ_+^\infty}$ problem.*

## 6.2.2 Illustrations

In order to illustrate Theorem 6.2.2, let us consider system (6.14) with $Q$ and $R$ given by (6.15). The solution of the ARE and the Hamiltonian matrix are given by

$$P_+ = \begin{bmatrix} 1.1731 & 0.5669 \\ 0.5669 & 1.5879 \end{bmatrix} \ \text{and} \ H = \begin{bmatrix} -2 & 1 & -0.1923 & -0.0769 \\ 1 & -2 & -0.0769 & -0.2308 \\ -4 & -0.01 & 2 & -1 \\ -0.01 & -6 & -1 & 2 \end{bmatrix}.$$

Let us denote

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \ \text{and} \ \tilde{x} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix}$$

and observe that

- $P_+\,\tilde{x} = \begin{bmatrix} 1.1731\,\tilde{x}_1 + 0.5669\,\tilde{x}_2 \\ 0.5669\,\tilde{x}_1 + 1.5879\,\tilde{x}_2 \end{bmatrix}$;
- $(P_+\,\tilde{x})^T x = x_1\,(1.1731\,\tilde{x}_1 + 0.5669\,\tilde{x}_2) + x_2\,(0.5669\,\tilde{x}_1 + 1.5879\,\tilde{x}_2)$ ;
- $\tilde{x}^T [P_+ \quad I_n]\, H \begin{bmatrix} I_n \\ P_+ \end{bmatrix} x = x_1\,(-5.0827\,\tilde{x}_1 - 5.694\,\tilde{x}_2) + x_2\,(-4.5906\,\tilde{x}_1 - 11.879\,\tilde{x}_2)$.

Solving $\begin{cases} x \geq 0 \\ P_+\,\tilde{x} \geq 0 \\ (P_+\,\tilde{x})^T x = 0, \end{cases}$ in (6.20) by means of MAPLE gives three solutions :

1. $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\tilde{x} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix}$ such that
$$\tilde{x}_1 = -\tilde{x}_2\,\frac{5669\,x_1 + 15879\,x_2}{11731\,x_1 + 5669\,x_2}$$
with $x_1,\ x_2,\ \tilde{x}_2 \in \mathbb{R}$.
2. $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\tilde{x} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix}$ such that
$$\begin{cases} 1.1731\tilde{x}_1 + 0.5669\tilde{x}_2 & \geq & 0 \\ 0.5669\tilde{x}_1 + 1.5879\tilde{x}_2 & \geq & 0 \end{cases}$$
i.e. $\tilde{x}_1 \geq -0.4832\,\tilde{x}_2$ with $\tilde{x}_2 \in \mathbb{R}$.
3. $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\tilde{x} = \begin{bmatrix} \tilde{x}_1 \\ 0 \end{bmatrix}$ where $\tilde{x}_1 \geq 0$.

Then, for these three solutions $x,\ \tilde{x} \in \mathbb{R}^n$, one has
$$\tilde{x}^T [P_+ \quad I_n]\, H \begin{bmatrix} I_n \\ P_+ \end{bmatrix} x \geq 0$$
such that condition (6.20) is verified. Therefore, the conclusion of Theorem 6.2.2 holds, as was to be expected.

### 6.2.3 Using graphs

Now some additional notations are used to rewrite this condition more briefly. Let $G(X) := \{\, z = \begin{bmatrix} I_n \\ P_+ \end{bmatrix} x \ :\ x \in \mathbb{R}^n \}$, the *graph* of $P_+$. Let $C_+ = \mathbb{R}^n_+$ the nonnegative orthant of $\mathbb{R}^n$. Let
$$\begin{array}{rcccl} P_{C_+} & : & C_+ & \rightarrow & \mathbb{R}^n \\ & & x & \rightsquigarrow & P_{C_+} x := P_+\, x \end{array}$$
be the restriction of $P_+$ on the cone $C_+$. And finally let $L := \begin{bmatrix} 0_n & I_n \\ I_n & 0_n \end{bmatrix}$ a permutation matrix. With these notations, we obtain the following result :

**Theorem 6.2.3** *Consider the* $\boldsymbol{LQ^\infty}$ *problem (5.1)-(5.2). The* $\boldsymbol{LQ^\infty}$ *closed-loop system is positive if and only if*

$$\begin{array}{c} \forall\, z \in G(P_{C_+}),\, \forall\, \tilde{z} \in G(P_+) \cap (I\!\!R^n \times C_+), \\ (L\,\tilde{z})^T\, z = 0 \;\Rightarrow\; (L\,\tilde{z})^T H\, z \geq 0. \end{array} \tag{6.21}$$

*Therefore the solution of the* $\boldsymbol{LQ^\infty}$ *problem is solution of the* $\boldsymbol{LQ_+^\infty}$ *problem.*

**Proof :** Condition (6.21) is equivalent to, with $z = \begin{bmatrix} I_n \\ P_+ \end{bmatrix} x,\ x \geq 0$ and $\tilde{z} = \begin{bmatrix} I_n \\ P_+ \end{bmatrix} \tilde{x}$, $P_+\,\tilde{x} \geq\ 0$,

$$\tilde{x}^T [P_+ \quad I_n]\, H \begin{bmatrix} I_n \\ P_+ \end{bmatrix} x \geq 0$$

since

$$L\,\tilde{z} = \begin{bmatrix} 0_n & I_n \\ I_n & 0_n \end{bmatrix} \begin{bmatrix} I_n \\ P_+ \end{bmatrix} \tilde{x} = \begin{bmatrix} P_+ \\ I_n \end{bmatrix} \tilde{x}$$

and

$$(L\,\tilde{z})^T z = \left( \begin{bmatrix} P_+ \\ I_n \end{bmatrix} \tilde{x} \right)^T \begin{bmatrix} I_n \\ P_+ \end{bmatrix} x = \begin{bmatrix} \tilde{x}^T P_+ x \\ \tilde{x}^T P_+ x \end{bmatrix}.$$

□

Finally, we can reinterpret this condition in terms of basis. Consider $\mathcal{L}^-(H)$ the $H$-invariant subspace spanned by the (generalized) eigenvectors associated to the stable eigenvalues, i.e. eigenvalues with negative real parts. Let us construct a basis of this subspace :

$$\mathcal{Z} = \begin{bmatrix} \vdots & & \vdots \\ z_1 & \cdots & z_n \\ \vdots & & \vdots \end{bmatrix},\ z_i \in \mathrm{I\!R}^{2n}.$$

Consider the following decomposition $z_i = \begin{bmatrix} u_i \\ \cdots \\ v_i \end{bmatrix},\ u_i,\, v_i \in \mathrm{I\!R}^n$.

Therefore $\mathcal{Z} = \begin{bmatrix} U \\ V \end{bmatrix}$ et Im$\mathcal{Z}$ = $\mathcal{L}^-(H)$. Whence $P_+ = V\,U^{-1}$ and $\mathcal{Z}(\mathrm{I\!R}^n) = G(P_+)$. In fact, $\mathcal{Z}$ can be considered as a basis of $\mathcal{L}^-(H)$ but $\mathcal{Z}\,V$ also, with $V$ a nonsingular matrix. Then there exist an infinity of choices of $U$ and $V$ such that $P_+ = V\,U^{-1}$. We obtain therefore the following condition :

**Theorem 6.2.4** *Consider the* $\boldsymbol{LQ^\infty}$ *problem (5.1)-(5.2). The* $\boldsymbol{LQ^\infty}$ *closed-loop system is positive if and only if*

$$\forall\, z = \begin{bmatrix} U \\ V \end{bmatrix} v,\ v \in I\!\!R^n \ \textit{ such that } U\,v \geq 0,$$

$$\forall\, \tilde{z} = \begin{bmatrix} U \\ V \end{bmatrix} \tilde{v},\ \tilde{v} \in I\!\!R^n \ \textit{ such that } V\tilde{v} \geq 0,$$

$$(U\,\tilde{v})^T (V\,v) = -(V\,\tilde{v})^T (U\,v) \;\Rightarrow\; 2\,(U\,\tilde{v})^T \left(-C^T C\,(U\,v) - A^T (V\,v)\right) \geq 0.$$

*Therefore the solution of the* $\boldsymbol{LQ^\infty}$ *problem is solution of the* $\boldsymbol{LQ_+^\infty}$ *problem.*

**Proof :** With the previous considerations, condition (6.21) can be reinterpreted in terms of $U$ and $V$ :

- $z \in G(P_{C_+}) \;\Leftrightarrow\; z = \begin{bmatrix} I_n \\ P_+ \end{bmatrix} x = \begin{bmatrix} I_n \\ V\,U^{-1} \end{bmatrix} x = \begin{bmatrix} U \\ V \end{bmatrix} v$ with $x := U\,v \geq 0$ ;

- $\tilde z \in G(P_+) \cap (\mathbb{R}^n \times C_+) \;\Leftrightarrow\; \tilde z = \begin{bmatrix} I_n \\ P_+ \end{bmatrix} \tilde x$ such that $P_+ \tilde x = V\,U^{-1}\tilde x \geq 0$. Then with $\tilde x := U\,\tilde v$ that gives $\tilde z = \begin{bmatrix} U \\ V \end{bmatrix} \tilde v$, where $\tilde v \in \mathbb{R}^n$ such that $V\,\tilde v \geq 0$ ;

- $$\begin{aligned} (L\,\tilde z)^T z &= \left(\begin{bmatrix} V \\ U \end{bmatrix} \tilde v\right)^T \begin{bmatrix} U \\ V \end{bmatrix} v \\ &= \tilde v^T \begin{bmatrix} V^T & U^T \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} v \\ &= \tilde v^T \left(V^T U + U^T V\right) v \\ &= (V\,\tilde v)^T (U\,v) + (U \tilde v)^T (V\,v)\,; \end{aligned}$$

- $$\begin{aligned} (L\,\tilde z)^T H\,z &= \left(\begin{bmatrix} V \\ U \end{bmatrix} \tilde v\right)^T \begin{bmatrix} A & -B\,R^{-1}B^T \\ -C^T C & -A^T \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} v \\ &= \tilde v^T \begin{bmatrix} V^T & U^T \end{bmatrix} \begin{bmatrix} A\,U - B\,R^{-1}B^T V \\ -C^T C\,U - A^T V \end{bmatrix} v \\ &= \tilde v^T \left(V^T A\,U - V^T B\,R^{-1}B^T V - U^T C^T C\,U - U^T A^T V\right) v \\ &= \tilde v^T U^T \left((V\,U^{-1})^T A - (V\,U^{-1})^T B\,R^{-1}B^T V\,U^{-1} - C^T C - A^T V\,U^{-1}\right) U\,v \\ &= \tilde v^T U^T \left(P_+ A - P_+ B\,R^{-1}B^T P_+ - C^T C - A^T P_+\right) U\,v \\ &= \tilde v^T 2\,U^T \left(-C^T C - A^T P_+\right) U\,v \\ &= 2\,\tilde v^T \left(-U^T C^T C\,U - U^T A^T V\right) v \\ &= 2\,(U\,\tilde v)^T \left(-C^T C\,(U\,v) - A^T (V\,v)\right). \end{aligned}$$

□

## 6.3 Diagonal solution for the ARE

As for the $\boldsymbol{LQ_+^{t_f}}$ problem in finite horizon, a diagonal solution for the ARE is considered, in order to keep the Metzler property of the closed-loop matrix.

**Lemma 6.3.1** *Consider the* $\boldsymbol{LQ^\infty}$ *problem (5.1)-(5.2). If the solution* $P_+$ *of the ARE is such that,* $-B\,R^{-1}B^T P_+$ *is a diagonal matrix, then the* $\boldsymbol{LQ^\infty}$ *closed-loop system is positive and therefore the solution of the* $\boldsymbol{LQ^\infty}$ *problem is solution of the* $\boldsymbol{LQ_+^\infty}$ *problem.*

By the analysis on the Hamiltonian matrix, one can obtain the following result :

**Theorem 6.3.2** *Assume that we can find a basis* $\begin{bmatrix} U \\ V \end{bmatrix}$ *of* $\mathcal{L}^-(H)$ *such that*

$$B\,R^{-1}B^T V = D\,U$$

*where $D$ is a diagonal matrix. Then the $\boldsymbol{LQ^\infty}$ closed-loop system is positive and therefore the solution of the $\boldsymbol{LQ^\infty}$ problem is solution of the $\boldsymbol{LQ_+^\infty}$ problem.*

One easy way to get the condition of Lemma 6.3.1 is to impose notably that $P_+$ is a diagonal matrix, provided that the matrix $B\,R^{-1}B^T$ be also diagonal. The following result gives sufficient conditions for achieving this goal ; it is a generalization of the finite horizon case, see Remark 4.2.5 b).

**Theorem 6.3.3** *Consider the $\boldsymbol{LQ^\infty}$ problem (5.1)-(5.2) where $B$ is equal to $I_n$. Choose a constant $\alpha$ such that $\alpha > \max\{0,\, \lambda_F\}$. Define*

$$\mathcal{A}_\alpha := A\,A^T - (\alpha\, I_n + A)\,(\alpha\, I_n + A)^T.$$

*Assume that*

$$\begin{cases} S &= \alpha\, I_n, \\ R &= r\, I_m, \\ C^T C &= \alpha^2\,(\frac{1}{r} + 1)\, I_n + \mathcal{A}_\alpha, \end{cases}$$

*with $r > 0$ such that*

$$\frac{1}{r} > -\frac{\lambda_{\min}(\mathcal{A}_\alpha)}{\alpha^2} - 1, \tag{6.22}$$

*where $\lambda_{\min} := \min\{\lambda \,:\, \lambda \in \sigma(\mathcal{A}_\alpha)\}$.*
*Then the solution of ARE is given by $P_+ = \alpha\, I_n$ and the $\boldsymbol{LQ^\infty}$ closed-loop system is positive and therefore the solution of the $\boldsymbol{LQ^\infty}$ problem is solution of the $\boldsymbol{LQ_+^\infty}$ problem.*

**Proof :** The proof follows the lines of the finite horizon case and is therefore omitted. □

# Chapter 7

# The Inverse Input/State-Invariant LQ Problem

In this chapter, the inverse input/state-invariant LQ problem is studied. First, the standard inverse LQ problem is stated. This problem consists of, for a fixed matrix $K$, determining weighting matrices $Q$ and $R$ such that the control $u = K\,x$ is optimal for the resulting LQ problem, see e.g. [Loc01] and [CD91]. The problem is solved by means of linear matrix inequalities, see e.g. [BEFB94] and [SW05]. Then, the inverse state-invariant LQ problem is studied by means of the computation of an invariant stabilizing feedback $K$ such that the resulting control is optimal for the corresponding LQ problem. The resolution of this problem leads to linear and bilinear matrix inequalities (BMI), see e.g. [VB00]. Bilinear matrix inequalities were popularized by Safonov and co-workers in a series of proceedings papers, see e.g. [SGL94]. In particular, the inverse positive LQ problem is solved by using LMIs. Finally, the inverse input/state-invariant LQ problem is described and is solved by LMI and also BMI.

## 7.1 The inverse standard LQ problem

### 7.1.1 Problem statement

Recall that the standard $\boldsymbol{LQ^\infty}$ problem consists of, for fixed weighting matrices $Q$ and $R$, finding the control $u = K\,x$ which minimizes the cost defined by these weighting matrices. The inverse LQ problem, denoted by $\boldsymbol{LQ^\infty_{\text{inv}}}$, is a reciprocal approach to the $\boldsymbol{LQ^\infty}$ problem. It consists of determining weighting matrices $Q$ and $R$ such that the control $u$ given by a fixed state-feedback $K$ is solution of the corresponding LQ problem. Furthermore, as in the $\boldsymbol{LQ^\infty}$ problem, the assumptions of stabilizability and detectability for the given system have to hold to obtain a solution, these assumptions will also hold for the inverse problem, see [AM90, Section 5.6] for a discussion on the general inverse optimal control problem.

The $\boldsymbol{LQ^\infty}$ problem can be stated as follows : given a system $[A,\,B]$ such that $(A,\,B)$ is stabilizable. Let $K$ be a fixed matrix in $\mathbb{R}^{m\times n}$. The inverse standard $\boldsymbol{LQ^\infty_{\text{inv}}}$ problem consists of determining symmetric matrices $Q = C^TC$ and $R$, (with $Q$ positive semidefinite and $R$ positive definite respectively) such that

1. the pair $(Q,\, A)$ is detectable ;
2. the control $u = K\,x$ is optimal for the corresponding LQ problem, i.e. minimizes the following quadratic cost

$$J(x_0,\, u) = \frac{1}{2}\left(\int_0^\infty (\|R^{1/2}u(t)\|^2 + \|C\,x(t)\|^2)\,\mathrm{dt}\right). \tag{7.1}$$

### 7.1.2 Matrix inequalities approach

In this section, the $\boldsymbol{LQ}^{\infty}_{\mathrm{inv}}$ problem is stated by using linear matrix inequalities, see e.g. [BEFB94].

**Definition 7.1.1**
*a) A linear matrix inequality (LMI) is an inequation of the form*

$$F(x) \prec 0 \;\; (\textit{ or } F(x) \succ 0)$$

*where $F$ is an affine function, from $\mathcal{X}$ (a linear space of finite dimension) to $\mathcal{S}$ (the set of symmetric matrices) and where $\prec 0$ ($\succ 0$) means "negative definite" ("positive definite", respectively).*

*b) A finite set of LMIs, $F_1(x) \prec 0,\ F_2(x) \prec 0, \ldots,\ F_n(x) \prec 0$, can be written as a single LMI*

$$F(x) = \begin{pmatrix} F_1(x) & 0 & \cdots & 0 \\ 0 & F_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_n(x) \end{pmatrix} \prec 0.$$

In the sequel, the numerical implementation of LMIs is done with YALMIP, which is a modeling language for advanced modeling and solution of convex and nonconvex optimization problems. It is implemented as a free toolbox for MATLAB. See e.g. [Yal] for details. Furthermore, this tool allows to mix LMIs of different types. Thus one can create a LMI containing different LMIs of type $\prec 0$, $\succ 0$, $\preceq$ or $\succeq 0$, as for example the following single LMI :

$$F(x) = \begin{pmatrix} F_1(x) \prec 0 & 0 & \cdots & 0 \\ 0 & F_2(x) \succ 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_n(x) \preceq 0 \end{pmatrix}.$$

Note that equation $F(x) = 0$ can be translated in two LMIs : $F(x) \preceq 0$ and $F(x) \succeq 0$.
The conditions to guarantee the existence of a solution to the $\boldsymbol{LQ}^{\infty}_{\mathrm{inv}}$ problem can be written with LMIs. However the analysis to transform these conditions into a single LMI of the same type will not be performed since it is not necessary for the numerical solving (see [Jac09] for

details). So consider conditions (7.1.1) and rewrite them in a matrix form. First the condition of detectability of $(Q,\, A)$ can be translated as, see [BEFB94, Section 10.6],

$$\textit{there exists a positive definite matrix } P_1 \textit{ such that } A^T P_1 + P_1 A \prec Q \tag{7.2}$$

where $X \prec Y$ means that $X$ is less than $Y$ in the order of symmetric matrices. Now the fact that the control $u = K\,x$ is optimal for the $\boldsymbol{LQ^\infty}$ problem means that $u(t) = K\,x(t) = -R^{-1}B^T P_+ x(t)$ that is $K = -R^{-1}B^T P_+$ or equivalently

$$B^T P_+ + R\,K = 0. \tag{7.3}$$

Condition (7.3) requires the computation of the unique stabilizing positive semidefinite solution $P_+$ of ARE. Then with $A_+ := A + B\,K = A - B\,R^{-1}B^T P_+$, i.e $A = A_+ + B\,R^{-1}B^T P_+$, the algebraic Riccati equation (5.4) becomes

$$\begin{aligned}
& (A_+ + B\,R^{-1}B^T P_+)^T P_+ + P_+\,(A_+ + B\,R^{-1}B^T P_+) - P_+ B\,R^{-1}B^T P_+ + Q = 0 \\
\Leftrightarrow\ & A_+^T P_+ + P_+ A_+ + P_+ B\,R^{-1}B^T P_+ + Q = 0 \\
\Leftrightarrow\ & (A + B\,K)^T P_+ + P_+(A + B\,K) + \underbrace{P_+ B\,R^{-1}}_{-K^T} R\, \underbrace{R^{-1}B^T P_+}_{-K} + Q = 0.
\end{aligned}$$

or equivalently

$$(A + B\,K)^T P_+ + P_+(A + B\,K) + K^T R\,K + Q = 0 \tag{7.4}$$

Hence, by considering conditions (7.2)-(7.4), solving the $\boldsymbol{LQ^\infty_{\text{inv}}}$ problem is equivalent to the resolution of the following set of LMIs :

**LMI 2 :**

$$\boxed{\begin{gathered}
R \succ 0 \\
Q \succeq 0 \\
P_+ \succeq 0 \\
P_1 \succ 0 \\
(A + B\,K)^T P_+ + P_+(A + B\,K) + K^T R\,K + Q = 0 \\
B^T P_+ + R\,K = 0 \\
A^T P_1 + P_1\,A \prec Q
\end{gathered}} \tag{7.5}$$

In the sequel, this set of LMIs is called “LMI 2”. In the following sections, we study the inverse input and state-invariant LQ problems. For solving these problems, we compute first a matrix $K$ in order to obtain the invariance (of the state and/or the input) of the closed-loop system before determining matrices $Q$ and $R$ such that the resulting control is optimal (given by LMI 2).

## 7.2 The inverse state-invariant LQ problem

### 7.2.1 Problem statement and matrix inequalities approach

Consider an invariant stabilizable system $[A,\, B]$, see Definition 1.3.2. Let $K$ be an invariant stabilizing feedback, i.e. a matrix $K$ such that $A + B\,K$ is a stable Metzler matrix and $(A + B\,K)\,\bar{x} \geq 0$, see Theorem 1.3.3. The inverse state-invariant LQ problem, which is denoted by $\boldsymbol{LQ}_{\bar{\boldsymbol{x}}}^{\text{inv}}$, consists of determining symmetric matrices $Q = C^T C$ and $R$, such that $(Q,\, A)$ is detectable and such that the control $u = K\,x$ is optimal for the corresponding $\boldsymbol{LQ}^{\boldsymbol{\infty}}$ problem. The difference with the $\boldsymbol{LQ}_{\text{inv}}^{\boldsymbol{\infty}}$ problem is the determination of an invariant stabilizing feedback matrix $K$. This step can be stated under the form of matrix inequalities by using Theorem 1.3.4 which gives a characterization of invariant stabilizability by using a Lyapunov equation :

**BMI 1a :**
$$\boxed{\begin{array}{c} \text{diag}[(A + B\,K)_{ij}]_{i\neq j} \succeq 0 \\ \text{diag}[(A + B\,K)\,\bar{x}] \succeq 0 \\ P \succ 0 \\ P\,(A + B\,K)^T + (A + B\,K)\,P \prec 0 \end{array}} \tag{7.6}$$

Then the $\boldsymbol{LQ}_{\bar{\boldsymbol{x}}}^{\text{inv}}$ problem is solved first by computing a matrix $K$ which is solution of (7.6) and next by solving LMI 2 ; see (7.5), which gives weighting matrices $Q$ and $R$ such that $u = K\,x$ is optimal for the corresponding $\boldsymbol{LQ}^{\boldsymbol{\infty}}$ problem. However, in (7.6), the inequality $P\,(A + B\,K)^T + (A + B\,K)\,P \prec 0$ is bilinear in $P$ and $K$, see e.g. [VB00]. This bilinear matrix inequality (BMI) is not easy to handle as it is written but the solver YALMIP is able to compute a solution, see Section 7.4 for numerical examples. Moreover, in the positive case, where $\bar{x} = 0$, the parameterization of the matrix $K$ as $K = Y\,P^{-1}$ allows us to achieve the following LMI, which is much easier to handle :

$$P\,A^T + Y^T B^T + A\,P + B\,Y \prec 0,$$

see the following subsection. The second step of the methodology consists of finding $Q$ and $R$ such that $K$ is an LQ-optimal feedback. The usefulness of the LQ problem is notably the stabilization of the closed-loop system together with robustness, see e.g. [AM90, Section 5.3] and also Chapter 9 where we allude to these properties.

### 7.2.2 The inverse positive LQ problem

Consider the particular case where $\bar{x} = 0$ and define the inverse positive LQ problem, which is denoted by $\boldsymbol{LQ}_{+}^{\text{inv}}$, as follows : let a positively stabilizable system $[A,\, B]$. Let $K$ be a fixed matrix such that $A + B\,K$ is a stable Metzler matrix. The inverse positive $\boldsymbol{LQ}_{+}^{\text{inv}}$ problem consists of determining symmetric matrices $Q = C^T C$ and $R$ such that $(Q,\, A)$ is detectable and the control $u = K\,x$ is optimal for the corresponding LQ problem. Using Theorem 2.3.5 which gives LMI characterizations of the determination of the matrix $K$, the $\boldsymbol{LQ}_{+}^{\text{inv}}$ can be summarized as the resolution of the two following LMIs, see [Jac09] ; the first one for the computation of the matrix $K$, which can be written as follows :

**LMI 1a :**

$$\begin{array}{c} \mathrm{diag}[(A\,P + B\,Y)_{ij}]_{i\neq j} \succeq 0 \\ P \succ 0 \\ P\,A + Y^T B^T + A\,P + B\,Y \prec 0 \end{array} \tag{7.7}$$

where $P$ is a diagonal matrix in this case, see Theorem 2.3.5, and the second LMI for the computation of $Q$ and $R$ for an optimal solution of $\boldsymbol{LQ^{\infty}}$, that is LMI 2.

## 7.3 The inverse input/state-invariant LQ problem

Let us define the inverse input/state-invariant LQ problem, which is denoted by $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$, as follows : consider a stabilizable system $[A,\,B]$. Our aim is to compute a matrix $K$ such that the corresponding state $x(t)$ is such that $x(t) \geq \bar{x}$ and the control $u(t) = K\,x(t)$ is such that $u(t) \geq \bar{u}$. Such matrix $K$ exists by Theorem 1.1.13 (which gives necessary and sufficient conditions for the input/state-invariance) and Corollary 1.1.15 (which gives sufficient conditions). The inverse input/state-invariant $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem consists of determining symmetric matrices $Q = C^T C$ and $R$ such that $(Q,\,A)$ is detectable and the control $u = K\,x$ is optimal for the corresponding LQ problem, given by LMI 2. First, by using Theorem 1.1.13, the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ can be summarized as the resolution of the following BMI (before solving of LMI 2) :

**BMI 1b :**

$$\begin{array}{c} \mathrm{diag}[(\mathcal{H})_{ij}]_{i\neq j} \succeq 0 \\ \begin{bmatrix} -I_n \\ -K \end{bmatrix} (A + B\,K) - \mathcal{H} \begin{bmatrix} -I_n \\ -K \end{bmatrix} = \begin{bmatrix} 0_{n\times n} \\ 0_{m\times n} \end{bmatrix} \\ \mathrm{diag}\left[\mathcal{H} \begin{bmatrix} -\bar{x} \\ -\bar{u} \end{bmatrix}\right] \leq 0_{(m+n)\times 1} \end{array} \tag{7.8}$$

Now, using the sufficient conditions of Corollary 1.1.15, gives the following way for the resolution of the first part of the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem :

**LMI 1b :**

$$\begin{array}{c} \mathrm{diag}[(A + B\,K)_{ij}]_{i\neq j} \succeq 0 \\ \mathrm{diag}[(A + B\,K)\,\bar{x}] \succeq 0 \\ K\,A \geq 0 \\ \mathrm{diag}[(K\,B)_{ij}]_{i\neq j} \succeq 0 \\ \mathrm{diag}[K\,(A\,\bar{x} + B\,\bar{u})] \succeq 0 \end{array} \tag{7.9}$$

**Remarks 7.3.1**

*a) Assume that for all* $x(0) = x_0 \geq \bar{x},\ x(t) \geq \bar{x}$ *for all* $t \geq 0$*, i.e.* $A + B\,K$ *is a Metzler matrix such that* $(A + B\,K)\,\bar{x} \geq 0$*. Now assume that* $K$ *is a nonnegative matrix. Then*

$$\begin{array}{rcl} u = K\,x & \geq & K\,\bar{x} \geq \bar{u}, \\ \textit{whence}\quad K\,\bar{x} - \bar{u} & \geq & 0. \end{array}$$

*Then another alternative for LMI 1 is the following :*

**LMI 1c :**
$$\begin{array}{c} \mathrm{diag}[(A+B\,K)_{ij}]_{i\neq j} \succeq 0 \\ \mathrm{diag}[(A+B\,K)\,\bar{x}] \succeq 0 \\ K \geq 0 \\ \mathrm{diag}[K\,\bar{x}-\bar{u}] \succeq 0 \end{array} \tag{7.10}$$

*b) Assume that for all $x(0) = x_0 \geq \bar{x},\ x(t) \geq \bar{x}$ for all $t \geq 0$, i.e. $A + B\,K$ is a Metzler matrix such that $(A + B\,K)\,\bar{x} \geq 0$. If $u(t) = K\,x(t) \geq \bar{u}$ for all time $t$, then the condition is also satisfied for the initial time, i.e. $K\,x_0 - \bar{u} \geq 0$. This condition can be seen as giving a suitable initial condition for the input trajectories. Then we can reasonably hope that this starting boost will be sufficient to guarantee that $u(t) = K\,x(t) \geq \bar{u}$ for all larger times. This necessary condition implies the following alternative for LMI 1 :*

**LMI 1d :**
$$\begin{array}{c} \mathrm{diag}[(A+B\,K)_{ij}]_{i\neq j} \succeq 0 \\ \mathrm{diag}[(A+B\,K)\,\bar{x}] \succeq 0 \\ \mathrm{diag}[K\,x_0-\bar{u}] \succeq 0 \end{array} \tag{7.11}$$

# 7.4 Numerical examples

In this section, the inverse positive $\boldsymbol{LQ}_{+}^{\mathrm{inv}}$ problem and the inverse input/state-invariant $\boldsymbol{LQ}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}^{\mathrm{inv}}$ problem are illustrated by solving the different LMIs or BMIs introduced in the previous sections.

## 7.4.1 The inverse positive LQ problem

Consider the stable nonpositive system described by

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} u(t), \tag{7.12}$$

with the initial condition $x_0 = [0.1 \quad 0.1]^T \geq 0$, under the constraints

$$\forall\, x_0 \geq 0,\ \forall\, t \geq 0, \quad \begin{cases} x_1(t) \geq 0 \\ x_2(t) \geq 0 \end{cases}$$

The open-loop state trajectories are drawn in Figure 7.1, for the initial states $x_0 = [1 \quad 0]^T$ (graphs on the left) and $x_0 = [0 \quad 1]^T$ (graphs on the right) respectively, i.e. the columns of $e^{A\,t}$ at the sampling times.

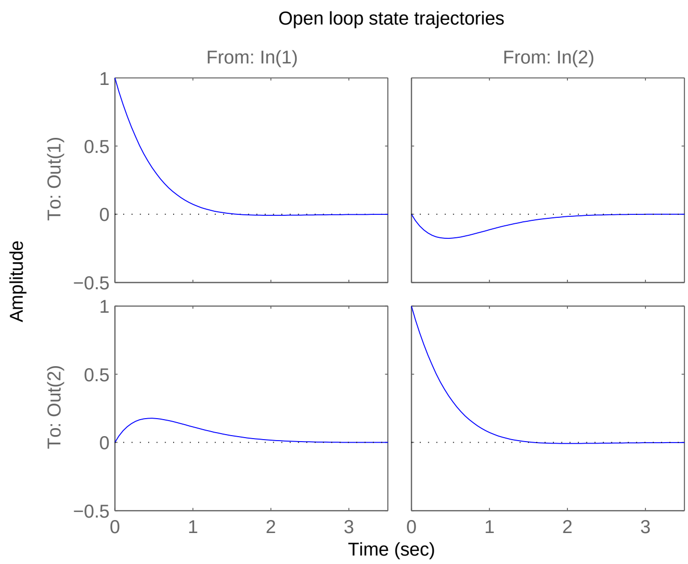



Figure 7.1: Open-loop state trajectories $x(t)$ for system (7.12).

Solving LMI 1a given by (7.7) gives :

$$P = \begin{bmatrix} 35.106 & 0 \\ 0 & 35.106 \end{bmatrix}$$

and

$$Y = \begin{bmatrix} 52.659 & 42.127 \\ -28.085 & 52.659 \end{bmatrix}$$

and therefore, we obtain :

$$K = Y\,P^{-1} = \begin{bmatrix} 1.5 & 1.2 \\ -0.8 & 1.5 \end{bmatrix}$$

and

$$A + B\,K = \begin{bmatrix} -0.5 & 0.2 \\ 0.2 & -0.5 \end{bmatrix}.$$

Unfortunately, when solving LMI 2, it is not possible to find matrices $Q$ and $R$ such that $K$ gives an LQ-optimal control. Indeed, LMI 1a delivers only one matrix $K$, among many others, which may not be admissible for LMI 2. An iterative process, which is an heuristic approach, is introduced to compute another state-feedback $K$. This iterative process is summarized in Table 7.1, where the maximum number of iterations is fixed to 100. This heuristic approach has been used on other numerical examples in [Jac09].

```
1. Init :
   • Compute P_0 and Y_0 by solving LMI 1a
   • Let K_0 = Y_0 P_0^{-1}
   • Solve LMI 2 with K_0 in order to obtain weighting matrices Q_0 and R_0
     such that condition (7.5) holds for R ← R_0 and Q ← Q_0 for a fixed
     tolerance ε.

2. If LMI 2 has a solution (Q_0, R_0)
   Then
        u_0 = K_0 x is the optimal control corresponding to (Q_0, R_0) → STOP.
   Else
        Let i = 1.
        While i < 100, Do :
             ⋆ Compute P_i and Y_i by solving LMI 1a
               with the additional condition P_i ≺ P_{i-1}
             ⋆ Let K_i = Y_i P_i^{-1}
             ⋆ Solve LMI 2 with K_i in order to obtain weighting matrices
               Q_i and R_i such that condition (7.5) holds for R ← R_i
               and Q ← Q_i for a fixed tolerance ε.
             If LMI 2 has a solution (Q_i, R_i)
             Then
                  u_i = K_i x is the optimal control corresponding to (Q_i, R_i)
                  → STOP.
             Else
                  Let i = i + 1.
             End
        End
   End
```

Table 7.1: Heuristic iterative process

Now, back to the numerical example (7.12). After solving the LMI 1a for the first time, seven iterations are needed in the iterative process (Table 7.1 with $\varepsilon = 10^{-8}$) to obtain weighting matrices $(Q,\, R)$ such that condition (7.5) holds and $u = K\,x$ is the optimal control of the corresponding LQ problem. Therefore, we obtain :

$$K = \begin{bmatrix} -0.86229 & 2.1449 \\ 0.14491 & -0.86229 \end{bmatrix} \tag{7.13}$$

and

$$Q = \begin{bmatrix} 0.62003 & -6.5301 \\ -6.5301 & 74.6 \end{bmatrix} \text{ and } R = \begin{bmatrix} 2.7673 & 9.2843 \\ 9.2843 & 40.959 \end{bmatrix}. \tag{7.14}$$

The optimal control is depicted in Figure 7.2. In addition, Figure 7.3 gives a comparison of the state trajectories before the iterative process and the LQ-optimal state trajectories after the iterative process. As for Figure 7.1, this figure represents the state trajectories, for the initial states $x_0 = [1 \quad 0]^T$ (graphs on the left) and $x_0 = [0 \quad 1]^T$ (graphs on the right) respectively. We can observe that the state trajectories are both nonnegative but the LQ-optimal state trajectories converge faster towards zero. Finally, we have also checked that solving the standard $\boldsymbol{LQ^\infty}$ problem with $Q$ and $R$ given by (7.14) leads to the matrix $K$ given by (7.13).

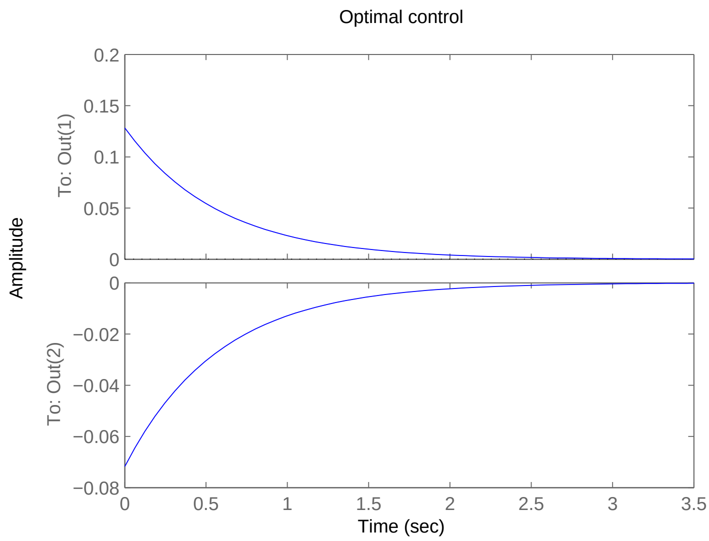



Figure 7.2: Optimal control $u(t)$ for system (7.12) with (7.13) and (7.14).

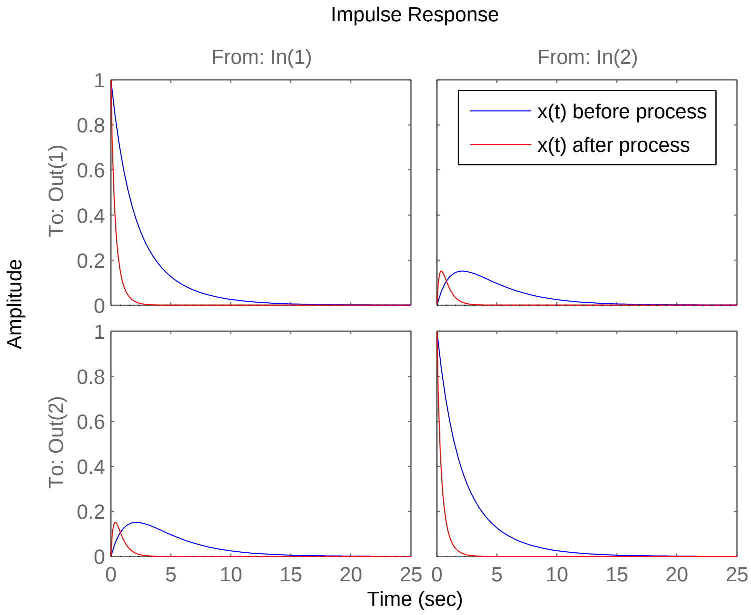



Figure 7.3: State trajectories $x(t)$ for system (7.12) before and after the iterative process.

### 7.4.2 The inverse input/state-invariant LQ problem

Consider system (7.12) under the following constraints

$$\forall\, x_0 \geq \bar{x},\; \forall\, t \geq 0, \quad \begin{cases} x(t) \geq \bar{x} \\ u(t) \geq \bar{u} \end{cases} \tag{7.15}$$

where $\bar{x} = [-1 \quad -1]^T$, $\bar{u} = [-0.2 \quad -0.2]^T$ and $x_0 = [0.1 \quad 0.1]^T$. Solving BMI 1b, given by (7.8), with YALMIP in MATLAB, gives the following results :

$$H = \begin{bmatrix} -3.9320 & 0.046815 & 0.26920 & 0.1991 \\ 2.3942 & -4.2286 & 0.048942 & 0.19385 \\ 0.055613 & 0.89403 & -4.9881 & 0.010056 \\ 0 & 0 & 3.6223 & -3.6223 \end{bmatrix}$$

and

$$K = \begin{bmatrix} -2.2090 & 0.69072 \\ 1.5953 & -2.7225 \end{bmatrix}. \tag{7.16}$$

Then solving LMI 2 gives directly $Q$ and $R$,

$$Q = \begin{bmatrix} 2648.4 & -423 \\ -423 & 1197.4 \end{bmatrix} \text{ and } R = \begin{bmatrix} 202.12 & 31.776 \\ 31.776 & 77.286 \end{bmatrix}, \tag{7.17}$$

such that $K$ is optimal for the standard $\boldsymbol{LQ^\infty}$ problem ; it is not needed to go through the iterative process. In addition, one can verify that solving the standard $\boldsymbol{LQ^\infty}$ problem with $Q$ and $R$ gives the state feedback $K$ given by (7.16). The optimal state trajectories and the optimal control are drawn in Figures 7.4 and 7.5 respectively. One can observe that the constraints (7.15) are numerically verified.

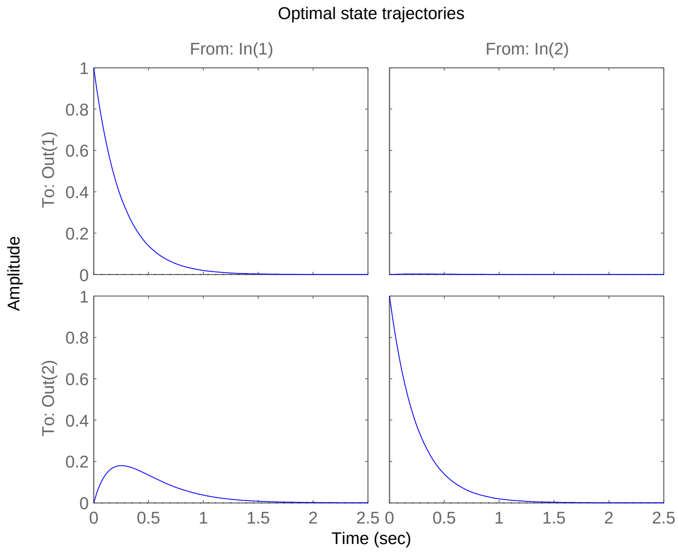



Figure 7.4: Optimal state trajectories $x(t)$ for system (7.12) with (7.16) and (7.17).

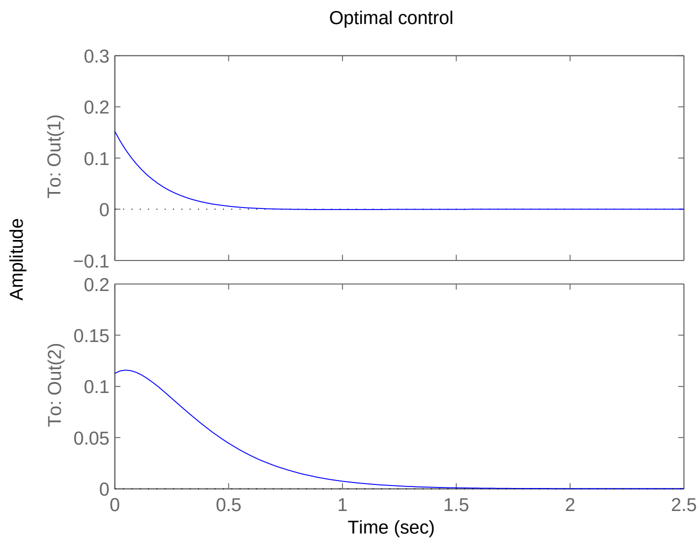



Figure 7.5: Optimal control $u(t)$ for system (7.12) with (7.16) and (7.17).

# Part III

# Application To Invariant Nonlinear Systems

This last part is devoted to the application of the LQ problem to locally positively invariant nonlinear systems. First, properties of locally positively invariant nonlinear systems are described in Chapter 8. A linear approximation of such nonlinear systems around an equilibrium is also studied, see e.g. [CBHB09, Kha02]. Next, in Chapter 9, the problem of coexistence of species, which are in competition for a single nutrient in a chemostat, is studied, see [SW95] for an overview on the chemostat model. The theory developed so far for the input/state-invariant LQ problem is applied to guarantee the local positive invariance of the chemostat model. The idea is to ensure the input/state-invariance of its linearized system around an equilibrium, by applying an appropriate LQ-optimal control (given either by the solution of an input/state-invariant $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem together with a receding horizon approach, or by the solution of an inverse input/state-invariant $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem obtained by solving LMIs and BMIs).

# Chapter 8

# Locally Positively Input/State-Invariant Nonlinear Systems

This chapter, devoted to locally positively input/state-invariant nonlinear systems, is an interlude between the theory developed so far and the next chapter devoted to the application (on the chemostat model) which is described by a nonlinear system. In this application, the objective is to guarantee the stability of the model and also a property of local positive input/state-invariance (see [CBHB09] and the references therein for the study of nonnegative solutions of a nonlinear system, applied there to kinetic equations). In this chapter, we first recall the classical notions of stability of an equilibrium of a nonlinear system, see e.g. [Kha02]. Then the concept of local positive invariance around an equilibrium is developed. Finally, conditions for the stability and the local positive input/state-invariance of a nonlinear system are established in terms of the stability and the input/state-invariance of the linearized system.

Notice that the notion of local positive nonlinear time-varying linear systems is introduced in [Ka 03]. There, the local positiveness of nonlinear systems implies the nonnegativity of the state trajectories in a neighborhood of an equilibrium. Here it implies the strict positivity of the state and the input trajectories. Moreover, the methodology is different here. First, the local positive input/state-invariance of the nonlinear system is studied by using the linearization of the system around an equilibrium. Then the design of a state feedback of the linearized system is studied such that it guarantees the local positive input/state-invariance of the resulting nonlinear closed-loop system.

## 8.1 Stability of nonlinear systems

Consider the following nonlinear system :

$$\dot{x} = f(x), \quad x(0) = x_0 \tag{8.1}$$

where $f(x)$ is a continuously differentiable function, (which guarantees the existence and uniqueness of the solution of (8.1), see e.g. [Kha02, Section 2.2]). Assume that there exists an equilibrium $x_e$ for the system (8.1), i.e. such that $f(x_e) = 0$. The stability of a nonlinear system is stated in terms of stability of its equilibrium, see e.g. [Kha02, Section 3.1] and [CBHB09] :

**Definition 8.1.1** *The equilibrium $x_e$ of system (8.1) is said to be*

- **(Lyapunov) stable** *if, for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that*

$$\|x(0) - x_e\| < \delta \;\Rightarrow\; \|x(t) - x_e\| < \varepsilon,\, \forall\, t \geq 0.$$

- **asymptotically stable** *if it is stable and $\delta$ can be chosen such that*

$$\|x(0) - x_e\| < \delta \;\Rightarrow\; \lim_{t\to\infty} x(t) = x_e.$$

The Lyapunov stability of an equilibrium means that solutions starting close enough to the equilibrium (within a distance $\delta$ from it) remain close enough to it forever (within a distance $\varepsilon$ from it). Note that this must be true for any $\varepsilon$ that one may choose. Asymptotic stability means that solutions that start close enough to the equilibrium not only remain close enough to it but also eventually converge to the equilibrium.

Now in a small neighborhood of the equilibrium $x_e$, the nonlinear system (8.1) can be approximated by a linear one, see e.g. [Kha02, Sections 3.3 and 11.2] and [Ka 03]. Consider $f(x) = A\,x + N_f(x)$ where

$$A = \frac{\partial f}{\partial x}(x_e),$$

is the Jacobian matrix of $f(x)$ at $x_e$, $N_f(x)$ is the nonlinear part of $f(x)$ and

$$\frac{\|N_f(x)\|}{\|x - x_e\|} \to 0 \;\text{ as }\; \|x - x_e\| \to 0.$$

Then the linearized system

$$\dot{\tilde{x}} = \frac{\partial f}{\partial x}(x_e)\,\tilde{x} = A\,\tilde{x} \tag{8.2}$$

where $\tilde{x} := x - x_e$ is called a linear approximation of the nonlinear system (8.1) in the neighborhood of $x_e$. The following theorem gives conditions under which the stability of the equilibrium of the nonlinear system can be investigated by the study of its stability as an equilibrium for the linearized system, see [Kha02, Theorem 3.7] :

**Theorem 8.1.1** *Let $x_e$ be an equilibrium for the nonlinear system (8.1). Let $A = \frac{\partial f}{\partial x}(x_e)$ be the Jacobian matrix of $f(x)$ at $x_e$. Then $x_e$ is asymptotically stable if $\mathcal{R}e\lambda_i < 0$ for all $\lambda_i$ eigenvalues of $A$.*

Theorem 8.1.1 states that the stability of the linear system (8.2) implies the asymptotic stability of the equilibrium $x_e$ of system (8.1).

## 8.2 Locally positively invariant nonlinear systems

In this section, the concept of locally positively input/state-invariant nonlinear system is defined. Consider the following nonlinear system :

$$\dot{x} = F(x,\, u) := f(x) + G(x)\, u, \quad x(0) = x_0 \tag{8.3}$$

where $G(x) = [g_1(x) \ldots g_m(x)] \in \mathbb{R}^{n\times m}$ and $f(\cdot)$ and $g_i(\cdot)$, $i = 1, \ldots, m$, are continuously differentiable functions. Assume that there exists an equilibrium $x_e$ corresponding to an input $u_e$ for system (8.3), i.e. such that $F(x_e,\, u_e) = f(x_e) + G(x_e)\, u_e = 0$. Assume, for the context of the application developed in Chapter 9, that $x_e \gg 0$ and $u_e \gg 0$.

Consider the linear approximation of system (8.3) in the neighborhood of $(x_e,\, u_e)$ :

$$\dot{x} = F(x,\, u) = A\, x + B\, u + N_F(x,\, u) \tag{8.4}$$

where

$$A = \frac{\partial F}{\partial x}(x_e,\, u_e) = \frac{\partial f}{\partial x}(x_e) + \frac{\partial G}{\partial x}(x_e)\, u_e,$$
$$B = \frac{\partial F}{\partial u}(x_e,\, u_e) = G(x_e),$$

and $N_F(x,\, u)$ is the nonlinear part of $F(x,\, u)$ such that

$$\frac{\|N_F(x,\, u)\|}{\|(x - x_e,\, u - u_e)\|} \to 0 \text{ as } \|(x - x_e,\, u - u_e)\| \to 0 \tag{8.5}$$

Then one has the following linearized system

$$\dot{\tilde{x}} = A\, \tilde{x} + B\, \tilde{u} \tag{8.6}$$

with $\tilde{x} := x_L - x_e$ and $\tilde{u} := u_L - u_e$ where $x_L$ and $u_L$ are called the (shifted) linearized state and input trajectories, respectively.

Now consider the linearization of (8.3) about $(x_e,\, u_e)$ which results in the linear system (8.6). Assume that the pair $(A,\, B)$ is stabilizable. Let us design a matrix $K$ such that all the eigenvalues of $A + B\, K$ have negative real parts (for stability of the linear closed-loop system). Let us apply the linear state feedback control $\tilde{u} = K\, \tilde{x}$ to the nonlinear system, i.e.

$$u = \tilde{u} + u_e = K\, \tilde{x} + u_e = K(x - x_e) + u_e = K\, x - (K\, x_e - u_e) = K\, x + v,$$

that is an affine feedback for the nonlinear system, which gives the following closed-loop system

$$\dot{x} = f(x) + G(x)\, u = f(x) + G(x)\, (K\, x - (K\, x_e - u_e)) = F(x, K\, x - (K\, x_e - u_e)) \tag{8.7}$$

Clearly, $(x_e,\, u_e)$ is an equilibrium of the closed-loop system. The linearization of system (8.7) about $(x_e,\, u_e)$ is given by :

$$\dot{\tilde{x}} = (A + B\, K)\, \tilde{x} \tag{8.8}$$

Since $K$ is such that $A + B\, K$ is stable, it follows by Theorem 8.1.1 that $(x_e,\, u_e)$ is an asymptotically stable equilibrium of the closed-loop system (8.7). In fact, this equilibrium is exponentially stable, see [Kha02, Theorem 3.11].

Now, let us define the concept of local positive input/state-invariance of a nonlinear system. This definition is inspired by the definition of (global) input/state-invariance of a linear system, see Definition 1.1.6.

**Definition 8.2.1** *System (8.7) is said to be* **locally positively input/state-invariant** *around the equilibrium* $(x_e,\, u_e)$ *where* $u_e = K\, x_e + v$ *if there exists a neighborhood* $V_e$ *of the equilibrium* $x_e$ *such that*

$$\forall\, x_0 \in V_e \;\textit{ such that }\; x_0 \gg 0,\; \forall\, t \geq 0, \qquad \begin{cases} x(t) \gg 0 \\ u(t) = K\, x(t) + v \gg 0 \end{cases}$$

The concept of state-invariance of nonlinear systems (the fact that state trajectories starting in a set will stay in this set for all future times) has been developed in [Kha02, Section 3.2] and [CBHB09]. Moreover, conditions for the local positiveness of nonlinear time-varying systems are established in [Ka 03].

Here, the aim is to find a linear feedback control law $K$ for the linearized system such that the resulting closed-loop nonlinear system is locally positively input/state-invariant and stable. The following theorem states that it suffices that a linear feedback $K$ for the linearized system be a stabilizing input/state-invariant feedback to guarantee the stability and the local positive input/state-invariance of the resulting nonlinear system. Recall the cone $C_{\bar{x},\bar{u}}$ used in the definition of the input/state-invariance for linear systems, see Definition 1.1.6 :

$$C_{\bar{x},\bar{u}} := \left\{ x \in \mathbb{R}^n \;:\; \begin{bmatrix} I_n \\ K \end{bmatrix} x \geq \begin{bmatrix} \bar{x} \\ \bar{u} \end{bmatrix} \right\}.$$

**Theorem 8.2.1** *If there exists a linear feedback control law* $K$ *such that the linearized closed-loop system (8.8) is stable (i.e.* $\mathcal{R}e\lambda < 0$ *for all* $\lambda \in \sigma(A + B\, K)$*) and input/state-invariant with respect to* $(x_e,\, u_e)$ *, i.e. such that*

$$\forall\, \tilde{x}_0 \in C_{\bar{x},\bar{u}},\; \forall\, t \geq 0, \quad \tilde{x}(t) \in C_{\bar{x},\bar{u}},$$

*where*

$$\bar{x} = -x_e + x_\varepsilon \;\textit{ and }\; \bar{u} = -u_e + u_\varepsilon$$

*with* $x_\varepsilon \gg 0$ *and* $u_\varepsilon \gg 0$*, and where* $\tilde{x}(t)$ *is the solution of system (8.8) and* $\tilde{u} = K\, \tilde{x}$*, then the resulting nonlinear closed-loop system (8.7) is locally positively input/state-invariant, i.e. there exists a neighborhood* $V_e$ *of the equilibrium* $x_e$ *such that*

$$\forall\, x_0 \in V_e \;\textit{ such that }\; x_0 \gg 0,\; \forall\, t \geq 0, \qquad \begin{cases} x(t) & \gg & 0 \\ u(t) & \gg & 0 \end{cases} \tag{8.9}$$

*where* $x(t)$ *is the solution of system (8.7) and* $u(t) = K\, x - (K\, x_e - u_e)$.

**Proof :** The fact that the closed-loop system (8.8) is stable implies that $x_e$ is an asymptotically stable equilibrium for the nonlinear system (8.7) (see Theorem 8.1.1), i.e.

$$\forall\, \epsilon > 0,\; \exists\, \delta > 0 \text{ such that } \|x(0) - x_e\| < \delta \Rightarrow \begin{cases} \|x(t) - x_e\| < \epsilon \\ \lim\limits_{t \to \infty} x(t) = x_e \end{cases} \quad \forall\, t \geq 0,$$

where $x(t)$ is the solution of system (8.7). Let us define $\mathcal{B}(x_e,\, \delta)$ a ball centered at $x_e$ of radius $\delta > 0$. Then, since for all $x_0 \in \mathcal{B}(x_e,\, \delta)$, $x(t) \to x_e$, with $x_e \gg 0$, it implies that $x(t) \gg 0$ for $t$ sufficiently large, that is :

$$\text{there exists } \; T > 0 \; \text{ such that for all } \; t > T,\; x(t) \gg 0. \tag{8.10}$$

Therefore, since $u(t) = K\, x(t) + v \to u_e$ and $u_e \gg 0$,

$$\text{there exists } \; T > 0 \; \text{ such that for all } \; t > T,\; u(t) \gg 0. \tag{8.11}$$

Hence $x(t) \gg 0$ and $u(t) \gg 0$ hold for $t$ sufficiently large. It remains to be shown that it also holds for $t \in [0,\, T]$ for all initial states in a sufficiently small neighborhood of $x_e$. First, let us show that the state trajectories $x(t)$ are strictly positive on $[0,\, T]$ for $x_0$ sufficiently close to the equilibrium $x_e$. Consider the linearized state $x_L(t)$ such that $x_L(t) = \tilde{x}(t) + x_e$ with, by assumption, $\tilde{x}(t) \geq -x_e + x_\varepsilon$ for all time $t$. Then, in particular, $x_L(t) \gg 0$ for all $t \in [0,\, T]$ since $x_\varepsilon \gg 0$. Let $z(t) = (x - x_L)(t)$ and computing $\dot{z}(t)$ with

$$\begin{aligned}
\dot{x} &= A\, x + B\, u + N_F(x,\, u) = A\, x + B\, K\, x - B\, K\, x_e + B\, u_e + \underbrace{N_F(x,\, K\, x + v)}_{:= N_F(x)} \\
\dot{x}_L &= \dot{\tilde{x}} = (A + B\, K)\, \tilde{x} = (A + B\, K)\, (x_L - x_e) = (A + B\, K)\, x_L - (A + B\, K)\, x_e
\end{aligned}$$

leads to

$$\begin{aligned}
\dot{z} &= \dot{x} - \dot{x}_L \\
&= A\, x + B\, K\, x - B\, K\, x_e + B\, u_e + N_F(x) - A\, x_L - B\, K\, x_L + A\, x_e + B\, K\, x_e \\
&= A\, (x - x_L) + B\, K\, (x - x_L) + A\, x_e + B\, u_e + N_F(x) \\
&= (A + B\, K)\, z + M_F(x,\, u_e)
\end{aligned}$$

where $M_F(x,\, u_e) = A\, x_e + B\, u_e + N_F(x)$ such that $M_F(x_e,\, u_e) = F(x_e,\, u_e) = 0$. Therefore, with $z(0) = x(0) - x_L(0) = x_0 - x_0 = 0$, for $t \in [0,\, T]$, it follows that

$$\begin{aligned}
z(t) &= \int_0^t e^{(A + B\, K)(t - \tau)} (A\, x_e + B\, u_e + N_F(x(\tau)))\, \mathrm{d}\tau\, ; \\
\|z(t)\| &\leq \int_0^t M\, e^{\sigma(t - \tau)} \|A\, x_e + B\, u_e + N_F(x(\tau))\|\, \mathrm{d}\tau, \quad \text{with } M > 0 \text{ and } \sigma < 0 \\
&\leq \underbrace{M\, e^{\sigma t}}_{\leq M} \int_0^T e^{-\sigma \tau} \|A\, x_e + B\, u_e + N_F(x(\tau))\|\, \mathrm{d}\tau.
\end{aligned}$$

Hence,

$$\begin{aligned}
\|z\|_\infty &:= \max_{t \in [0,\, T]} \|z(t)\| = \max_{t \in [0,\, T]} \max_{1 \leq i \leq n} |z_i(t)| \\
&\leq M \int_0^T e^{-\sigma \tau} \|A\, x_e + B\, u_e + N_F(x(\tau))\|\, \mathrm{d}\tau \\
&\leq M\, e^{-\sigma T} \int_0^T \max_{\tau \in [0,\, T]} \|A\, x_e + B\, u_e + N_F(x(\tau))\|\, \mathrm{d}\tau \\
&\leq M\, T\, e^{-\sigma T} \max_{\tau \in [0,\, T]} \|A\, x_e + B\, u_e + N_F(x(\tau))\|
\end{aligned}$$

where $\|A\,x_e + B\,u_e + N_F(x(\tau))\|$ tends uniformly to zero on $[0,\,T]$. Indeed, by the continuous dependence of the state trajectories with respect to the initial condition $x_0 \to x_e$ (see e.g. [Kha02, Theorem 2.6]), the state trajectory corresponding to any initial condition $x_0$ converges to the state trajectory corresponding to $x_e$ uniformly on any compact interval $[0,\,T]$ as $x_0$ tends to $x_e$. Then $x(\tau)$ converges uniformly to $x_e$ on $[0,\,T]$ and so $N_F(x(\tau))$ converges uniformly to $N_F(x_e)$. Therefore, as $x_0$ tends to $x_e$, $\max\limits_{\tau\in[0,\,T]} \|A\,x_e + B\,u_e + N_F(x(\tau))\|$ converges to $F(x_e,\,u_e) = 0$ on $[0,\,T]$. Let us denote $r := \min\limits_{1\le i\le n}[x_L]_i(t) > 0$. Then, there exists a neighborhood $W_e$ of the equilibrium $x_e$ (for example : a ball centered at $x_e$ of radius $\eta > 0$), such that, for all $x_0 \in \mathcal{B}(x_e,\,\eta)$,

$$0 \le \max_{\substack{1\le i\le n\\ t\in[0,\,T]}} |x_i(t) - [x_L]_i(t)| < r$$

that is,

$$\forall\, i,\ \forall\, t\in[0,\,T],\ -r < x_i(t) - [x_L]_i(t) < r$$

which implies that

$$\forall\, i,\ \forall\, t\in[0,\,T],\ x_i(t) > [x_L]_i(t) - r \ge 0$$

Therefore $x(t) \gg 0$ for all $t \in [0,\,T]$. Now consider the input trajectories $u(t)$ on $[0,\,T]$. By the uniform convergence of $x(t)$ to $x_e$ on $[0,\,T]$, $u(t) = K\,x(t) + v$, an affine function of $x$, will uniformly converge to $u_e$. Consider the linearized input trajectories $u_L(t)$ such that $u_L(t) = \tilde{u} + u_e$ where by assumption $\tilde{u} \ge u_e + u_\varepsilon$, with $u_\varepsilon \gg 0$, for all time $t \ge 0$. In particular on $[0,\,T]$, $u_L(t) \gg 0$. Let $\check{u}(t) = (u - u_L)(t)$ with $\tilde{u} = K\,\tilde{x}$,

$$u_L = \tilde{u} + u_e = K\,(x_L - x_e) + u_e = K\,x_L - K\,x_e + u_e$$

and

$$\begin{aligned}
\check{u} &= u - u_L\\
&= K\,x + v - K\,x_L + K\,x_e - u_e\\
&= K\,(x - x_L) + \underbrace{K\,x_e + v}_{=u_e} - u_e\\
&= K\,z
\end{aligned}$$

where $\|z\|_\infty$ tends to zero. So $\|\check{u}\|_\infty$ also converges to zero. Define $s := \min\limits_{1\le i\le m}[u_L]_i(t) > 0$. Therefore, there exists a neighborhood $Z_e$ of the equilibrium $x_e$ (e.g. $\mathcal{B}(x_e,\,\varepsilon)$, with $\varepsilon > 0$) such that, for all $x_0 \in \mathcal{B}(x_e,\,\varepsilon)$,

$$0 \le \max_{\substack{1\le i\le m\\ t\in[0,\,T]}} |u_i(t) - [u_L]_i(t)| < s$$

that is,

$$\forall\, i,\ \forall\, t\in[0,\,T],\ -s < u_i(t) - [u_L]_i(t) < s$$

and implies that

$$\forall\, i,\ \forall\, t\in[0,\,T],\ u_i(t) > [u_L]_i(t) - s \ge 0$$

Hence $u(t) \gg 0$ for all $t \in [0,\,T]$. Since we have shown that it also holds for $t$ sufficiently large, condition (8.9) holds for all time $t \ge 0$ on the ball $V_e = \mathcal{B}(x_e,\,\rho)$ where $\rho = \min\{\delta,\,\eta,\,\varepsilon\}$. $\square$

**Remark 8.2.1** *a) Applying a state feedback to the linearized system* $\tilde{u} = K\,\tilde{x}$*, leads to an invariant stabilizing feedback* $K$ *for the linearized system. Then the existence of this matrix* $K$ *gives the local positive input/state-invariance and the stability of the nonlinear system around the equilibrium. The principle of computing a state feedback for the linearized system such that the resulting nonlinear closed-loop system is locally positively input/state-invariant and stable is applied in Chapter 9. This chapter is devoted to the study of the coexistence of species in a chemostat model, which is described by a nonlinear system.*
*b) Note that Theorem 8.2.1 holds for any nonlinear system. There is no assumption of positivity on the open-loop nonlinear system (8.3). Now, if the open-loop system is positive, if the input trajectories are nonnegative for all time, this implies automatically that the state trajectories are nonnegative. However, it is not guaranteed that they are strictly positive. But the strict positivity of the state trajectories is essential in the application since the objective is the coexistence of species in a chemostat model. We will see in this application that the input/state-invariance of the linearized system is also paramount since, actually, the linearized system description of the chemostat model is not a positive system.*

# Chapter 9

# The Chemostat Model

The chemostat model is a perfectly mixed tank operated in continuous conditions and in which (bio)chemical reactions take place. The chemostat model may be used in particular to describe the interaction of microbial species which are competing for a single nutrient, see [SW95] for a detailed survey on this topic and see e.g. [BD90] for a survey on control of bioreactors. This model has also been used for different systems such as lakes, waste-water treatment processes and biological reactors producing genetically altered organisms.

A central result in microbial ecology theory is the *competitive exclusion principle* which states that the competition process yields at best a single winning species in the long run, see e.g. [SW95]. Yet, in nature, many species may coexist (see for example the paradox of the plankton in [Hut61]). This contradiction between the theory and the real world leads to modifications of the model in order to try to bring theory and practice in better accordance. There is a large literature devoted to modifying the chemostat model to ensure coexistence of the organisms. These studies are based on suitable manipulations of the two natural operating parameters, the dilution rate [BHW85, SFA79] or the input nutrient concentration [SFA79, Hsu80, Smi95, HS83, Smi81], that are taken to be time-varying rather than constant. Also, feedback control of the dilution rate has been used to allow coexistence in the chemostat [DS03, DS02]. Recently, in [RDH09], it is shown that the coexistence of multiple species, with growth functions close to each other, competing in a chemostat for a single resource, can occur in the long run. Finally, a design problem of a series of two chemostats is revisited in [RHM07] when more than one species are present for a single resource : they give conditions under which coexistence of two species is possible for such configurations.

Here, an LQ-optimal control is designed for the chemostat model with appropriate choice of the inputs, notably the input concentrations of the species. It is shown that in this case, coexistence of the species may occur. The theory of the input/state-invariant LQ problem (with direct approach (see Chapter 5) or inverse approach (see Chapter 7)) together with the properties of local positive invariance of nonlinear systems, developed in Chapter 8, are applied to the chemostat model in order to guarantee the coexistence of the species.

# 9.1 Description of the chemostat

## 9.1.1 Model description

The chemostat is a well-known model which is used to describe the interaction between microbial species which are competing for a single nutrient, see the scheme of a chemostat in Figure 9.1.

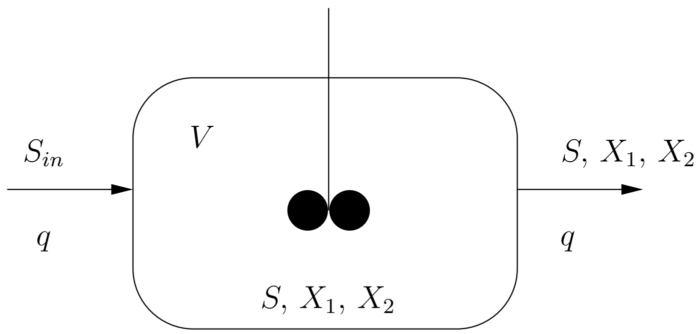



Figure 9.1: Scheme of a chemostat

It is a continuous stirred tank reactor with, for example in this figure, two species $X_1$ and $X_2$ growing on one limited substrate $S$. The basic assumption about the chemostat is that it is perfectly stirred, and, as a consequence, that each individual has an equal access to the resources. Consider a general model of a chemostat with $\tilde{n}$ species and a single resource :

$$\dot{x} = f(x) + G(x)\,u := F(x,\,u) \tag{9.1}$$

where $x = \begin{bmatrix} S \\ X_1 \\ \vdots \\ X_{\tilde{n}} \end{bmatrix},\ f = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{\tilde{n}} \end{bmatrix},\ u = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},1} \\ \vdots \\ X_{\text{in},\tilde{m}} \end{bmatrix},$

$f_0(x) = -D\,S - \sum_{i=1}^{\tilde{n}} \frac{\mu_i(S)\,X_i}{Y_i},\ \ f_i(x) = (\mu_i(S) - D)\,X_i,\ i = 1, \ldots, \tilde{n}$

$G(x) = B = \begin{bmatrix} D\,I_{m\times m} \\ 0_{(n-m)\times m} \end{bmatrix},\ \ n = 1 + \tilde{n},\ \ m = 1 + \tilde{m}$ with $\tilde{m} \le \tilde{n}$, and

| | | |
|---|---|---|
| $X_i$ | = | concentration of the species $i$ $(gr/l)$ |
| $S$ | = | concentration of the nutrient (substrate) $(gr/l)$ |
| $S_{\text{in}}$ | = | concentration of the nutrient in the input flow $(gr/l)$ |
| $X_{\text{in, i}}$ | = | concentration of the species $i$ in the input flow $(gr/l)$ |
| $D$ | = | dilution rate of the nutrient and the species $(1/t)$ |
| | = | $\frac{q}{V}$ where $q$ is the input flow rate and $V$ the volume of the tank |
| $Y_i$ | = | yield constant reflecting the conversion of nutrient to organism i.e. species $i$. (constant which can be taken to one by using a suitable choice of units) |

This constant can be determined (in batch culture) by measuring the ratio $\dfrac{\text{mass of the organism formed}}{\text{mass of the substrate used}}$ and hence is dimensionless.

$\mu_i(S)$ = the growth rate of the population $i$ $(1/t)$

where the functions $\mu_i(S)$ satisfy the following properties :

(H1) The function $S \mapsto \mu_i(S)$ is defined for all $S \geq 0$ and is differentiable.
(H2) $\mu_i(S) \geq 0$ and $\mu_i(0) = 0$.
(H3) The function $S \mapsto \mu_i(S)$ is increasing.

**Remark 9.1.1** *There exist several models for the definition of the growth rate function $\mu_i$, see [LH06], [BW85] and [SW95]. In the sequel, we consider the most common growth rate model, namely the Monod model (or Michaelis-Menten), which expresses the dependence of $\mu_i$ with respect to the substrate concentration $S$ as follows :*

$$\mu_i(S) = \mu_{\max,i} \frac{S}{K_{S,i} + S} \tag{9.2}$$

*where* $\mu_{\max,i}$ = *maximum growth rate of the population $i$ (when $S = \infty$) $(1/t)$*
$K_{S,i}$ = *half-saturation constant (or Michaelis-Menten constant), which represents the nutrient concentration such that the growth rate is half maximum (less than half its maximum), $(gr/l)$ also known as the affinity constant (of the substrate towards species)*

*See Figure 9.2 which gives the classical graph of a growth rate function.*

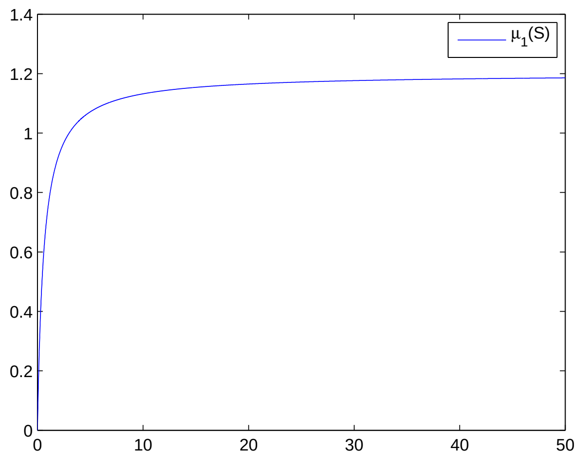



Figure 9.2: Growth rate function $\mu_i(S)$

### 9.1.2 The competitive exclusion principle

Consider system (9.1) with $Y_i = 1$, without loss of generality (by replacing $X_i$ by $X_i/Y_i$), and with $u = \begin{bmatrix} S_{\text{in}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ such that (9.1) becomes

$$\left\{ \begin{array}{rcl} \dot{S} & = & D\,(S_{\text{in}} - S) - \displaystyle\sum_{i=1}^{\tilde{n}} \mu_i(S)\,X_i \\ \\ \dot{X}_i & = & (\mu_i(S) - D)\,X_i \quad i = 1, \dots, \tilde{n} \end{array} \right. \tag{9.3}$$

Let us denote by $x = \begin{bmatrix} S \\ X \end{bmatrix}$ the solution of system (9.3) where $X = \begin{bmatrix} X_1 \\ \vdots \\ X_{\tilde{n}} \end{bmatrix}$ and let us define the set

$$\Omega := \left\{ \begin{bmatrix} S \\ X \end{bmatrix} \in \mathbb{R}^n \text{ such that } S \geq 0 \text{ and } X \gg 0 \right\}.$$

The *competitive exclusion principle (CEP)*, probably the most important result for chemostat models, is now stated. Assume that $\mu_i(S_{\text{in}}) > D$ otherwise it would imply the extinction of the $i$th organism even without competition. System (9.3) has $n$ equilibria, see [SW95, DS03] :

$$\left\{ \begin{array}{l} E_0 := (S_{\text{in}},\ 0, \dots,\ 0) \\ E_1 := (\lambda_1,\ (S_{\text{in}} - \lambda_1),\ 0, \dots,\ 0) \\ \vdots \\ E_{\tilde{n}} := (\lambda_{\tilde{n}},\ 0, \dots,\ (S_{\text{in}} - \lambda_{\tilde{n}})) \end{array} \right.$$

where the parameters $\lambda_i$, $i = 1, \dots, \tilde{n}$, called the *break-even concentrations*, are defined as follows :

$$\lambda_i(D) = \left\{ \begin{array}{ll} S & \text{such that } \mu_i(S) = D \\ +\infty & \text{if } \mu_i(S) < D \text{ for any } S \geq 0 \end{array} \right.$$

In fact, computing the equilibrium of second equation of system (9.3) gives

$$(\mu_1(S) - D)\,X_1 = 0$$

that is $X_1 = 0$ or $\mu_1(S) = D$, or equivalently for Monod's model, $S = \frac{D\,K_{S,1}}{\mu_{\max,1} - D} := \lambda_1$. In the same way, the third equation of system (9.3) gives $X_2 = 0$ or $\mu_2(S) = D$, i.e. $S = \frac{D\,K_{S,2}}{\mu_{\max,2} - D} := \lambda_2$ and it will be the same for the following equations in $X_i$. Therefore, by the first equation, it follows that :

- $S = S_{\text{in}}$ when $X_i = 0, \quad i = 1, \dots, \tilde{n}$ ;
- $X_1 = (S_{\text{in}} - \lambda_1)$ when $S = \lambda_1$ and $X_i = 0, \quad i = 2, \dots, \tilde{n}$ ;
- $X_2 = (S_{\text{in}} - \lambda_2)$ when $S = \lambda_2$ and $X_i = 0, \quad i = 1, \dots, \tilde{n},\ i \neq 2$ ;

  $\vdots$
- $X_{\tilde{n}} = (S_{\text{in}} - \lambda_{\tilde{n}})$ when $S = \lambda_{\tilde{n}}$ and $X_i = 0, \quad i = 1, \dots, \tilde{n} - 1.$

In the sequel, unless otherwise stated, the following values of the constants are used in the chemostat model, as in [LH06] :

| $D$ | $\mu_{\max,1}$ | $\mu_{\max,2}$ | $K_{S,1}$ | $K_{S,2}$ |
|---|---|---|---|---|
| $0.2$ | $1.2$ | $0.83$ | $0.6$ | $0.2$ |

Table 9.1: Values of the constants used in the chemostat model (9.5).

So, in general, only one species will win the competition and survive. In order to illustrate this fact, the behavior of two species in competition for one nutrient is illustrated with Figures 9.3 and 9.4, which represent, respectively, the growth curves compared to the dilution rate $D$ and the corresponding concentrations of the species $X_1$ and $X_2$. One can observe that the growth curve which crosses first the value of $D$ will imply that the corresponding species will win the competition. So one can say that the winner species is the one which has the best affinity with the nutrient or equivalently the smallest break-even concentration.

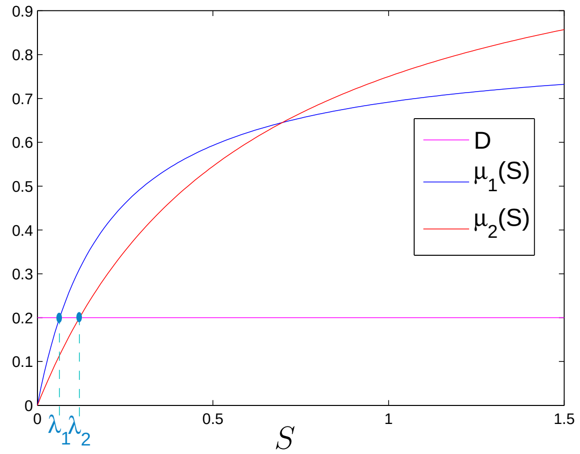



Figure 9.3: Growth rate functions $\mu_1(S)$ and $\mu_2(S)$ compared to $D$.

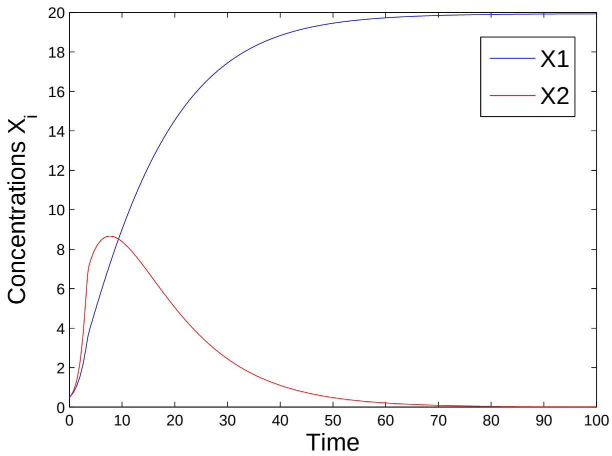



Figure 9.4: Trajectories of $X_1$ and $X_2$.

More formally, the competitive exclusion principle (CEP) can be stated as follows, see e.g. [SW95] :

**Theorem 9.1.1 (Competitive exclusion principle)** *Suppose that* $0 < \lambda_1 < S_{in}$ *and* $0 < \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_{\tilde{n}} \leq \infty$*. Then the equilibrium* $E_1$ *is a globally asymptotically stable equilibrium for system (9.3) with any initial condition* $x(0) \in \Omega$*. In other words, any solution of the system (9.3) with* $S(0) \geq 0$ *and* $X_i(0) > 0$ *satisfies*

$$\begin{cases} \lim\limits_{t\to\infty} S(t) & = & \lambda_1 \\ \lim\limits_{t\to\infty} X_1(t) & = & (S_{in} - \lambda_1) \\ \lim\limits_{t\to\infty} X_2(t) & = & 0 \\ & \vdots & \\ \lim\limits_{t\to\infty} X_{\tilde{n}} & = & 0. \end{cases}$$

The competitive exclusion principle states that, when several species are competing for the same substrate, only one of the species survives in the long run, see [SW95]. On the other hand, in nature, many species seem to coexist. An example of this fact is the paradox of the plankton to which many papers have been devoted. Notably the one of Hutchinson, see [Hut61], which observed that a great number of different species of planktons could survive on a very limited number of resources. This contradiction between theory and real world has triggered a lot of research aimed at bringing theory and practice in better accordance, see e.g. [BHW85, SFA79, Hsu80, Smi95] and the references therein. The aim of the following section is to find conditions such that the coexistence of the $\tilde{n}$ species is guaranteed.

## 9.2 The coexistence of species

### 9.2.1 Definition of coexistence

Let us define the concept of coexistence of species in a chemostat (inspired by [DAS06] and [RHM07]) as follows. Assume that $x(t)$ is the solution of system (9.1) with respect to the initial condition $x(0) = x_0 \in \Omega$ and the corresponding input $u \in \mathcal{U}$. We define the concept of coexistence w.r.t an admissible initial state $x_0$ as follows :

**Definition 9.2.1** *System (9.1) is said to be* **coexistent w.r.t.** $x_0 \in \Omega$ *if there exists an input* $u \in \mathcal{U}$ *such that* $\liminf_{t\to\infty} x(t) \in \Omega$.

If the coexistence holds for every admissible initial condition $x_0$, one gets the following concept of (global) coexistence :

**Definition 9.2.2** *System (9.1) is said to be* **(globally) coexistent** *if*

$$\forall\, x_0 \in \Omega,\ \exists\, u \in \mathcal{U},\ \ \textit{such that}\ \ \liminf_{t\to\infty} x(t) \in \Omega.$$

Now, let $(x_e,\, u_e)$ be an equilibrium for system (9.1), i.e. such that $F(x_e,\, u_e) = 0$.

**Definition 9.2.3** *System (9.1) is said to be* **locally coexistent** *around the equilibrium* $(x_e,\, u_e)$ *if there exists a neighbourhood* $V_e$ *of* $x_e$*, such that system (9.1) is coexistent w.r.t. every initial state* $x_0 \in \Omega \cap V_e$*, i.e.*

$$\forall\, x_0 \in \Omega \cap V_e,\ \exists\, u \in \mathcal{U},\ \ \textit{such that}\ \ \liminf_{t\to\infty} x(t) \in \Omega.$$

It is clear that coexistence implies local coexistence. Furthermore, in Chapter 8, we defined the concept of local positive input/state-invariance of a nonlinear system around an equilibrium. This concept leads to the local coexistence with in addition the positivity of the input. In fact, it is important to remark that the input also represents concentrations. So it is meaningful to guarantee the nonnegativity of the state and the input trajectories. Furthermore, the concept of local positive input/state-invariance system forces the strict positivity of the state and input trajectories and not only their nonnegativity. The following result obviously holds :

**Proposition 9.2.1** *Consider system (9.1). Assume that there exists a neighbourhood* $V_e$ *of the equilibrium* $x_e$ *such that*

$$\forall\, x_0 \in V_e\ \textit{such that}\ \ x_0 \gg 0,\ \forall\, t \geq 0,\quad \begin{cases} x(t) \geq x_\varepsilon \gg 0 \\ u(t) = K\,x(t) + v \gg 0, \end{cases}$$

*whence system (9.1) is locally positively input/state-invariant, in the sense of Definition 8.2.1. Then system (9.1) is locally coexistent with strictly positive input trajectories.*

**Proof :** Let $x_0 \in V_e$ such that $x_0 \gg 0$ and $u(t) = K\,x(t) + v \gg 0$. Since for all time $t \geq 0$, $x(t) \gg x_\varepsilon$, the following inequality holds : $\inf\limits_{\tau > t} x(\tau) \geq x_\varepsilon$. Observe that $t \mapsto \inf\limits_{\tau > t} x(\tau)$ is an increasing function. By taking its limit, with $x_\varepsilon \gg 0$, it follows that

$$\liminf_{t\to\infty} x(t) := \lim_{t\to\infty} \inf_{\tau>t} x(\tau) \geq x_\varepsilon \gg 0,$$

whence $\liminf\limits_{t\to\infty} x(t) \in \Omega$.

□

In the sequel, the term of coexistence is used for the species and the term of input/state-invariance is used for the system.

### 9.2.2 Problem statement

In this section, the coexistence problem for system (9.1) is stated in terms of optimal control in order to guarantee the coexistence of the species and more precisely the local positive input/state-invariance of system (9.1). We define the state $x(t)$ and the input $u(t)$ as follows :

$$x(t) = \begin{bmatrix} S \\ X_1 \\ \vdots \\ X_{\tilde{n}} \end{bmatrix} \text{ and } u(t) = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},1} \\ \vdots \\ X_{\text{in},\tilde{m}} \end{bmatrix}.$$

Assume that system (9.1) has an equilibrium

$$x_e := \begin{bmatrix} S_e \\ X_{1e} \\ \vdots \\ X_{\tilde{n}e} \end{bmatrix} \gg 0$$

corresponding to an input $u_e$ given by

$$u_e := \begin{bmatrix} S_{\text{in,e}} \\ X_{\text{in,1,e}} \\ \vdots \\ X_{\text{in},\tilde{m},\text{e}} \end{bmatrix} \gg 0.$$

Consider the linearized system, as studied in Chapter 8,

$$\dot{\tilde{x}} = \frac{\partial F}{\partial x}(x_e,\, u_e)\,\tilde{x} + \frac{\partial F}{\partial u}(x_e,\, u_e)\,\tilde{u} = A\,\tilde{x} + B\,\tilde{u} \tag{9.4}$$

where $\tilde{x} := x - x_e$ and $\tilde{u} := u - u_e$ and the Jacobian matrices $A$ and $B$ are given by :

$$
\begin{aligned}
A &= \begin{bmatrix} -D - \frac{\partial \mu_1(S_e)}{\partial S} X_{1e} - \frac{\partial \mu_2(S_e)}{\partial S} X_{2e} & -\mu_1(S_e) & \dots & -\mu_{\tilde{n}}(S_e) \\ \frac{\partial \mu_1(S_e)}{\partial S} X_{1e} & \mu_1(S_e) - D & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ \vdots & \vdots & & \vdots \\ \frac{\partial \mu_{\tilde{n}}(S_e)}{\partial S} X_{\tilde{n}e} & 0 & \dots & \mu_{\tilde{n}}(S_e) - D \end{bmatrix} \\
&= \begin{bmatrix} -D - \frac{\mu_{\max,1} K_{S,1}}{(K_{S,1}+S_e)^2} X_{1e} - \frac{\mu_{\max,2} K_{S,2}}{(K_{S,2}+S_e)^2} X_{2e} & -\frac{\mu_{\max,1} S_e}{K_{S,1}+S_e} & \dots & -\frac{\mu_{\max,\tilde{n}} S_e}{K_{S,\tilde{n}}+S_e} \\ \frac{\mu_{\max,1} K_{S,1}}{(K_{S,1}+S_e)^2} X_{1e} & \frac{\mu_{\max,1} S_e}{K_{S,1}+S_e} - D & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ \vdots & \vdots & & \vdots \\ \frac{\mu_{\max,\tilde{n}} K_{S,\tilde{n}}}{(K_{S,\tilde{n}}+S_e)^2} X_{\tilde{n}e} & 0 & \dots & \frac{\mu_{\max,\tilde{n}} S_e}{K_{S,\tilde{n}}+S_e} - D \end{bmatrix}
\end{aligned}
$$

$$
B = \begin{bmatrix} D\, I_{m \times m} \\ 0_{(n-m)\times m} \end{bmatrix}.
$$

Our first aim is to achieve the coexistence of system (9.1) with respect to an admissible fixed initial state $x_0$ by solving a finite horizon input/state-invariant $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\boldsymbol{\tilde{u}},\boldsymbol{\tilde{x}}}$ problem (as studied in Chapter 3) for the linearized system. This problem is solved as an optimization problem (using the function `quadprog` in MATLAB). It will be shown that the stabilization property of the LQ problem numerically guarantees the convergence of the linear trajectories $x_L = \tilde{x} + x_e$ ($u_L = \tilde{u} + u_e$) to the equilibrium $x_e$ ($u_e$, respectively), which are both strictly positive. Hence, with a receding horizon approach, i.e. with $t_f$ sufficiently large, we can numerically ensure the local coexistence w.r.t. $x_0$ for the resulting nonlinear system, in the sense of Definition 9.2.1. This method is developed below in Section 9.3.

The second aim is to find a feedback control law $K$ such that the resulting closed-loop system is locally positively input/state-invariant in the sense of Definition 8.2.1. One way to study the local positive input/state-invariance of the nonlinear system (9.1) is to consider its equilibrium $(x_e,\, u_e)$ where $x_e \gg 0$ and $u_e \gg 0$ and to linearize the system around this equilibrium (since at $(x_e,\, u_e)$, the input/state-invariance is ensured). Then one can stabilize the linearized system (9.4) around this equilibrium with an appropriate optimal control law $\tilde{u}$ such that the resulting linearized closed-loop system is input/state-invariant, i.e. such that (see Definition 1.1.6) :

$$
\forall\, t \geq 0,\ \forall\, \tilde{x}_0 \geq -x_e, \quad \begin{cases} \tilde{x}(t) \geq -x_e \\ \tilde{u}(t) \geq -u_e. \end{cases}
$$

As we have seen in Theorem 8.2.1, if there exists a linear feedback $K$ such that the linearized closed-loop system is input/state-invariant and stable, then the resulting nonlinear closed-loop system is locally positively input/state-invariant around its stable equilibrium $x_e$. Hence, by Proposition 9.2.1, with $x_\varepsilon$ sufficiently close to $x_e \gg 0$, the local positive input/state-invariance of the nonlinear closed-loop system ensures its local coexistence around the equilibrium $x_e$. Therefore, a stabilizing input/state-invariant feedback $K$ is computed for the linearized system (9.4), such that $\tilde{u} = K\tilde{x}$, by solving an inverse input/state-invariant $\boldsymbol{LQ}^{\mathrm{inv}}_{\boldsymbol{\tilde{u}},\boldsymbol{\tilde{x}}}$ problem with the aid of matrix inequalities (LMIs or BMIs), as studied in Chapter 7. Theorem 8.2.1 guarantees that the resulting closed-loop system is locally positively input/state-invariant and by Proposition 9.2.1, with $x_\varepsilon$ sufficiently close to $x_e$, the local coexistence of the closed-loop system is ensured. The results of this second method are developed in Section 9.4. Notice that in this method the optimal control of the linearized system is of state-feedback type whereas it is not the case in the first method. Furthermore, the first method is strongly inspired by the second one and by Theorem 8.2.1. These methods are illustrated by some numerical simulations.

Before developing these methods, we first describe precisely the problem of coexistence of two species competing for a single substrate. The calculation of the equilibria and the corresponding linearized systems are described. Finally, the behavior of trajectories is studied in an open-loop design with a constant input.

### 9.2.3 The coexistence of two species

**– <u>The chemostat model</u>**

Consider the competition between two species for one substrate. Assume that the growth rate functions $\mu_i(S)$ are given by Monod's model (9.2). Then system (9.3) reads

$$\begin{aligned} \begin{bmatrix} \dot{S} \\ \dot{X}_1 \\ \dot{X}_2 \end{bmatrix} &= \begin{bmatrix} -D\,S - \mu_1(S)\,X_1 - \mu_2(S)\,X_2 \\ (\mu_1(S) - D)\,X_1 \\ (\mu_2(S) - D)\,X_2 \end{bmatrix} + B\,u \\ \dot{x} &= f(x)\,x + B\,u \end{aligned} \tag{9.5}$$

where $B = P \begin{bmatrix} D\,I_{m\times m} \\ 0_{(n-m)\times m} \end{bmatrix}$ and $u = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},1} \\ X_{\text{in},2} \end{bmatrix}$ or a subvector with $n = 3$ ; $m \leq 3$ and $P$ is a permutation matrix. Thus the input vector $u$ can be chosen by three different ways :

- <u>Case 1</u> : $B = B_1 = \begin{bmatrix} D & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & D \end{bmatrix}$ ; $u = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},1} \\ X_{\text{in},2} \end{bmatrix}$ ;

- <u>Case 2</u> : $B = B_2 = \begin{bmatrix} D & 0 \\ 0 & D \\ 0 & 0 \end{bmatrix}$ ; $u = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},1} \end{bmatrix}$ ;

- <u>Case 3</u> : $B = B_3 = \begin{bmatrix} D & 0 \\ 0 & 0 \\ 0 & D \end{bmatrix}$ ; $u = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},2} \end{bmatrix}$.

For each case, the equilibrium $(x_e,\ u_e)$ is computed.

**– Computation of the equilibria**

Let us compute the equilibria of system (9.5) in the three cases depending on the choice of the input $u$.

CASE 1 : $u = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},1} \\ X_{\text{in},2} \end{bmatrix}$

The equilibrium equations in the first case are given by :

$$\begin{cases} D\,(S_{in,e} - S_e) - \mu_1(S_e)\,X_{1e} - \mu_2(S_e)\,X_{2e} & = & 0 \\ D\,X_{in1,e} + (\mu_1(S_e) - D)\,X_{1e} & = & 0 \\ D\,X_{in2,e} + (\mu_2(S_e) - D)\,X_{2e} & = & 0 \end{cases} \tag{9.6}$$

where $\mu_i(S_e) = \mu_{\max,i} \frac{S_e}{K_{S,i}+S_e}$ and

$$\begin{bmatrix} S_{in,e} \\ X_{in1,e} \\ X_{in2,e} \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \\ 5 \end{bmatrix} \tag{9.7}$$

is chosen such that $S_{in,e}$ is sufficiently large and such that the input concentration of $X_1$ is twice the one of $X_2$ since $X_2$ has the best affinity to win the competition, see Figures 9.3 and 9.4. Then, by using MAPLE to solve this system, we obtain the following equilibrium :

$$\begin{cases} S_e & = & 0.042533882 \\ X_{1e} & = & 16.58877203 \\ X_{2e} & = & 18.36869408. \end{cases}$$

In fact, this equilibrium can be found by isolating $X_1$ and $X_2$ in (9.6), which gives, after substitutions, the following expression depending only on $S_e$ :

$$4 - 0.2\,S_e + \frac{2.4\,S_e}{(0.6 + S_e)\,(\frac{1.2\,S_e}{0.6+S_e} - 0.2)} + \frac{0.83\,S_e}{(0.2 + S_e)\,(\frac{0.83\,S_e}{(0.2+S_e)} - 0.2)} = 0. \tag{9.8}$$

This expression gives a third order equation in $S_e$ which admits three roots ($S_e = 0.0425$, $0.0927$ and $38.6355$). This function of $S_e$ is drawn in Figure 9.5 according to $S_e$ and in comparison with $S_{\text{in}}$. First of all, one can observe that one root ($S_e = 38.6355$) is larger than $S_{in,e} = 20$ and is therefore not admissible. Then, Figure 9.6 is obtained by zooming on the transient part of the last figure in order to identify the two other roots. In this figure, the expressions of $X_{1e}$ and $X_{2e}$ according to $S_e$ are also depicted. This allows us to see that even if the two other roots of (9.8) are smaller than $S_{in,e}$, only one is admissible ($S_e = 0.0425$) since the root $S_e = 0.0927$ gives a negative value of $X_{2e}$.

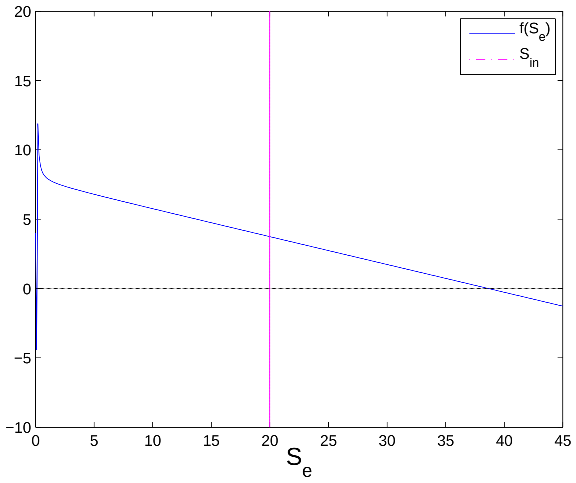



Figure 9.5: Expression of $S_e$ given by (9.8).

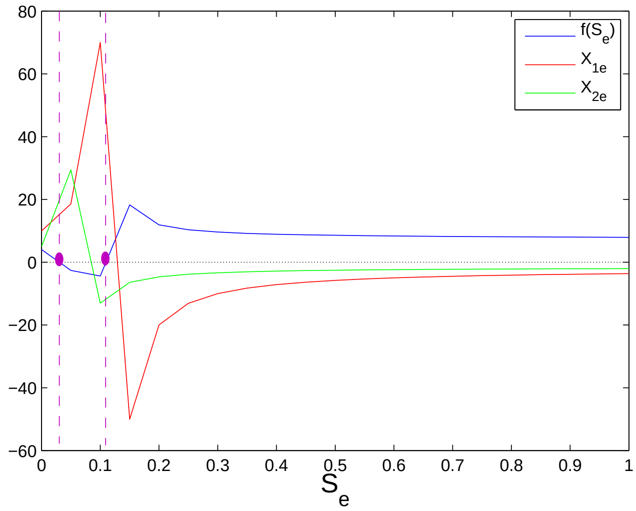



Figure 9.6: Expression of $S_e$ given by (9.8) in comparison with the expressions of $X_{1e}$ and $X_{2e}$ solving equation (9.6).

<u>CASE 2</u> : $u = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},1} \end{bmatrix}$

The equilibrium equations in the second case are given by :

$$\begin{cases} D\,(S_{in,e} - S_e) - \mu_1(S_e)\,X_{1e} - \mu_2(S_e)\,X_{2e} & = & 0 \\ D\,X_{in1,e} + (\mu_1(S_e) - D)\,X_{1e} & = & 0 \\ (\mu_2(S_e) - D)\,X_{2e} & = & 0 \end{cases}$$

Therefore $\mu_2(S_e) = D$, which gives, with Monod's model (9.2), $S_e = \dfrac{D\,K_{S,2}}{\mu_{\max,2} - D} = 0.0635.$

Let us fix arbitrarily

$$\begin{bmatrix} S_{in,e} \\ X_{in1,e} \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \end{bmatrix}, \tag{9.9}$$

which gives the following equilibrium :

$$\begin{cases} S_e & = & \dfrac{D\, K_{S,2}}{\mu_{\max,2} - D} = 0.0635 \\[2ex] X_{1e} & = & \dfrac{-D X_{\text{in},1}}{\mu_1 - D} = 23.48314607 \\[2ex] X_{2e} & = & \dfrac{D(S_{\text{in}} - S_e) - \mu_1 X_{1e}}{\mu_2} = 6.453361870 \end{cases}$$

where

$$\begin{cases} \mu_1(S_e) & = & \dfrac{\mu_{\max,1} S_e}{K_{S,1} + S_e} = 0.1148 \\ \mu_2(S_e) & = & \dfrac{\mu_{\max,2} S_e}{K_{S,2} + S_e} = D = 0.2 \end{cases}$$

<u>Case 3</u> : $u = \begin{bmatrix} S_{\text{in}} \\ X_{\text{in},2} \end{bmatrix}$

The equilibrium equations in the third case read :

$$\begin{cases} D\,(S_{in,e} - S_e) - \mu_1(S_e)\, X_{1e} - \mu_2(S_e)\, X_{2e} & = & 0 \\ (\mu_1(S_e) - D)\, X_{1e} & = & 0 \\ D\, X_{in2,e} + (\mu_2(S_e) - D)\, X_{2e} & = & 0 \end{cases}$$

Therefore $\mu_1(S_e) = D$, which gives, with Monod's model (9.2), $S_e = \dfrac{D\, K_{S,1}}{\mu_{\max,1} - D} = 0.12$. Let us fix arbitrarily $\begin{bmatrix} S_{in,e} \\ X_{in2,e} \end{bmatrix} = \begin{bmatrix} 20 \\ 5 \end{bmatrix}$, which gives the following equilibrium :

$$\begin{cases} S_e & = & \dfrac{D\, K_{S,1}}{\mu_{\max,1} - D} = 0.12 \\[2ex] X_{2e} & = & \dfrac{-D X_{\text{in},2}}{\mu_2 - D} = -8.988764045 \\[2ex] X_{1e} & = & \dfrac{D(S_{\text{in}} - S_e) - \mu_2 X_{2e}}{\mu_1} = 33.86876404 \end{cases}$$

Then it is impossible in this case, with the parameters given as in Table 9.1, to obtain an equilibrium $(x_e,\, u_e)$ such that $x_e \gg 0$. This fact is illustrated in Figure 9.7.

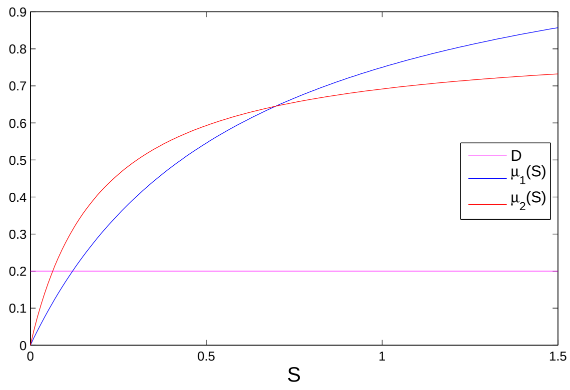



Figure 9.7: Growth rate functions $\mu_1(S)$ and $\mu_2(S)$ compared to $D$ for constants values of Table 9.1.

One can observe in this figure that the species $X_2$ has the best affinity. Then by the competitive exclusion principle, the species $X_2$ will win the competition if we consider the case $u = S_{\text{in}}$, see Figure 9.8.

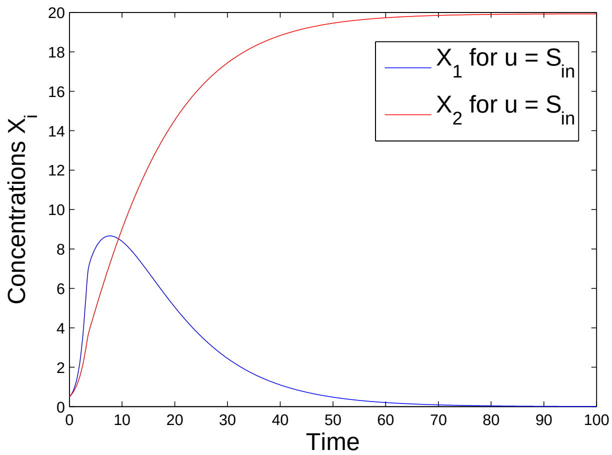



Figure 9.8: Trajectories of $X_1$ and $X_2$ for constants values of Table 9.1 with $u = S_{\text{in}}$.

By using the input $u = [S_{\text{in}} \;\; X_{\text{in},2}]^T$, the species $X_1$ does not receive any "external support" whereas $X_2$ receives some help and has initially the best affinity to win the competition. Therefore in this case, it is impossible to guarantee the coexistence of the two species, see Figure 9.9 which represents the concentrations $X_1$ and $X_2$, obtained by solving system (9.5) with a constant input $u = [S_{\text{in}} \;\; X_{\text{in},2}]^T$. One can observe that $X_1$ numerically tends to zero two times faster than in the previous case. Then this case is not interesting in the aim of coexistence of the species and that is why it is no more considered in the sequel.

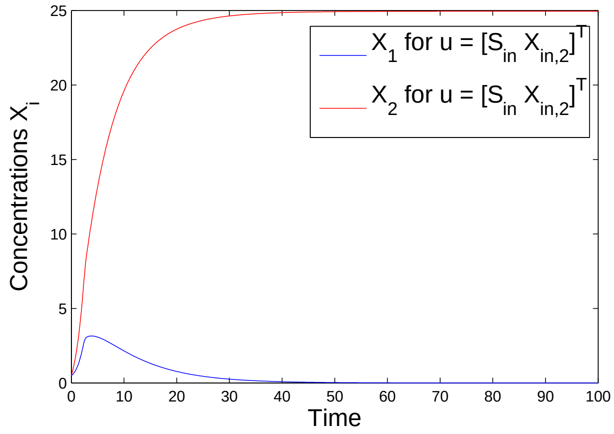



Figure 9.9: Trajectories of $X_1$ and $X_2$ for constants values of Table 9.1 in CASE 3.

Now, solving system (9.5) with a constant input $u$ given by (9.7) (CASE 1) and (9.9) (CASE 2) gives the following trajectories for the concentrations of the species over time $t$ in an open-loop design, see Figures 9.10 and 9.11. In these cases, coexistence of the species may occur. In CASE 1, $X_1$ and $X_2$ have similar values while in CASE 2, $X_1$ numerically tends to a larger value than $X_2$. Indeed, in this case, $X_1$ obtains initially some help at the expense of $X_2$. Nevertheless, the coexistence of the two species is ensured.

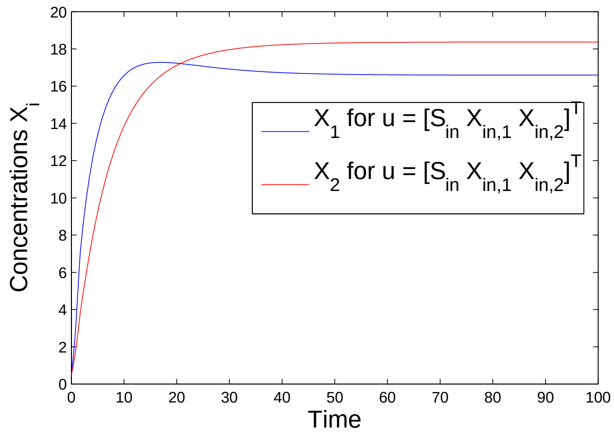



Figure 9.10: Trajectories of $X_1$ and $X_2$ for constants values of Table 9.1 in CASE 1.

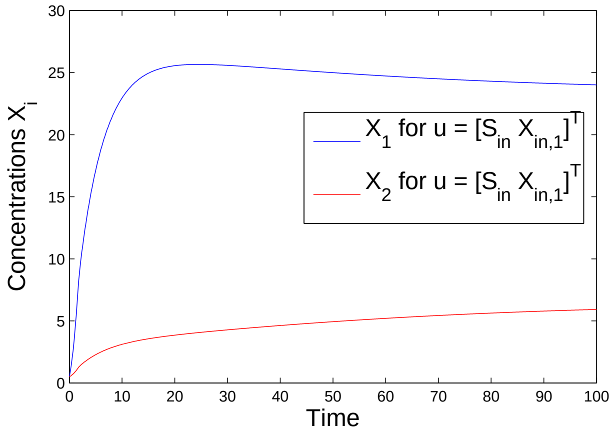



Figure 9.11: Trajectories of $X_1$ and $X_2$ for constants values of Table 9.1 in CASE 2.

**Remark 9.2.1** *In* CASE 1, *the values of* $X_{in,1}$ *and* $X_{in,2}$ *have been fixed such that* $X_{in,1} = 2\,X_{in,2}$. *Conversely, let us consider* $X_{in,2} = 2\,X_{in,1}$, *then the equilibrium is given by* $x_e = [0.036 \quad 7.58 \quad 27.38]^T$. *Therefore, computing the trajectories of* $X_1$ *and* $X_2$ *in open-loop as in Figure 9.10 will also guarantee the coexistence but there exists a larger gap between the values of* $X_1$ *and* $X_2$ *at the equilibrium. Similarly, in* CASE 2, *let us fix the value of* $X_{in,1} = 5$ *instead of* $10$ *and the resulting equilibrium is given by* $x_e = [0.063 \quad 11.74 \quad 13.19]^T$. *Then* $X_1$ *and* $X_2$ *have similar values whereas previously,* $X_1$ *numerically tended to a larger value than* $X_2$.

Hence, as we have seen in the previous figures, coexistence of the two species is possible in CASE 1 and CASE 2. For these cases, computing the equilibria and the resulting Jacobian matrices $A$ and $B$, which define the linearized system (9.4), gives the following results, which are summarized in Table 9.2.

| | CASE 1 $u = [S_{\text{in}} \quad X_{\text{in},1} \quad X_{\text{in},2}]^T$ | CASE 2 $u = [S_{\text{in}} \quad X_{\text{in},1}]^T$ |
|---|---|---|
| $x_e = \begin{bmatrix} S_e \\ X_{1e} \\ X_{2e} \end{bmatrix}$ | $\begin{bmatrix} 0.0425 \\ 16.5888 \\ 18.3687 \end{bmatrix}$ | $\begin{bmatrix} 0.0635 \\ 23.48314607 \\ 6.453361870 \end{bmatrix}$ |
| $u_e$ | $\begin{bmatrix} 20 \\ 10 \\ 5 \end{bmatrix}$ | $\begin{bmatrix} 20 \\ 10 \end{bmatrix}$ |
| $A$ | $\begin{bmatrix} -80.9676 & -0.0794 & -0.1456 \\ 28.9304 & -0.1206 & 0 \\ 51.8372 & 0 & -0.0544 \end{bmatrix}$ | $\begin{bmatrix} -54.0374 & -0.1148 & -0.2000 \\ 38.4076 & -0.0852 & 0 \\ 15.4298 & 0 & 0 \end{bmatrix}$ |
| $B$ | $\begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$ | $\begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \\ 0 & 0 \end{bmatrix}$ |

Table 9.2: Equilibria and associated Jacobian matrices used in the linearized system (9.4).

The next two sections are devoted to the application of input/state-invariant LQ controls to the chemostat model in order to improve the coexistence of the species in these two cases of choice of the input $u$. Indeed, Figures 9.10-9.11 have been realized with an open-loop design for a constant $u$. Now solving an $\boldsymbol{LQ}^{\text{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem allows us to deal with a stabilizing control. In that case, the optimal control is of state feedback type and we can therefore benefit of key properties of closed-loop systems, which are notably tracking, disturbance and noise suppression, sensitivity to structured plant parameter variations, see e.g. [AM90, Section 5.3] for details. Furthermore, it is stated in [AM90, Section 5.5] that for the optimal state feedback $K$ arising from an LQ-optimal design, the optimal closed-loop system maintains asymptotic stability when sectorial nonlinearities are introduced. Moreover, the nonlinearities may be time-varying. The properties of the LQ problem which are highlighted in the sequel are robustness, desensitization due to small variations in some parameters and stabilization. Numerical simulations that follow are testing the effectiveness of the LQ problem in comparison with the results obtained in Figures 9.10-9.11.

In order to highlight the robustness of the LQ problem with respect to pertubations of parameters, an analysis of perturbations is done in the sequel. It is shown that, despite a small variation in some parameter at a fixed time, the LQ-optimal control is able to reestablish the convergence to the equilibrium in order to guarantee the coexistence of the species. There are several types of perturbations, notably, those due to the laboratory conditions (e.g. on the dilution rate $D$) or to the uncertainty on some biological parameters (e.g. on $K_{S,i}$ or $\mu_{\max,i}$).

First, let us apply a perturbation on the dilution rate $D$ for the open-loop system (9.5) with a constant input $u$ given by (9.7) (CASE 1). Formally, at time $t = 50$ (corresponding to the time where the trajectories had numerically converge, see Figure 9.10), the value of the dilution rate is changed to $D = 0.7$ instead of $D = 0.2$. That means that, by applying the conditions of the CEP, the winner of the competition has changed. Indeed, computing the growth curves $\mu_1(S)$ and $\mu_2(S)$ with $D = 0.7$, one can observe in Figure 9.12, that the smallest break-even concentration is now $\lambda_1$ and no more $\lambda_2$. Then $X_1$ will win the competition in the case of an open-loop design with a constant input $u = S_{\text{in}}$.

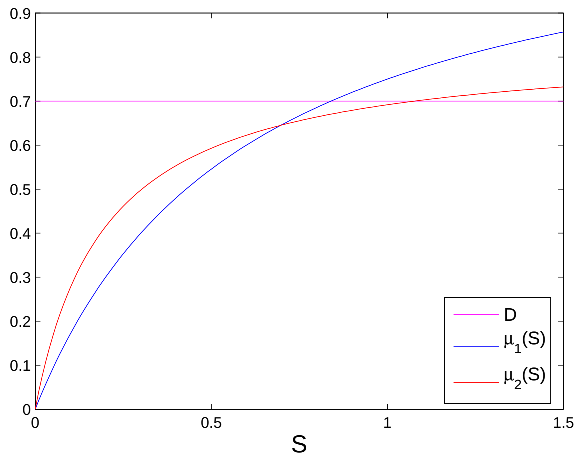



Figure 9.12: Growth rate functions $\mu_1(S)$ and $\mu_2(S)$ compared to $D = 0.7$ for constants values of Table 9.1.

Now, the state trajectories $X_1$ and $X_2$ of system (9.5) with a constant input $u$ given by (9.7) and with an instantaneous increase of $D$ at time $t = 50$, such that $D = 0.7$, are depicted in Figure 9.13. By comparison with Figure 9.10, the trajectories are numerically divergent. Indeed, since the input is constant and in open-loop, the control can not react to the perturbation in order to make the state trajectories numerically converge. The great advantage of the LQ design is the fact that the input can react to perturbations. Moreover, as we will see in Section 9.4, since in the $\boldsymbol{LQ}^{\text{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem, the optimal control is of state-feedback type, the closed-loop allows the system to better react to perturbations.

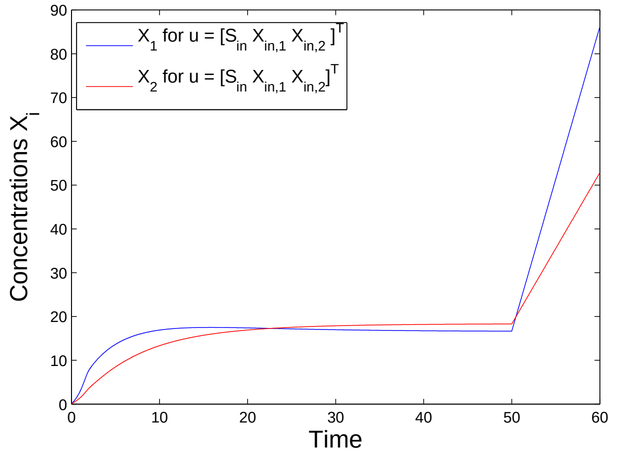



Figure 9.13: Perturbed trajectories of $X_1$ and $X_2$ for constants values of Table 9.1 in CASE 1 with $D = 0.7$ at $t = 50$.

Next, let us insert a perturbation in the growth function $\mu_1(S)$ for the open-loop system (9.5) with a constant input $u$ given by (9.7) (CASE 1). Formally, at time $t = 50$, the value of the maximum growth rate $\mu_{\max,1}$ is changed to $\mu_{\max,1} = 1$ instead of $\mu_{\max,1} = 1.2$ while the value of the dilution rate is kept to $0.2$ as initially. This can be seen as the fact that we have some uncertainty on the nominal parameters. Here, applying the CEP, the winner of the competition has not changed. Indeed, computing the growth curves $\mu_1(S)$ and $\mu_2(S)$ with $\mu_{\max,1} = 1$, one can observe in Figure 9.14, that the smallest break-even concentration is $\lambda_2$ as previously. Then, with these values, $X_2$ will win the competition in the case of an open-loop design with a constant input $u = S_{\text{in}}$.

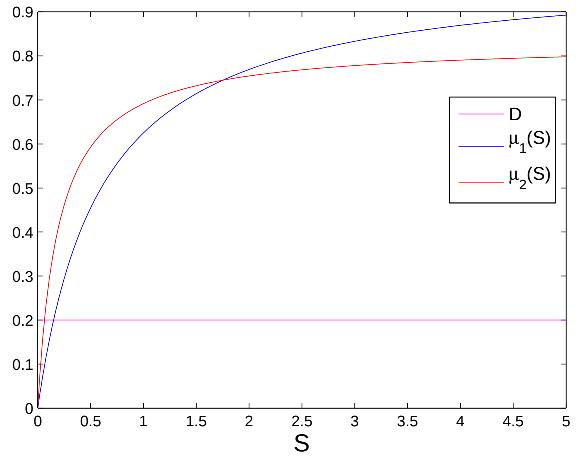



Figure 9.14: Growth rate functions $\mu_1(S)$ and $\mu_2(S)$ compared to $D = 0.2$ for $\mu_{\max,1} = 1$ and constants values of Table 9.1.

Now, the state trajectories $X_1$ and $X_2$ of system (9.5) with a constant input $u$ given by (9.7) and with an instantaneous increase of $\mu_{\max,1}$ at time $t = 50$, such that $\mu_{\max,1} = 1$, are depicted in Figure 9.15. By comparison with Figure 9.10, the trajectories are also numerically divergent. Indeed, as previously, since the input is a priori fixed at a constant value, the control can not react to the perturbation.

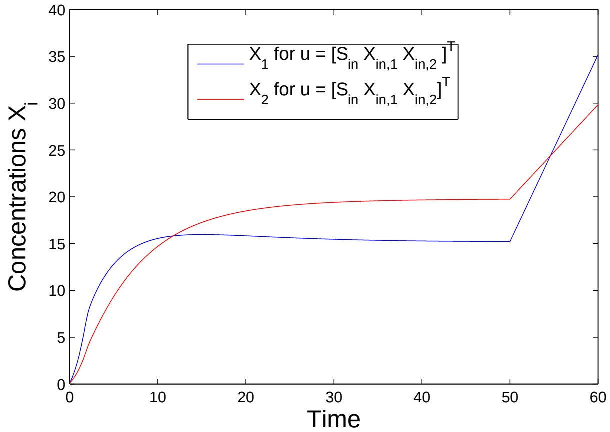



Figure 9.15: Perturbed trajectories of $X_1$ and $X_2$ for constants values of Table 9.1 in CASE 1 with $\mu_{\max,1} = 1$ at $t = 50$.

In the following two sections, two different methods are developed in order to compute an appropriate (and robust) optimal control law $\tilde{u}$ which ensures the coexistence of the species. These methods are based on the $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem and the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem respectively.

## 9.3 The input/state-invariant LQ problem

### 9.3.1 Problem statement

The finite horizon input/state-invariant $\boldsymbol{LQ^{t_f}_{\tilde{u},\tilde{x}}}$ problem applied to the chemostat model (where $W$ and $Z$ are equal to the identity matrix, see Chapter 3), consists of minimizing the quadratic functional :

$$J(\tilde{x}_0,\, \tilde{u}) = \frac{1}{2}\left(\int_0^{t_f} (\|R^{1/2}\tilde{u}(t)\|^2 + \|C\,\tilde{x}(t)\|^2)\,\mathrm{dt} + \tilde{x}(t_f)^T S\,\tilde{x}(t_f)\right) \tag{9.10}$$

for a given linear time-invariant system $[A,\, B]$ described by (9.4)

$$\dot{\tilde{x}} = A\,\tilde{x} + B\,\tilde{u},$$

which is the linearization of system (9.5) around the equilibrium $(x_e,\, u_e)$, with $x_e \gg 0$ and $u_e \gg 0$, and with the initial condition $\tilde{x}_0 \geq -x_e$, under the constraints

$$\forall\, t \in [0,\, t_f],\ \begin{cases} \tilde{x}(t) \geq -x_e \\ \tilde{u}(t) \geq -u_e, \end{cases} \tag{9.11}$$

where $t_f$ is a fixed final time, $R \in \mathbb{R}^{m\times m}$ is a symmetric positive definite matrix, $C \in \mathbb{R}^{p\times n}$ and $S \in \mathbb{R}^{n\times n}$ is a symmetric positive semidefinite matrix. Recall the result of Theorem 3.2.1 which gives the solution of the $\boldsymbol{LQ^{t_f}_{\tilde{u},\tilde{x}}}$ problem as follows :

**Theorem 9.3.1** *The control function $\tilde{u}(\cdot)$ is solution of the $\boldsymbol{LQ^{t_f}_{\tilde{u},\tilde{x}}}$ problem $\Leftrightarrow\ \exists\,\lambda(\cdot)$ and $\upsilon(\cdot)$ such that*

$$\tilde{u}(t) = -R^{-1}B^Tp(t)\ +\ R^{-1}\upsilon(t),\ t \in [0,\, t_f],$$

*where*

$$\begin{bmatrix} \dot{\tilde{x}}(t) \\ \dot{p}(t) \end{bmatrix} = H \begin{bmatrix} \tilde{x}(t) \\ p(t) \end{bmatrix} + \begin{bmatrix} B\,R^{-1}\upsilon(t) \\ \lambda(t) \end{bmatrix},\ t \in [0,\, t_f] \tag{9.12}$$

*with*

$$\begin{cases} \tilde{x}(0) = \tilde{x}_0 \\ p(t_f) = S\,\tilde{x}(t_f) - \lambda(t_f) \end{cases}$$

*where*

$$H = \begin{bmatrix} A & -B\,R^{-1}B^T \\ -C^TC & -A^T \end{bmatrix}$$

*is the Hamiltonian matrix, and for all $t \in [0,\, t_f]$*

$$\begin{aligned} \tilde{x}(t) &\geq -x_e, \\ \tilde{u}(t) &\geq -u_e, \\ \lambda(t) &\geq 0, \\ \upsilon(t) &\geq 0, \\ \lambda(t)^T\,(\tilde{x}(t) + x_e) &= 0 \quad \textit{(state complementarity condition).} \\ \upsilon(t)^T\,(\tilde{u}(t) + u_e) &= 0 \quad \textit{(input complementarity condition).} \end{aligned}$$

In the context of coexistence of species in a chemostat model, the $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\boldsymbol{\tilde{u}},\boldsymbol{\tilde{x}}}$ problem is solved for the linearized system (9.4) as an optimization problem with state and input constraints. This methodology is inspired by the results of Theorem 8.2.1. For a sufficiently large horizon $t_f$, the optimal control $u_L = \tilde{u} + u_e$ can be applied to the nonlinear system (9.5). Therefore, the strict compliance of the constraints on $\tilde{x}$ and $\tilde{u}$ will ensure the strict positivity of the nonlinear state and input trajectories by following the same reasoning as in Theorem 8.2.1. Numerical simulations are done to illustrate this and to ensure therefore the local coexistence w.r.t. the initial condition $x_0$ (which is chosen close to $x_e$), for a sufficiently large horizon $t_f$.

### 9.3.2 Numerical simulations

Consider system (9.4) with $A$ and $B$ given in Table 9.2, and the cost (9.10) where

$$C = 0_{3\times 3} \;\text{ and }\; R = I_m \tag{9.13}$$

where $m = 3$ or $m = 2$ depending on the choice of $u$ (CASE 1 or CASE 2 developed in Subsection 9.2.3). The numerical solution of this problem is computed by using `Matlab` and the function `quadprog`, as in Section 4.3.2. First, the continuous time problem is converted into a discrete time one by using sampling : for $i = 0, \ldots, N-1$, with $t_f = N\,h$, $\tilde{u}(t) = \tilde{u}(i\,h) =: \tilde{u}_i$, for $t \in [ih,\,(i+1)\,h]$, where $h$ is the sampling time. The resulting discrete time system is given by :

$$\tilde{x}_{i+1} = e^{A\,h}\,\tilde{x}_i + \left(\int_0^h e^{A\tau} B\,\mathrm{d}\tau\right)\,\tilde{u}_i, \quad i = 0, \ldots, N-1 \tag{9.14}$$

with the following discrete time cost $\dfrac{1}{2}\displaystyle\sum_{i=0}^{N-1} h\,\|\tilde{u}_i\|^2 + \tilde{x}_N^T S\,\tilde{x}_N$, see Appendix C for details on discretization. In the following numerical simulations and figures, $x_L$ denotes the shifted state trajectories coming from the linearized system, i.e. $x_L := \tilde{x} + x_e$ whereas $(S,\,X_1,\,X_2)$ denotes the nonlinear state trajectories of system (9.5).

Consider the final time $t_f = 50$ with the sampling time $h = 1$ and the initial condition $\tilde{x}_0 = [0.1 \quad 0.1 \quad 0.1]^T$ (which is chosen near zero since $x_L(0)$ has to be close enough to the equilibrium $x_e$). With this sampling time, we obtain the following matrices $A$ and $B$ defining the sampled data system (9.14), see Table 9.3.

| | CASE 1 $u = [S_{\text{in}} \quad X_{\text{in},1} \quad X_{\text{in},2}]^T$ | CASE 2 $u = [S_{\text{in}} \quad X_{\text{in},1}]^T$ |
|---|---|---|
| $A$ | $\begin{bmatrix} -0.0012 & -0.0008 & -0.0015 \\ 0.2808 & 0.8630 & -0.0444 \\ 0.5392 & -0.0435 & 0.8647 \end{bmatrix}$ | $\begin{bmatrix} -0.0021 & -0.0017 & -0.0032 \\ 0.5703 & 0.8494 & -0.1254 \\ 0.2506 & -0.0289 & 0.9474 \end{bmatrix}$ |
| $B$ | $\begin{bmatrix} 0.0022 & -0.0002 & -0.0003 \\ 0.0628 & 0.1860 & -0.0046 \\ 0.1163 & -0.0045 & 0.1862 \end{bmatrix}$ | $\begin{bmatrix} 0.0032 & -0.0004 \\ 0.1254 & 0.1846 \\ 0.0526 & -0.0029 \end{bmatrix}$ |

Table 9.3: Matrices $A$ and $B$ defining the discrete-time linearized system (9.14).

– CASE 1 :

First, consider the CASE 1 for the choice of $u$ with $S = I_3$. The optimization algorithm mentioned above leads to the optimal control $u_{L,i} = \tilde{u}_i + u_e$ applied to the nonlinear system (9.5) and depicted in Figure 9.16. The corresponding state trajectories $x_i$ of the nonlinear system (9.5) with the optimal control $u_i$ and the state trajectories $x_{L,i} = \tilde{x}_i + x_e$ of the linearized system (9.4) are depicted in Figure 9.17.

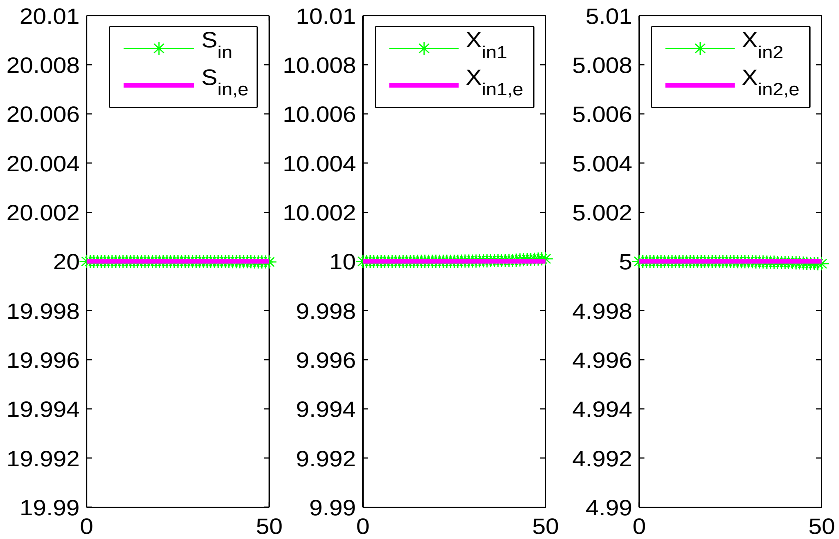



Figure 9.16: Optimal control for sampled data system (9.14) in CASE 1 with $S = I_3$.

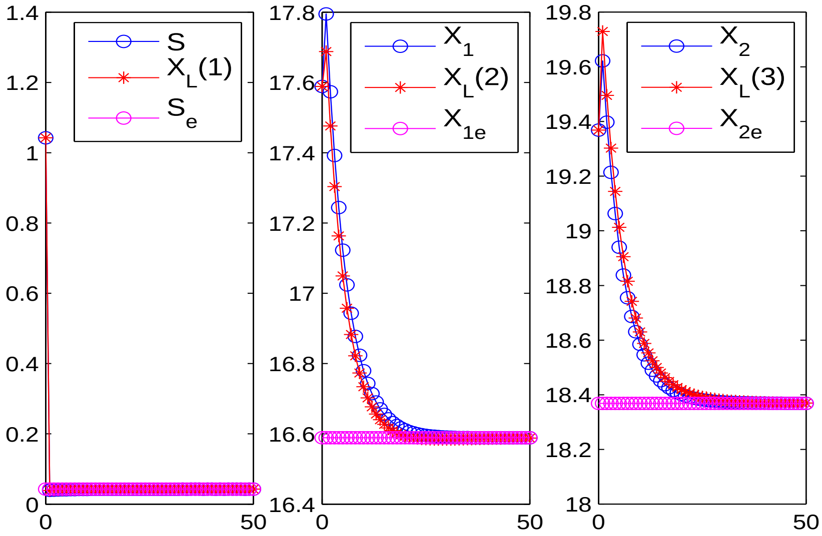



Figure 9.17: State trajectories for sampled data system (9.14) in CASE 1 with $S = I_3$.

In these figures, one can observe that the optimal control remains very close to the equilibrium. Unless otherwise stated, this behavior is always observed and is therefore not always mentioned in the sequel. On the other hand, Figure 9.17 shows that the substrate $S$ decreases quickly to give a boost to $X_1$ and $X_2$ which numerically tend smoothly to the equilibrium before the final fixed time $t_f = 50$.

Now an analysis can be done on a variation of the initial state $\tilde{x}_0$ to see how the model reacts with changes on the initial condition. First, consider $\tilde{x}_0 = [0.1 \quad 0.1 \quad -18.36]^T$ which is the case where $\tilde{x}_3(0)$ is close to $x_{e3}$ from below, that means that, for the nonlinear system, the initial condition $x_3(0)$ is near zero. In this case, with $S = I_3$ and $t_f = 50$, the corresponding optimal control $u_{L,i} = \tilde{u}_i + u_e$ applied to the nonlinear system (9.5) is close to the equilibrium as in the previous case and the corresponding state trajectories $x_i$ of the nonlinear system (9.5) with the optimal control $u_i$ and the state trajectories $x_{L,i} = \tilde{x}_i + x_e$ of the linearized system (9.4) are drawn in Figure 9.18.

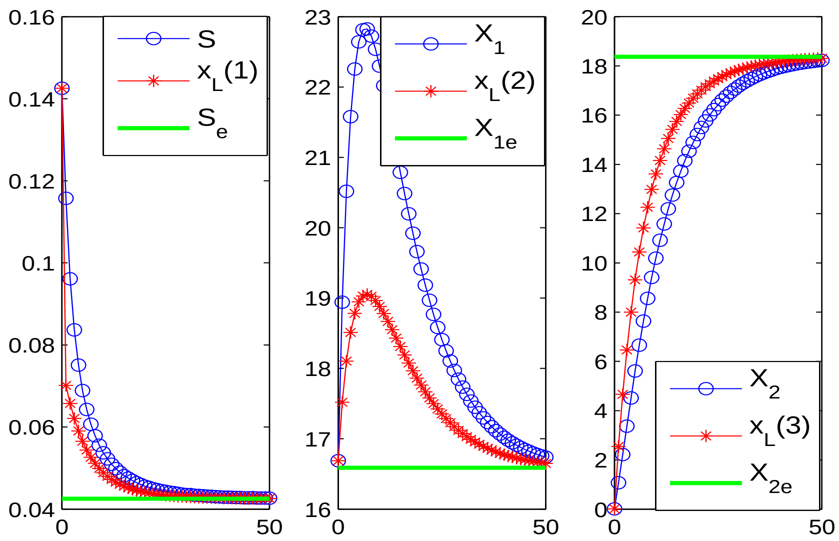



Figure 9.18: State trajectories for sampled data system (9.14) in CASE 1 with $S = I_3$ and $\tilde{x}_0 = [0.1 \quad 0.1 \quad -18.36]^T$.

One can observe in this case that the linear approximation of the nonlinear system is not so good and moreover the state trajectory $x_2(t) = X_1(t)$ is not exactly equal to the equilibrium at the final time. Then one can increase the final time to be $t_f = 100$ for example and then the equilibrium is reached in the long run while the linear approximation of the chemostat model is not so good for small time. In Figure 9.19, a comparison is done between the state trajectories with penalization matrix $S = I_3$ and with $S = 50\, I_3$ for $t_f = 50$. Instead of increasing the final time, one can increase the penalization of the final state in order to reach almost exactly the equilibrium at the final time. In addition, in order to make a precise comparison of these state trajectories for different values of $S$, the relative error, denoted $\varepsilon_r$, of these two curves is computed, which is defined by

$$\varepsilon_r := \frac{\|x_{S=I} - x_{S=50I}\|_2}{\|x_{S=I}\|_2}$$

where $x_{S=I}$ and $x_{S=50I}$ denote the nonlinear state trajectories for $S = I_3$ and for $S = 50\, I_3$, respectively. It is depicted in Figure 9.20 below.

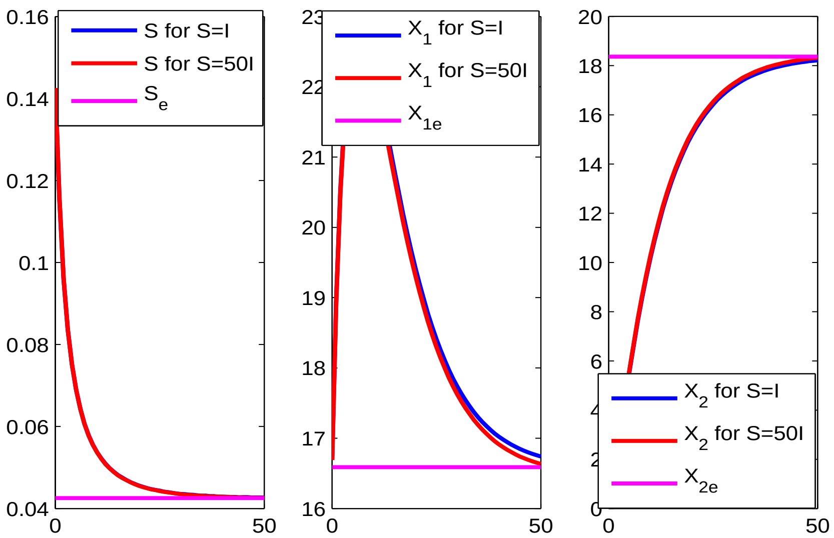



Figure 9.19: Comparison of the state trajectories for sampled data system (9.14) in CASE 1 with $S = I_3$ and $S = 50\, I_3$ for $\tilde{x}_0 = [0.1 \quad 0.1 \quad -18.36]^T$ and $t_f = 50$.

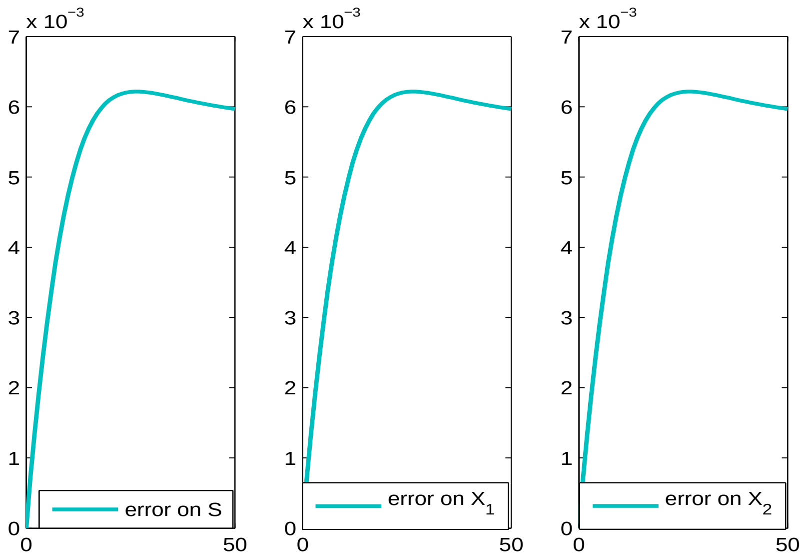



Figure 9.20: Relative error of the state trajectories for system (9.14) in CASE 1 with $S = I_3$ and $S = 50\, I_3$ for $\tilde{x}_0 = [0.1 \quad 0.1 \quad -18.36]^T$.

Another important question is how much can the initial condition $x_0$ be far from the equilibrium while always ensuring the validity of the linearized model. By values from above, a numerical analysis has been done with several values of initial conditions and the linearized model seems to be valid up to a value of $x_0 < x_e + 28$. Indeed, if we consider $x_0 = x_e + 28$, i.e. $\tilde{x}_0 = [28 \quad 28 \quad 28]^T$, the following state trajectories are obtained, see Figure 9.21.

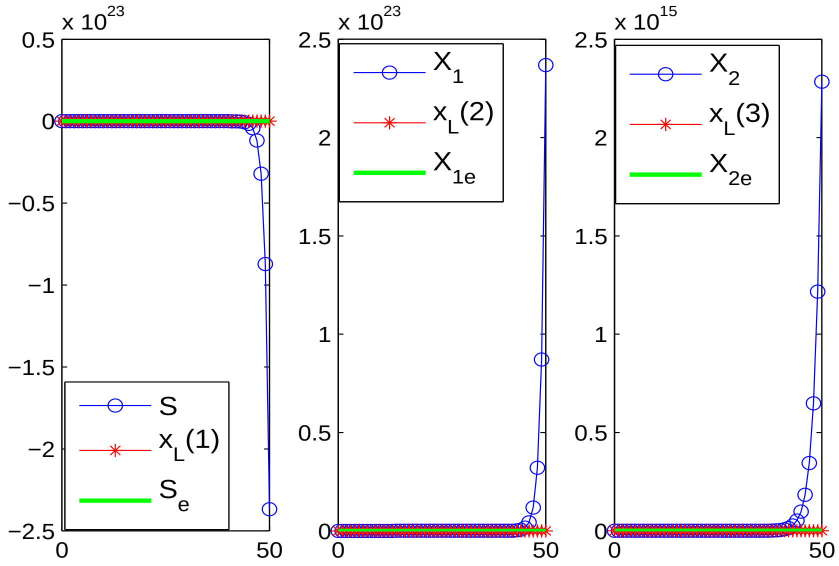



Figure 9.21: State trajectories for sampled data system (9.14) in CASE 1 with $S = I_3$ and $\tilde{x}_0 = [28 \quad 28 \quad 28]^T$.

One can see that the nonlinear state trajectories numerically diverge and therefore the linearization of the nonlinear system is no more a good approximation whereas the one for $x_0 = x_e + 27$ has a really good behavior. However, one can observe that the linear state trajectories $x_{L,i} = \tilde{x}_i + x_e$ have the right behavior since they numerically tend to the equilibrium in the long run. Moreover, increasing the penalization matrix $S$, as we have done in a previous analysis, does not improve the results.

Finally, an analysis of perturbations can be done here for the $\boldsymbol{LQ}^{t_f}_{\bar{u},\bar{x}}$ problem. As previously, consider a perturbation on the dilution rate $D$. Let us recall Figure 9.17 which depicts the state trajectories for the sampled data system (9.14) in CASE 1 with $S = I_3$ and $\tilde{x}_0 = [0.1 \quad 0.1 \quad 0.1]^T$. In this figure, one could see that the state trajectories $x(t)$ of the nonlinear system (9.5) numerically converged to the equilibrium around time $t = 40$. Then, at this time, the value of $D$ is changed to $0.7$, so that the role of the species is inverted. The associated optimal control and state trajectories are therefore computed for the perturbed resulting linearized system by using the optimization algorithm mentioned above for $S = I_3$, $t_0 = 40$, $t_f = 100$ and the initial condition equal to the last value of the previous state trajectories. Then the optimal control $u_{L,i} = \tilde{u}_i + u_e$ is applied to the nonlinear system (9.5) and is depicted, together with the previous control from time $0$ to time $100$, in Figure 9.22. The corresponding nonlinear state trajectories, from time $0$ to time $100$, are drawn in Figure 9.23.

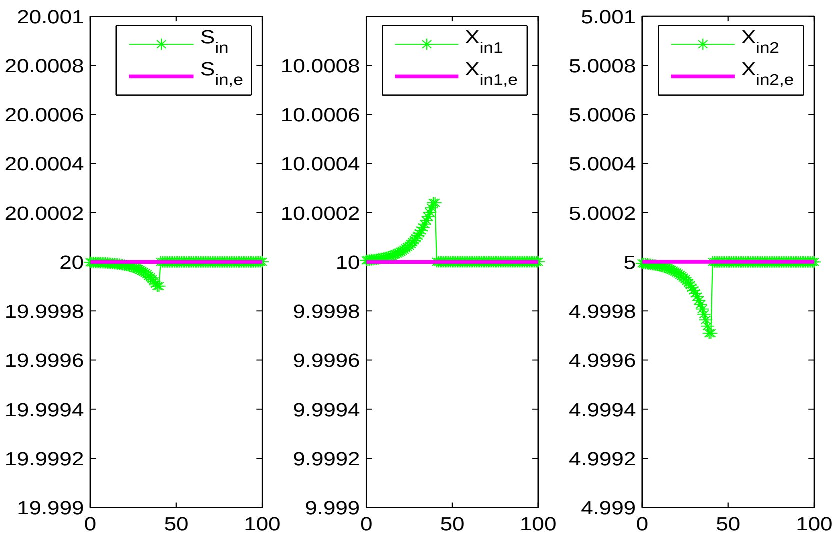



Figure 9.22: Optimal control for perturbed sampled data system (9.14) in CASE 1 with $S = I_3$ for $D = 0.7$ at time $t = 40$.

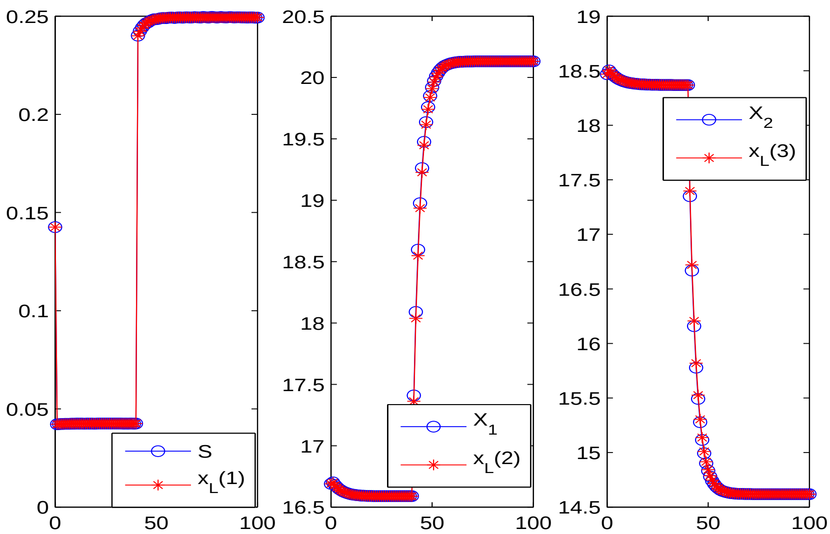



Figure 9.23: State trajectories for perturbed sampled data system (9.14) in CASE 1 with $S = I_3$ for $D = 0.7$ at time $t = 40$.

One can see that the optimal control has to be adapted in order to bring the state trajectories to the equilibrium. As a result, the linearized shifted state trajectories, together with the non-linear state trajectories, numerically converge without any difficulty to a new equilibrium, the one which corresponds to system (9.5) with $D = 0.7$. Anyway, the coexistence of the species is guaranteed.

Let us perform a similar analysis by considering a perturbation on the growth curve $\mu_1(S)$. At time $t = 40$, the value of $\mu_{\max,1}$ is changed to $1$ (instead of $1.2$). The associated optimal control and state trajectories are computed as previously and are depicted in Figures 9.24 and 9.25 respectively.

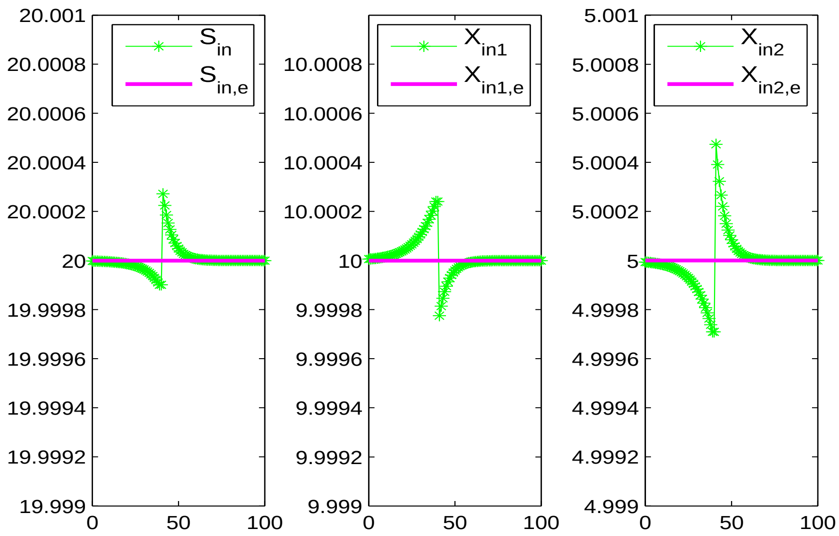



Figure 9.24: Optimal control for perturbed sampled data system (9.14) in CASE 1 with $S = I_3$ for $\mu_{\max,1} = 1$ at time $t = 40$.

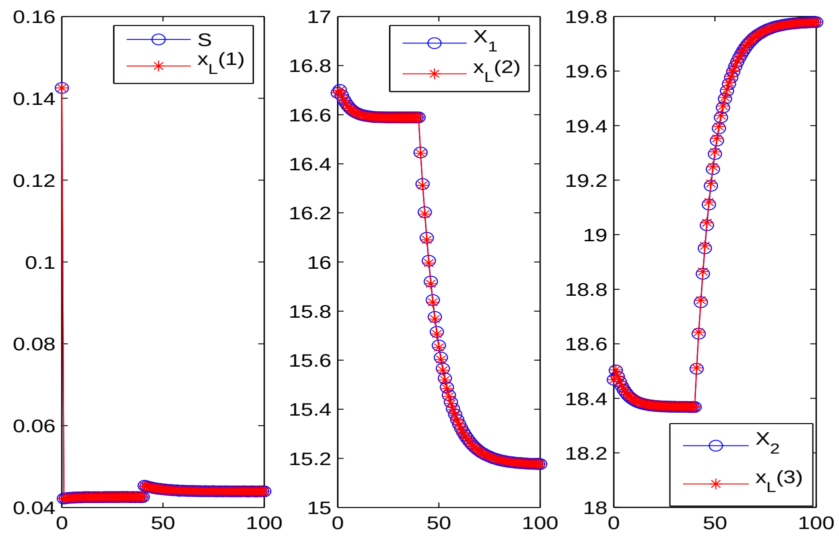



Figure 9.25: State trajectories for perturbed sampled data system (9.14) in CASE 1 with $S = I_3$ for $\mu_{\max,1} = 1$ at time $t = 40$.

Again, one can observe that the optimal control has to be adapted in order to bring the state trajectories to the new equilibrium, corresponding to the equilibrium of system (9.5) with $\mu_{\max,1} = 1$. As in the previous perturbation, the linearized state trajectories, together with the nonlinear one, numerically converge fast to this new equilibrium. And the coexistence of the species is still guaranteed.

– CASE 2 :

Let us consider the CASE 2 for the choice of $u$ with $S = I_3$, $\tilde{x}_0 = [0.1 \quad 0.1 \quad 0.1]^T$ and $t_f = 200$. The optimization algorithm mentioned above leads to the optimal control $u_{L,i} = \tilde{u}_i + u_e$ applied to the nonlinear system (9.5) and depicted in Figure 9.26. The corresponding state trajectories $x_i$ of the nonlinear system (9.5) with the optimal control $u_i$ and the state trajectories $x_{L,i} = \tilde{x}_i + x_e$ of the linearized system (9.4) are depicted in Figure 9.27.

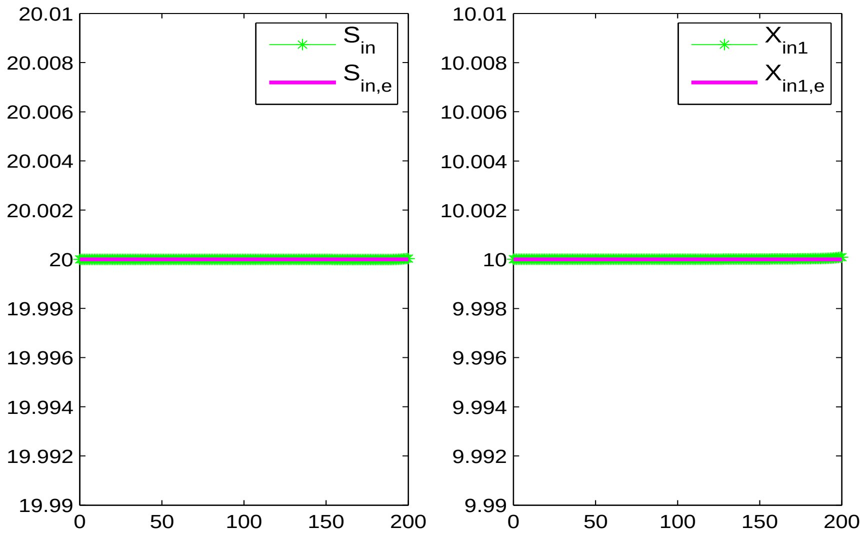



Figure 9.26: Optimal control for sampled data system (9.14) in CASE 2 with $S = I_3$.

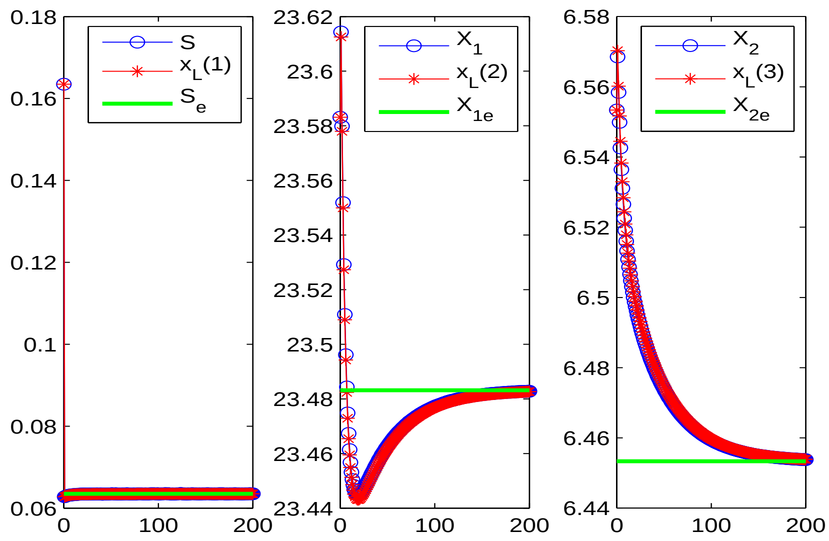



Figure 9.27: State trajectories for sampled data system (9.14) in CASE 2 with $S = I_3$.

As in CASE 1, simulations have been done for $t_f = 50$ and $t_f = 100$. However, the final time has to be more increased (up to $t_f = 200$) in order to reach more precisely the equilibrium in the long run. As already mentioned, the increasing of the penalization matrix $S$ can also help the state trajectories to come closer to the equilibrium and often with a smaller time. See Figure 9.28, which shows a comparison between the behavior of the state trajectories for $S = I_3$ and for $S = 50\,I_3$ with $t_f = 200$. As previously, the relative error is also drawn to give a more precise comparison, see Figure 9.29. In this case, this increasing of the penalization matrix $S$ has less impact than in the previous case since the two curves are close to each other.

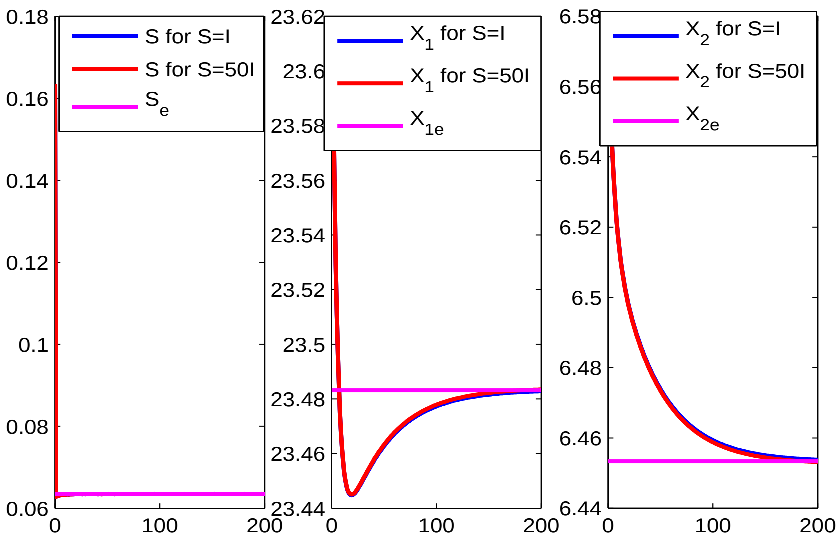



Figure 9.28: Comparison of the state trajectories for sampled data system (9.14) in CASE 2 with $S = I_3$ and $S = 50\,I_3$ for $\tilde{x}_0 = [0.1 \quad 0.1 \quad 0.1]^T$ and $t_f = 200$.

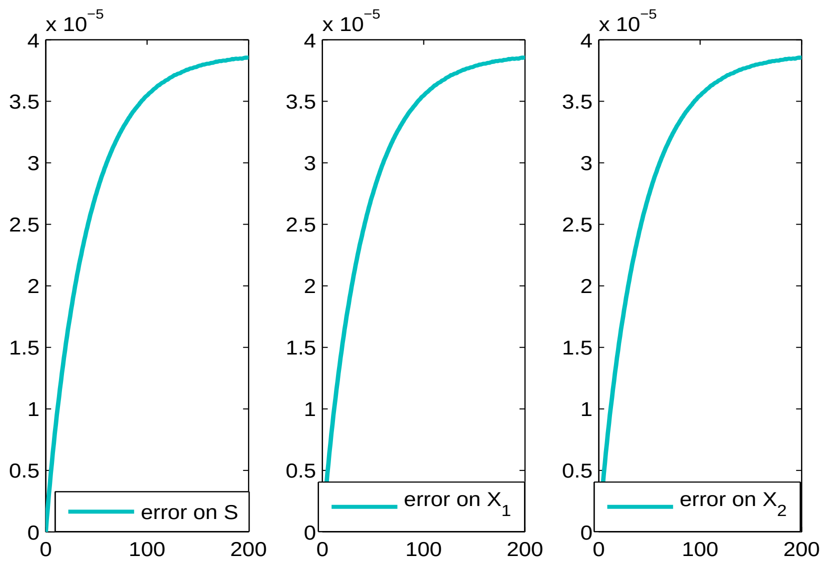



Figure 9.29: Relative error of the state trajectories for system (9.14) in CASE 2 with $S = I_3$ and $S = 50\,I_3$ for $\tilde{x}_0 = [0.1 \quad 0.1 \quad 0.1]^T$.

In general, one can observe that in CASE 2, it is more difficult or it takes more time to reach exactly the equilibrium in comparison with CASE 1. It can be explained by the fact that one has less direct action ("practical control") in this case than in the previous one. Indeed, in CASE 2, there is no control on $X_2$ while in CASE 1, there is a control on every variables, $S$, $X_1$ and $X_2$. Therefore, the CASE 2 is no more studied in the following section (for physical reasons but also for numerical reasons since it appears to react less efficiently).

# 9.4 The inverse input/state-invariant LQ problem

## 9.4.1 Problem statement

The objective in this method is to find an LQ-optimal control of state feedback type $\tilde{u} = K\,\tilde{x}$ such that the linearized closed-loop system $\dot{\tilde{x}} = (A + B\,K)\,\tilde{x}$ is stable and input/state-invariant. By Theorem 8.2.1, this guarantees the strict positivity of the state and input trajectories of the nonlinear system (9.5) in a neighbourhood of the equilibrium $(x_e,\, u_e)$ and this implies the local coexistence of the resulting closed-loop system, by Proposition 9.2.1, with $x_\varepsilon$ sufficiently close to the equilibrium $x_e$. As in the previous section, the initial state $x_0$ is chosen near $x_e$ while $\tilde{x}_0$ is close to zero. To determine an appropriate state feedback $K$ for the linearized system, an inverse input/state-invariant $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem is solved by using several LMIs and/or BMIs as developed in Section 7.3. The $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem can be summarized in two steps :

1. Find a matrix $K$ such that for all $t \geq 0$ and for all $\tilde{x}_0 \gg -x_e$, $\begin{cases} \tilde{x} \gg -x_e \\ \tilde{u} = K\,\tilde{x} \gg -u_e \end{cases}$

2. Determine the existence of $Q = C^T C$ and $R$ such that $(Q,\, A)$ is detectable and the control $\tilde{u} = K\,\tilde{x}$ is optimal for the corresponding LQ problem, i.e. minimizes the following quadratic cost

$$J(\tilde{x}_0,\, \tilde{u}) = \frac{1}{2}\left(\int_0^\infty (\|R^{1/2}\tilde{u}(t)\|^2 + \|C\,\tilde{x}(t)\|^2)\,\mathrm{dt}\right) \tag{9.15}$$

For the first step, the matrix $K$ is obtained by solving BMI 1b or LMI 1d (see Table 9.4 below). For the second step, solving LMI 2 gives appropriate matrices $Q$ and $R$, see Chapter 7.

| | BMI and LMI used for the computation of $K$ |
|---|---|
| **BMI 1b :** | $$\mathrm{diag}[(\mathcal{H})_{ij}]_{i\neq j} \succeq 0$$ $$\begin{bmatrix} -I_n \\ -K \end{bmatrix} (A + B\,K) - \mathcal{H} \begin{bmatrix} -I_n \\ -K \end{bmatrix} = \begin{bmatrix} 0_{n\times n} \\ 0_{m\times n} \end{bmatrix}$$ $$\mathrm{diag}\left[\mathcal{H} \begin{bmatrix} x_e \\ u_e \end{bmatrix}\right] \preceq 0$$ |
| **LMI 1d :** | $$\mathrm{diag}[(A + B\,K)_{ij}]_{i\neq j} \succeq 0$$ $$\mathrm{diag}[(A + B\,K)\,x_e] \preceq 0$$ $$\mathrm{diag}[K\,x_0 + u_e] \succeq 0$$ |
| | LMI used for the computation of $Q$ and $R$ for the LQ problem |
| **LMI 2 :** | $$R \succ 0$$ $$Q \succeq 0$$ $$P_+ \succeq 0$$ $$P_1 \succ 0$$ $$(A + B\,K)^T P_+ + P_+(A + B\,K) + K^T R\,K + Q = 0$$ $$B^T P_+ + R\,K = 0$$ $$A^T P_1 + P_1\,A \prec Q$$ |

Table 9.4: LMIs and BMI used in the resolution of the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem.

### 9.4.2 Numerical simulations

As already mentioned, the CASE 2 for the choice of $u$ is not studied for this methodology. The implementation of the previous method has revealed that one has less possible action in the CASE 2 than in the CASE 1. The latter is a better choice for the input to control the chemostat model and to ensure the coexistence of the species. Thus, consider CASE 1. First, determine the matrix $K$ by solving BMI 1b with YALMIP in MATLAB. That gives matrices $H_1$ and $K_1$. Unfortunately, when solving LMI 2, it is not possible to find corresponding matrices $Q$ and $R$ such that $K_1$ gives an LQ-optimal control. As we have seen in Chapter 7, the matrix $K$ solving the BMI 1b is not unique and may not be admissible for LMI 2. Therefore, another matrix $K_2$ is computed by solving again BMI 1b with an additional condition : $A + B\,K_2 \prec A + B\,K_1$ (by the same idea of the heuristic iterative process described in Table 7.1, by creating a decreasing sequence of matrices). This gives the following results :

$$H_2 = H = \begin{bmatrix} -80.815 & 0.020441 & 0.033481 & 0.078248 & 0.022940 & 0.044187 \\ 28.661 & -0.91434 & 0.022320 & 1.7724\,10^{-4} & 1.6114\,10^{-3} & 0 \\ 50.809 & 4.7736\,10^{-3} & -1.6606 & 0.012093 & 1.3123\,10^{-3} & 3.7499\,10^{-8} \\ 4.2681\,10^{-5} & 0.10419 & 1.6110 & -1.6001 & 4.6925\,10^{-5} & 0.13610 \\ 0.061421 & 0.061039 & 0.44864 & 3.3414\,10^{-4} & -0.92796 & 3.3381\,10^{-3} \\ 0.093693 & 0.34155 & 0.19632 & 0.020567 & 4.9874\,10^{-5} & -1.9375 \end{bmatrix}$$

and

$$K_2 = K = \begin{bmatrix} -0.89593 & 0.067073 & -1.4548 \\ -1.3595 & -4.0011 & 0.111211 \\ -5.2064 & 1.6701\,10^{-3} & -8.1182 \end{bmatrix} \tag{9.16}$$

such that

- $H$ is a Metzler matrix ;
- $\begin{bmatrix} -I_3 \\ -K \end{bmatrix} (A + B\,K) - \mathcal{H} \begin{bmatrix} -I_3 \\ -K \end{bmatrix} = 10^{-15}\,I_{6\times 3} \simeq \begin{bmatrix} 0_{3\times 3} \\ 0_{3\times 3} \end{bmatrix}$ ;
- $\mathcal{H} \begin{bmatrix} x_e \\ u_e \end{bmatrix} = \begin{bmatrix} -0.46802 \\ -13.519 \\ -28.008 \\ -8.9240\,10^{-4} \\ -3.1225\,10^{-16} \\ -6.0244\,10^{-8} \end{bmatrix} \leq 0_{6\times 1}.$

Then solving the LMI 2 gives the following weighting matrices such that $\tilde{u} = K\,\tilde{x}$ is optimal for the LQ problem with cost (9.15),

$$\begin{aligned} Q &= \begin{bmatrix} 1.5971\,10^{-3} & 5.1037\,10^{-5} & 2.6478\,10^{-3} \\ 5.1037\,10^{-5} & 3.6534\,10^{-5} & 6.6413\,10^{-5} \\ 2.6478\,10^{-3} & 6.6413\,10^{-5} & 4.4001\,10^{-3} \end{bmatrix}; \\ R &= \begin{bmatrix} 1.9001\,10^{-5} & 2.7891\,10^{-7} & 2.9418\,10^{-5} \\ 2.7891\,10^{-7} & 1.7648\,10^{-6} & -5.4882\,10^{-7} \\ 2.9418\,10^{-5} & -5.4882\,10^{-7} & 4.6255\,10^{-5} \end{bmatrix}. \end{aligned} \tag{9.17}$$

These results lead to the following optimal control law $u_L = K\,\tilde{x} + u_e$, which is depicted in Figure 9.30, with initial condition $\tilde{x}_0 = [0.1 \quad 0.1 \quad 0.1]^T$. This control, applied to system (9.5), gives the following closed-loop system

$$\dot{x} = f(x) + B\,(K\,x - (K\,x_e - u_e)). \tag{9.18}$$

and the associated state trajectories $x_L = \tilde{x} + x_e$ compared to the state trajectories of the nonlinear closed-loop system (9.18) are drawn in Figure 9.31.

In the sequel, unless otherwise stated, the behavior of the optimal control law $u_L = K\,\tilde{x} + u_e$ is similar to the one depicted in Figure 9.30. Now, it is interesting to compare the results which are obtained with $K_1$, a stabilizing input/state-invariant feedback which is not admissible for the resulting LQ problem, with the results which are obtained with $K_2$, the optimal state feedback, solution of the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem. See Figure 9.32 which gives a comparison between the nonlinear state trajectories for a first solving of BMI 1b (with $\tilde{u} = K_1\,\tilde{x}$), represented in the caption by "$S,\, X_1,\, X_2$ - LMI1", together with the optimal state trajectories after solving BMI 1b and LMI 2 (with $\tilde{u} = K_2\,\tilde{x}$), represented in the caption by "$S,\, X_1,\, X_2$ - LMI2". Furthermore, for a more precise comparison, Figure 9.33 represents the relative error between these two curves, defined as follows

$$\varepsilon_r := \frac{\|x_{LMI1} - x_{LMI2}\|_2}{\|x_{LMI1}\|_2}$$

where $x_{LMI1}$ ($x_{LMI2}$) denotes the state trajectories obtained after solving BMI 1b (BMI 1b and LMI 2, respectively). One can numerically observe in this case that solving an inverse $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem stabilizes faster the resulting closed-loop system. The optimal state trajectories numerically tend toward $x_e$ in a shorter time than the nonlinear state trajectories coming from the linearized system $\dot{\tilde{x}} = (A + B\,K_1)\,\tilde{x}$.

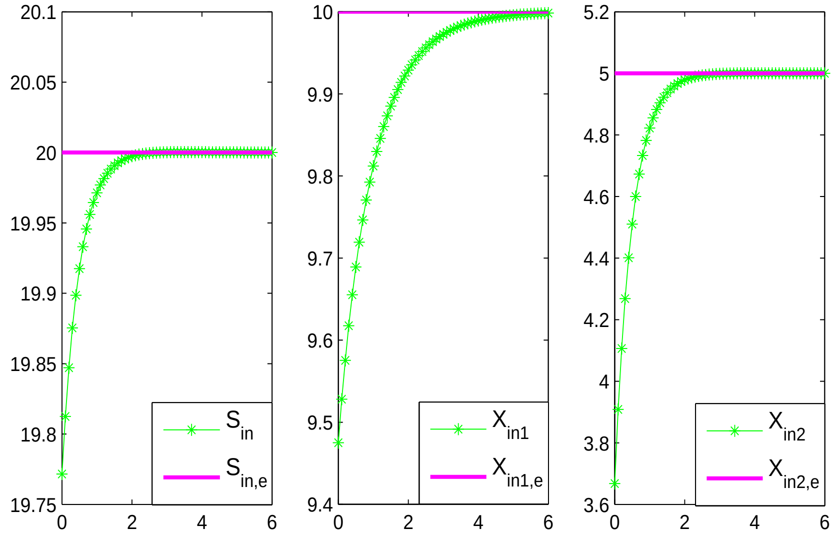



Figure 9.30: Optimal control $u_L = K\,\tilde{x} + u_e$ with $K$ given by (9.16).

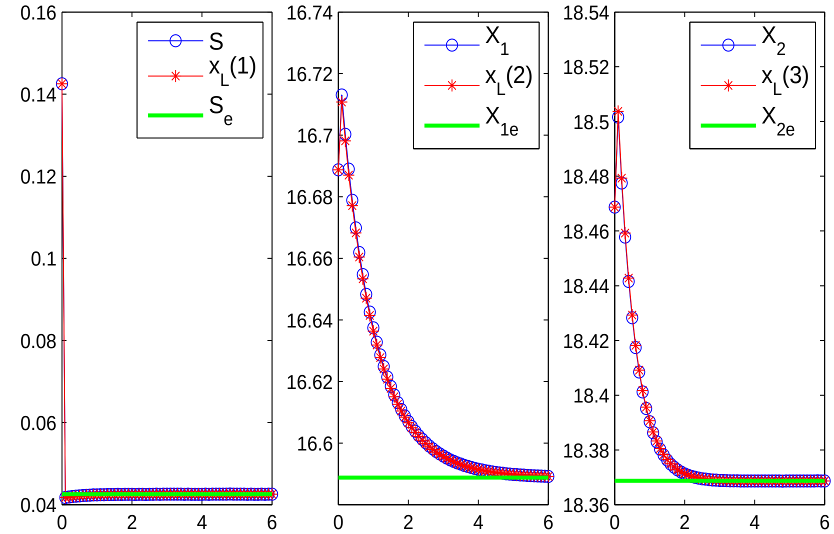



Figure 9.31: State trajectories for system (9.18) with $K$ given by (9.16).

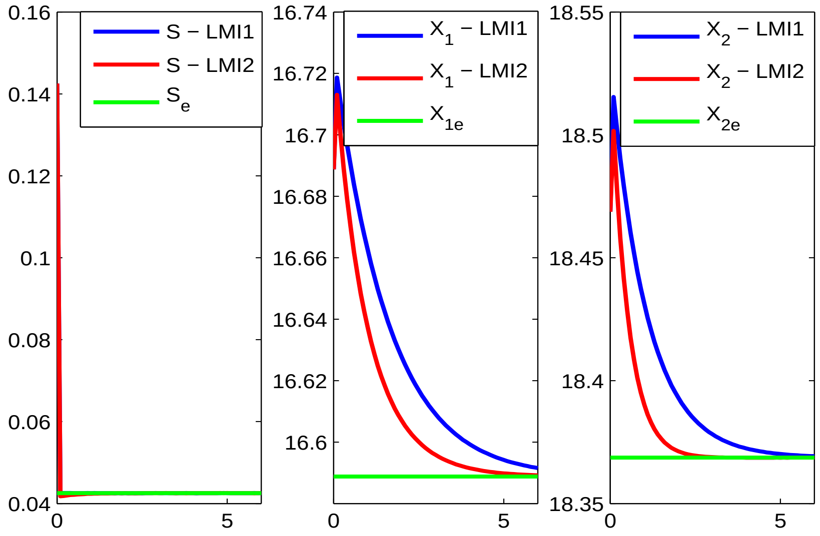



Figure 9.32: State trajectories for system (9.18) with $K_1$ and $K_2$.

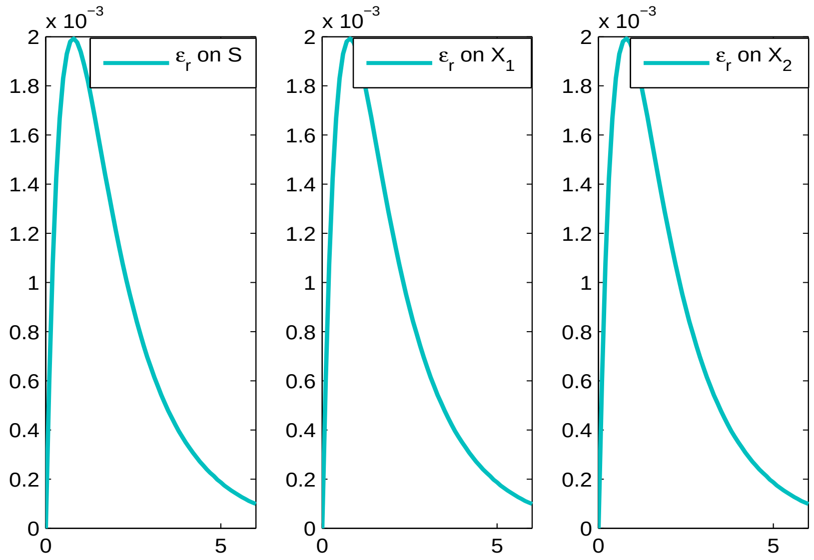



Figure 9.33: Relative error of the state trajectories of system (9.18) with $K_1$ and $K_2$.

Solving LMI 1d yields a state-feedback $K$ such that the resulting closed-loop system is state-invariant and such that the input trajectories have a starting boost at the initial condition which can be seen as a help to guarantee the input-invariance of the closed-loop system. We obtain the following state feedback $K$ :

$$K = \begin{bmatrix} -52.44 & 0.71008 & 1.0321 \\ 30.297 & -124.55 & 106.7 \\ 39.966 & 119.09 & -113.82 \end{bmatrix} \tag{9.19}$$

(which gives $\tilde{u} = K\,\tilde{x}$ directly optimal for the resulting LQ problem) such that the closed-loop matrix

$$A + B\,K = \begin{bmatrix} -91.456 & 0.062579 & 0.060861 \\ 34.99 & -25.031 & 21.34 \\ 59.83 & 23.818 & -22.819 \end{bmatrix}$$

is a Metzler matrix with

$$(A + B\,K)\,x_e = \begin{bmatrix} -1.7339 \\ -21.745 \\ -21.499 \end{bmatrix} \leq 0$$

and

$$K\,\tilde{x}_0 + u_e = \begin{bmatrix} 14.9302 \\ 11.2447 \\ 9.5236 \end{bmatrix} \geq 0.$$

Then, solving the LMI 2 gives the following weighting matrices such that $\tilde{u} = K\,\tilde{x}$ is optimal for the LQ problem of cost (9.15) :

$$\begin{aligned} Q &= \begin{bmatrix} 1034.6 & 138.79 & -74.251 \\ 138.79 & 240.3 & -214.15 \\ -74.251 & -214.15 & 195.52 \end{bmatrix}; \\ R &= \begin{bmatrix} 0.094638 & 00.050527 & 0.050030 \\ 0.050527 & 0.03.8781 & 0.030199 \\ 0.050030 & 0.03.0199 & 0.037606 \end{bmatrix}. \end{aligned} \tag{9.20}$$

These results give the following optimal control law $u_L = K\,\tilde{x} + u_e$, which is depicted in Figure 9.34 in comparison with the optimal control obtained with BMI 1b (drawn in Figure 9.30). The application of this optimal control to system (9.5) gives the resulting state trajectories for the nonlinear closed-loop system (9.18) drawn in Figure 9.35, in comparison with the state trajectories computed with BMI 1b (drawn in Figure 9.31). Moreover, the relative error between these two curves is depicted in Figure 9.36. One can observe that BMI 1b and LMI 1d give similar results.

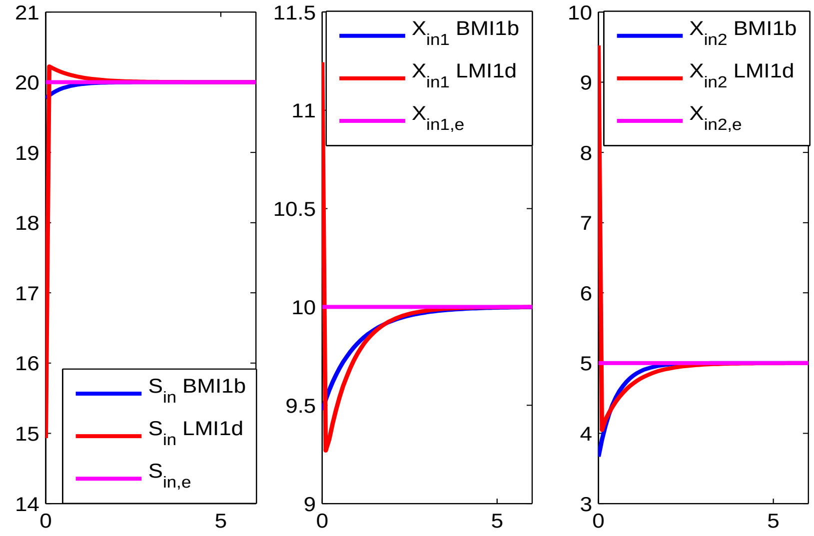



Figure 9.34: Optimal control $u_L = K\,\tilde{x} + u_e$ with $K$ given by (9.19) in comparison with (9.16).

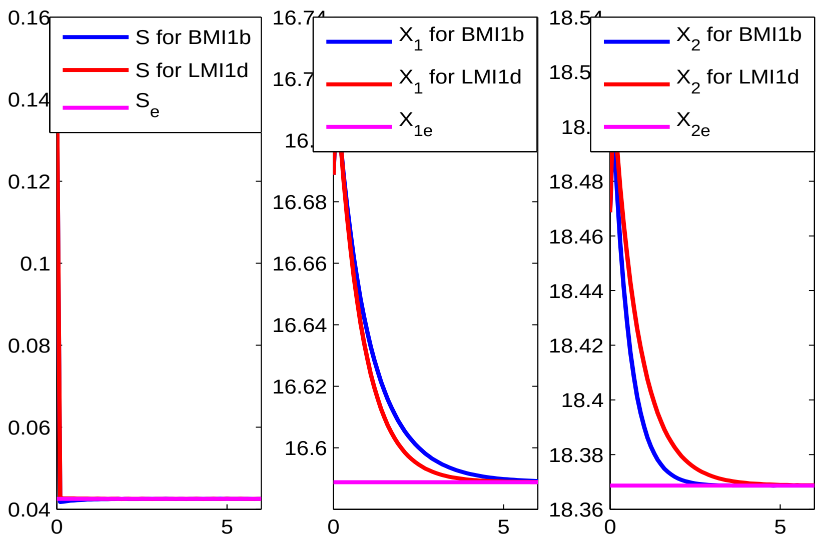



Figure 9.35: State trajectories for system (9.18) with $K$ given by (9.19) in comparison with (9.16).

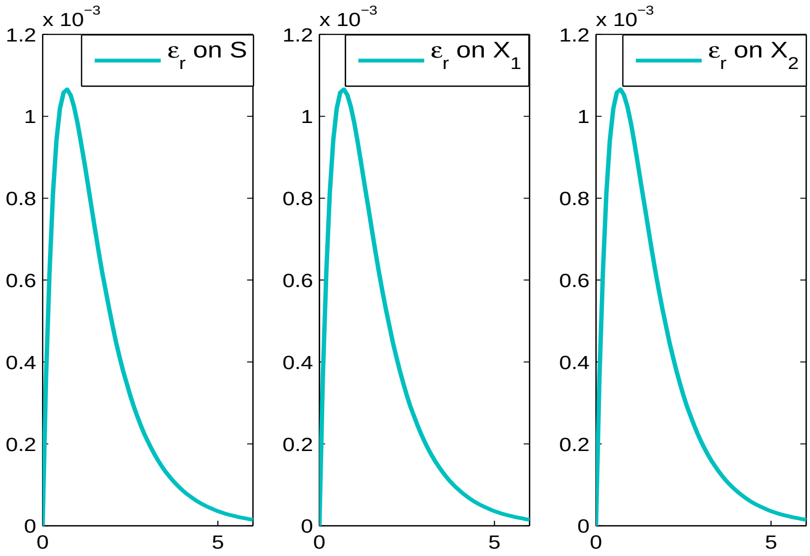



Figure 9.36: Relative error of the state trajectories for system (9.18) with $K$ given by (9.19) in comparison with (9.16).

Solving BMI 1b or LMI 1d gives a good matrix $K$ which guarantees the coexistence of the species and the admissibility of the input trajectories. Now using for example BMI 1b, an analysis can be realized on the variation of the initial state as done in the previous method. Consider the limit case of the previous method, $\tilde{x}_0 = [28 \quad 28 \quad 28]^T$. Recall that in this case the nonlinear state trajectories numerically diverge when solving a direct $\boldsymbol{LQ}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}^{\boldsymbol{t_f}}$ problem with a receding horizon approach. Here, using the state feedback $K$ given by (9.16), with the resolution of BMI 1b, leads to state trajectories which react very well. Therefore, by solving the $\boldsymbol{LQ}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}^{\mathrm{inv}}$ problem with different choices of $\tilde{x}_0$, the linear approximation of system (9.5) seems to be valid for $\tilde{x}_0 < [101 \quad 101 \quad 101]^T$ which corresponds to $x_0 < x_e + [101 \quad 101 \quad 101]^T$. See the corresponding state trajectories for $\tilde{x}_0 = [101 \quad 101 \quad 101]^T$ in Figure 9.37. One can observe the numerical divergence of the nonlinear state trajectories for $\tilde{x}_0 = [101 \quad 101 \quad 101]^T$ whereas the ones for $\tilde{x}_0 = [100 \quad 100 \quad 100]^T$ have a really good behavior.

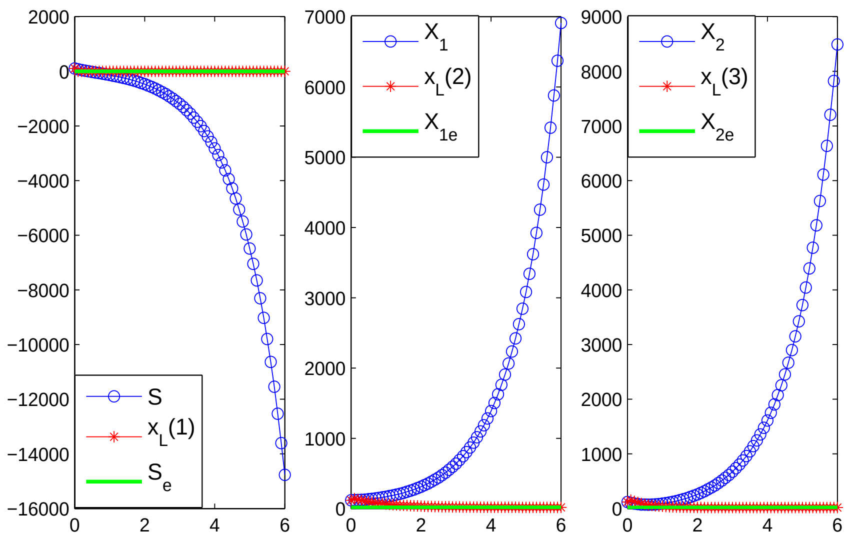



Figure 9.37: State trajectories for system (9.18) with $K$ given by BMI 1b with $\tilde{x}_0 = [101 \quad 101 \quad 101]^T$.

Finally, the robustness of the $\boldsymbol{LQ}^{\text{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem is numerically tested by applying perturbed values of the parameters (e.g. on the dilution rate and on the maximum growth rate value). First, we have observed in Figure 9.31, that by using BMI 1b together with LMI 2, the nonlinear state trajectories numerically converge to the equilibrium $x_e$ after time $6$. So, at this time, $t = 6$, a perturbation on the dilution rate $D$ or on the maximum growth rate $\mu_{\max,1}$ is introduced as previously. The resulting state and input trajectories are computed by using the same optimal state feedback $K$ given by (9.16). Figures 9.38 and 9.39 represent the input and state trajectories where at time $t = 6$, the dilution rate value has been changed to $D = 0.7$ instead of $0.2$. One can see that the optimal control has to be adapted in order to bring the state trajectories to the new equilibrium (corresponding to the equilibrium of system (9.18) with $D = 0.7$).

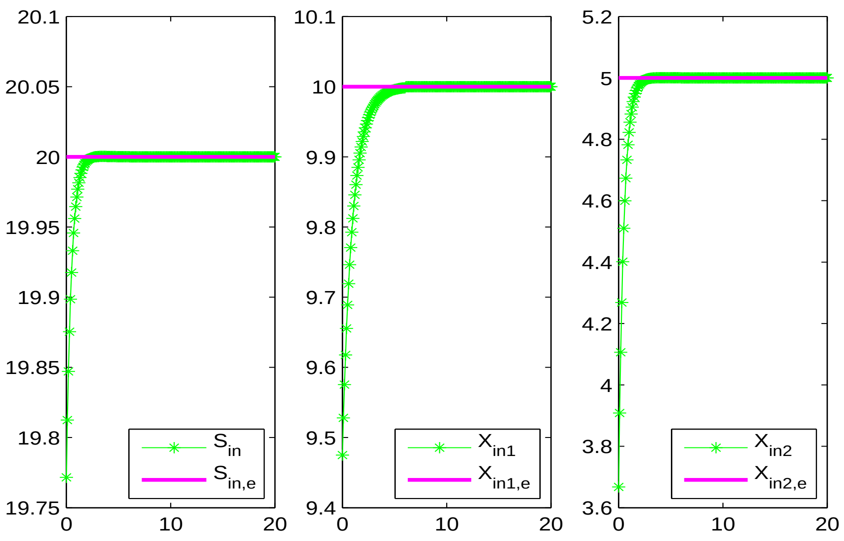



Figure 9.38: Perturbed optimal control $u_L = K\,\tilde{x} + u_e$ with $K$ given by (9.16).

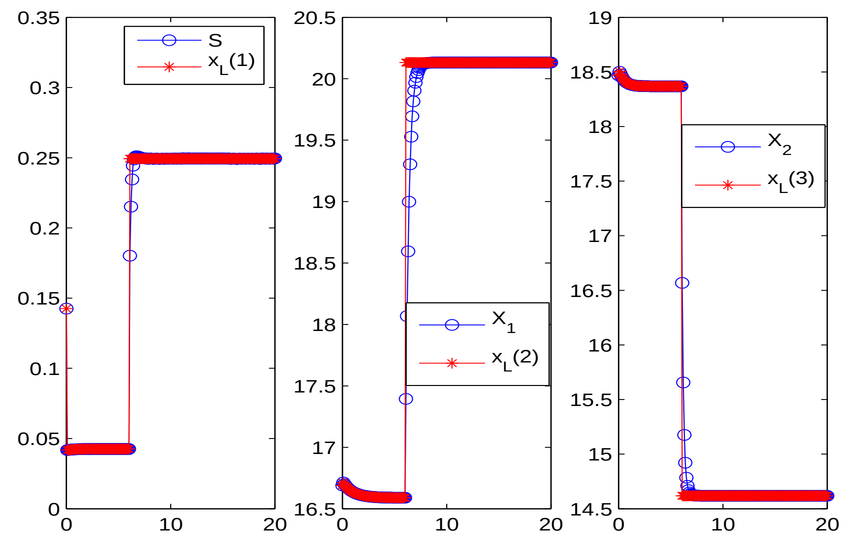



Figure 9.39: Perturbed state trajectories for system (9.18) with $K$ given by (9.16).

A similar analysis can be done by considering a small perturbation on the maximum growth rate $\mu_{\max,1}$. Figures 9.40 and 9.41 represent the input and state trajectories where at time $t = 6$, the maximum growth rate $\mu_{\max,1}$ has been changed to $\mu_{\max,1} = 1$ instead of $1.2$.

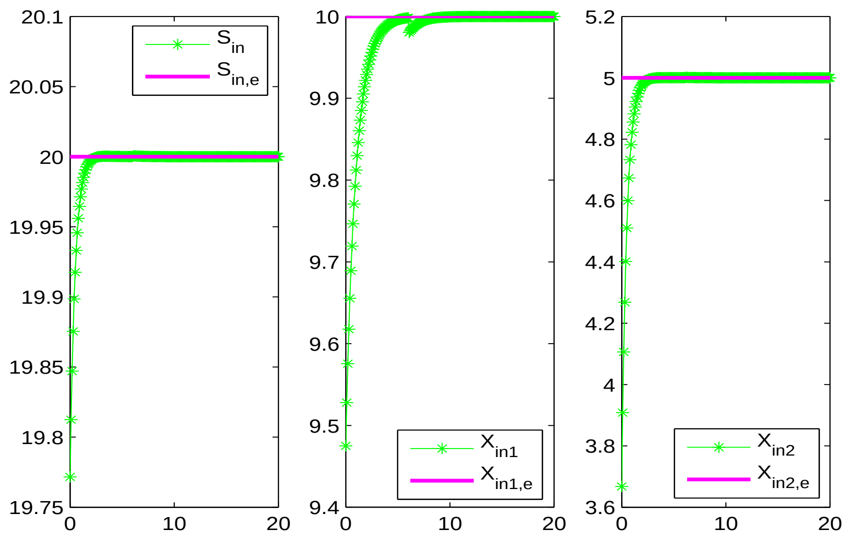



Figure 9.40: Perturbed optimal control $u_L = K\,\tilde{x} + u_e$ with $K$ given by (9.16).

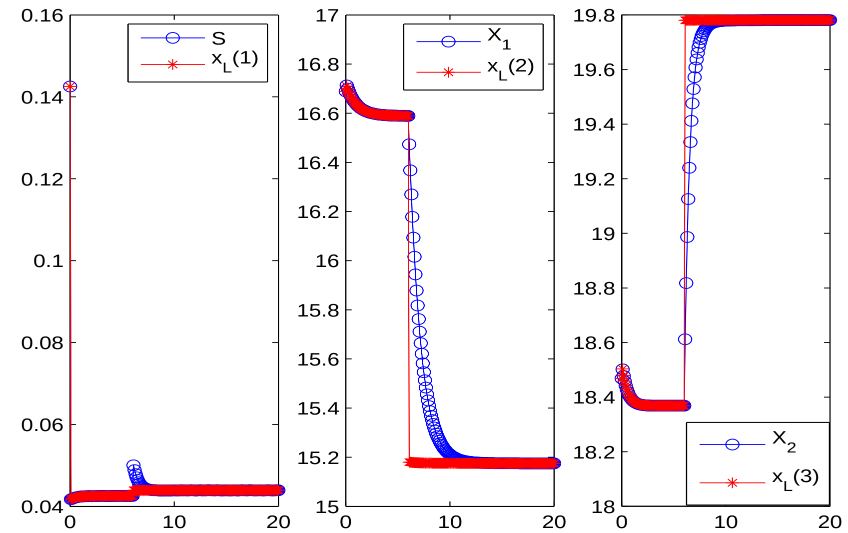



Figure 9.41: Perturbed state trajectories for system (9.18) with $K$ given by (9.16).

Again, one can observe that the optimal control numerically reacts to the perturbation in order to bring the state trajectories to the new equilibrium (corresponding to the equilibrium of system (9.18) with $\mu_{\max,1} = 1$) while always ensuring the coexistence of the species.

As a conclusion of this chapter, after all these numerical observations, we would like to make a comparison between the two methods used to guarantee the coexistence of the species. We tend to say that the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem seems to be more adapted to the problem of coexistence of species in a chemostat, for several reasons :

- the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem is in infinite horizon while for the other methodology, a receding horizon approach is needed ;
- the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem can deal with nonlinear systems, more specifically with the property of local positive input/state-invariance (by using Theorem 8.2.1) while ensures the local coexistence (by Proposition 9.2.1) ;
- the numerical simulations show that the neighborhood of the initial conditions for which the linear approximation of the chemostat model has the right behavior is larger in the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem than in the $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem.
- the numerical results also show that the coexistence of the species and the admissibility of the input trajectories are obtained for a smaller time with the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem whereas for the $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem a receding horizon approach is needed which implies an analysis for a large final time.

However, numerical simulations for the resolution of the $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem with the choice of CASE 2 for the input reveal that this case is harder to solve and has more difficulties to give good results than the resolution of the $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem. Furthermore, in the two methods, the robustness of the LQ problem has been numerically illustrated.

Finally, whatever the method used, we have shown in this chapter that it is possible to solve the problem of coexistence of species in a chemostat when the theory (with the competitive exclusion principle, see Theorem 9.1.1) states that there is only one winner in the competition.

# B. Discrete Time Case

# Chapter 10

# The Positive LQ Problem

In this last chapter, the finite-horizon linear quadratic optimal control problem with non-negative state constraints is studied for positive linear systems in discrete time. Necessary and sufficient optimality conditions are obtained by using the maximum principle. These conditions lead to a computational method for the solution of the positive LQ problem by means of the corresponding Hamiltonian system. In addition, necessary and sufficient conditions are proved for the positive LQ-optimal control to be given by the standard LQ-optimal state feedback law. In particular, such conditions are obtained for the problem of minimal energy control with penalization of the final state. Some results are direct adaptations of similar results for the continuous time case (see Chapter 4). Moreover, a positivity criterion for the LQ-optimal closed-loop system is derived specifically for positive discrete-time systems with a positively invertible (dynamics) generator which can be seen as an inverse time positive system. Monomial systems include the great class of compartmental systems (which are significant in applications, see e.g. [HCH10]). An algorithm is derived from the Hamiltonian system in order to compute a solution. Then the main results are illustrated by numerical examples.

The LQ problem with constraints has already been studied for positive linear systems in [CJ89] by using a controllable block companion transformation. Sufficient conditions on the weighting matrices of a quadratic cost criterion are derived to ensure that the closed-loop system is positive. This idea was generalized in [Joh94] in order to remove the restrictive positivity assumption that was required on such transformation.

## 10.1 Positive linear systems

First, let us recall some important results on positive linear systems in discrete time, see e.g. [FR00], [Ava00] or [HCH10] and the references therein. Consider the following linear time-invariant system description in discrete time, denoted by $[A,\, B]$ :

$$x_{i+1} = A\,x_i + B\,u_i,\ i = 0, \dots, N-1,\ \ x_0 = \hat{x}_0, \tag{10.1}$$

where the state $x_i$ and the control $u_i$ are in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, $A$ and $B$ are real matrices of compatible sizes, and $\hat{x}_0 \in \mathbb{R}^n$ denotes any fixed initial state.

**Definition 10.1.1** *The system $[A,\, B]$ given by (10.1) is said to be* **positive** *if, for all initial conditions $\hat{x}_0 \geq 0$ and for all controls $(u_i)_{i=0}^{N-1} \geq 0$, the state trajectories are nonnegative, i.e. for all $i = 0, \ldots,\, N$, $x_i \geq 0$.*

The following characterization of the positivity of discrete time systems is well-known (see e.g. [FR00], [Ka 02]).

**Proposition 10.1.1** *The system $[A,\, B]$ is positive if and only if $A$ and $B$ are nonnegative matrices.*

Now consider the following LTI homogeneous discrete time system

$$x_{i+1} = A\, x_i,\; i = 0, \ldots,\, N-1,\;\; x_0 = \hat{x}_0, \tag{10.2}$$

and recall the definition of stability for such systems and some useful results, see e.g. [CD91].

**Definition 10.1.2**

- *A LTI homogeneous system (10.2) is said to be* **asymptotically stable** *if for all $x_0 \in \mathbb{R}^n$, $x_i$ tends to zero as $i$ tends to infinity.*
- *A LTI homogeneous system (10.2) is said to be* **exponentially stable** *if there exist $\beta \in [0,\, 1[$ and $m > 0$ such that for all $i \geq 0$, $\|A^i\| \leq m\, \beta^i$.*

**Remark 10.1.1** *Recall that in the particular case of homogeneous time-invariant system, these two concepts of stability are equivalent, see [CD91]. Therefore in the sequel, the terms "asymptotic" and "exponential" are omitted.*

**Theorem 10.1.2 (Stability)** *A LTI homogeneous system (10.2) is stable if and only if all the eigenvalues of $A$ have a modulus strictly less than one, i.e.*

$$\forall\, \lambda \in \sigma(A)\; :\; |\lambda| < 1.$$

By using this result together with Theorem A.1.4 (Perron-Frobenius for nonnegative matrix), we obtain the following result on the stability of positive system :

**Theorem 10.1.3** *A positive LTI system (10.2) is stable if and only if its Frobenius eigenvalue $\rho(A)$ is less than one.*

Now, consider the following Lyapunov equation

$$A^T P\, A - P = -Q \tag{10.3}$$

where $A \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{n\times n}$ is symmetric positive definite and a unique symmetric positive definite solution $P$ has to be found for (10.3). The solvability of the Lyapunov equation relates directly to the exponential stability of system (10.2), see [CD91, pp. 214-216].

**Theorem 10.1.4** *A LTI homogeneous system (10.2) is stable if and only if for all symmetric positive definite matrices $Q$, the Lyapunov equation (10.3) has a unique symmetric positive definite solution $P$.*

In the case of positive systems, a stronger condition on the solution $P$ of the Lyapunov equation can be derived, namely the fact that $P$ is a diagonal matrix, see e.g. [FR00, p.41].

**Theorem 10.1.5** *A positive LTI system (10.2) is stable if and only if there exists a diagonal positive definite matrix $P$ such that the matrix $Q$, defined by*

$$-Q = A^T P\, A - P$$

*is positive definite.*

## 10.2 The positive LQ problem

This section is devoted to the LQ-optimal control problem for positive linear systems in discrete time. Some results are direct adaptations of similar results of the continuous time case and therefore their proofs will be omitted.

### 10.2.1 Problem statement

The finite horizon positive LQ problem in discrete time, which is denoted by $\boldsymbol{LQ_+^N}$, consists of minimizing the quadratic functional :

$$J(\hat{x}_0, (u_i)_{i=1}^{N-1}) := \frac{1}{2}\left(\sum_{i=0}^{N-1}(\|R^{1/2}u_i\|^2 + \|C\,x_i\|^2) + x_N^T S\, x_N\right) \tag{10.4}$$

for a given positive linear system described by (10.1), where the initial state $\hat{x}_0 \geq 0$ is fixed, under the constraints

$$\forall\, i \in \{0, \ldots, N\},\; x_i \geq 0, \tag{10.5}$$

where $N$ is a fixed final time, $R \in \mathbb{R}^{m\times m}$ is a symmetric positive definite matrix, $C \in \mathbb{R}^{p\times n}$ and $S \in \mathbb{R}^{n\times n}$ is a symmetric positive semidefinite matrix.

In other words, the $\boldsymbol{LQ_+^N}$ problem consists of minimizing a quadratic functional for a given positive system while requiring that the state trajectories be nonnegative for any fixed nonnegative initial state, whence the positivity property should be kept for the optimal state trajectories. In this framework, it is not required that the input function $(u_i)_{i=0}^{N-1}$ be nonnegative.

### 10.2.2 Optimality conditions

Assume that the inverse of $A$ exists. This assumption holds for example if $A$ comes from a discretization of a continuous-time system since $A = e^{A_c h}$, where $A_c$ denotes the matrix defining the continuous time system, see Appendix C. By applying the maximum principle in discrete time to this problem (see e.g. [HSV95]), i.e. the Karush-Kuhn-Tucker optimality conditions, the following discrete time version of Theorem 4.1.1 can be established for the $\boldsymbol{LQ_+^N}$-optimal control problem.

**Theorem 10.2.1 (Optimality conditions based on the Maximum Principle)**
*a) The $\boldsymbol{LQ_+^N}$ problem has a solution $(u_i)_{i=0}^{N-1}$ if and only if there exist multipliers $\lambda_i$ such that $u_i = -R^{-1}B^T p_i,\ i = 0, \ldots, N-1$, where $[x_i^T \quad p_i^T]^T \in \mathbb{R}^{2n}$ is the solution of the recurrent Hamiltonian equation*

$$\begin{bmatrix} x_i \\ p_i \end{bmatrix} = H \begin{bmatrix} x_{i+1} \\ p_{i+1} \end{bmatrix} - \begin{bmatrix} 0 \\ \lambda_i \end{bmatrix},\ i = N-1, \ldots, 0 \tag{10.6}$$

*with $x_0 = \hat{x}_0,\ p_N = S\,x_N - \lambda_N$, where*

$$H = \begin{bmatrix} A^{-1} & A^{-1}B\,R^{-1}B^T \\ C^TC\,A^{-1} & A^T + C^TCA^{-1}B\,R^{-1}B^T \end{bmatrix}$$

*is the Hamiltonian matrix, and for all $i = 0, \ldots, N$,*

$$x_i \geq 0, \tag{10.7}$$

$$\lambda_i \geq 0 \tag{10.8}$$

*and*

$$\lambda_i^T\,x_i = 0 \quad \textit{(complementarity condition)}. \tag{10.9}$$

*b) By using the matrix form of the recurrent Hamiltonian equation, $(u_i)_{i=0}^{N-1}$ is solution of the $\boldsymbol{LQ_+^N}$ problem if and only if there exist multiplier matrices $\Lambda_i$ such that $u_i = K_i(\hat{x}_0)\,x_i := -R^{-1}B^TY_i\,X_i^{-1}x_i,\ i = 0, \ldots, N-\ 1$, where*

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = H \begin{bmatrix} X_{i+1} \\ Y_{i+1} \end{bmatrix} - \begin{bmatrix} 0 \\ \Lambda_i \end{bmatrix},\ i = N-1, \ldots, 0$$

*with the final condition $X_N = I$ and $Y_N = S - \Lambda_N$, and for all $i = 0, \ldots, N$*

$$\Lambda_i\,X_0^{-1}\hat{x}_0 \geq 0, \tag{10.10}$$

$$\hat{x}_0^T X_0^{-T}\Lambda_i^T X_i\,X_0^{-1}\hat{x}_0 = 0 \quad \textit{(complementarity condition)} \tag{10.11}$$

*and*

$$X_i\,X_0^{-1}\hat{x}_0 \geq 0. \tag{10.12}$$

**Proof :** a) This result follows directly from the Karush-Kuhn-Tucker optimality conditions with state constraints (by using the discrete-time analogue of e.g. [HSV95, Theorem 4.1]), for necessity, and from the fact that the functional (10.4) is convex and the dynamics and inequality constraints (10.1) and (10.5) are defined by linear functions, for sufficiency.
b) This proof is a straightforward extension of the one of [CD91, Theorem 167, pp. 63-66]. The main fact is the invertibility of the matrices $X_i$, which can be proved by using an evaluation lemma, as in [CD91, Corollary 134, p. 61]. See also [Bea06, Chapter 5].

□

**Remark 10.2.1** *A priori, in view of conditions (10.10)-(10.12), the function $K_i(\hat{x}_0)$ in Theorem 10.2.1 (b) clearly depends upon the choice of the initial state $\hat{x}_0$. Stronger conditions are needed in order to make it independent of the initial state, i.e. such that the optimal control law is of the state feedback type $u_i = K_i\,x_i$. Such conditions are reported next.*

The following result follows easily from Lemma 4.1.3 (see the proof of Proposition 4.1.2).

**Proposition 10.2.2** *Conditions (10.10)-(10.12) are satisfied for all initial states $\hat{x}_0 \geq 0$ if and only if the following conditions hold for all $i = 0, \ldots, N$ :*

$$\Lambda_i X_0^{-1} \geq 0, \tag{10.13}$$

$$\Lambda_i^T X_i + X_i^T \Lambda_i = 0 \tag{10.14}$$

*and*

$$X_i X_0^{-1} \geq 0. \tag{10.15}$$

**Remarks 10.2.2** *a) Conditions (10.13)-(10.15) can be hard to check in general since the knowledge of $X_0$ is needed to check these conditions. However they obviously hold with $\Lambda_i = 0$ in an important particular case. See Corollary 10.2.3 below. Moreover, an algorithm is developed in Subsection 10.2.5 in order to make these conditions more computable.*
*b) The optimality conditions in Theorem 10.2.1 and Proposition 10.2.2 also hold for linear systems (10.1) that are not positive. However the positivity assumption plays a crucial role to obtain the criteria reported in Section 10.2.3.*

As in the continuous time problem, conditions can be obtained such that the $\boldsymbol{LQ_+^N}$ problem has a solution. Such conditions are based on the standard problem. The latter problem, denoted by $\boldsymbol{LQ^N}$ , consists of minimizing the quadratic functional (10.4) for a given positive linear system described by (10.1) (without any nonnegativity constraint on the state trajectory). Its solution is given by $u_i = K_i x_i = -R^{-1}B^T Y_i X_i^{-1} x_i,\ i = 0, \ldots, N-1$ where $[X_i^T \quad Y_i^T]^T \in \mathbb{R}^{2n \times n}$ is the solution of the matrix recurrent Hamiltonian equation, $i = 0, \ldots, N-1$,

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = H \begin{bmatrix} X_{i+1} \\ Y_{i+1} \end{bmatrix}, \ \begin{bmatrix} X_N \\ Y_N \end{bmatrix} = \begin{bmatrix} I \\ S \end{bmatrix}. \tag{10.16}$$

Equivalently the solution of the $\boldsymbol{LQ^N}$ problem is given, for all $i = 0, \ldots, N-1$, by

$$u_i = -R^{-1}B^T P_{i+1}[I + B\, R^{-1}B^T P_{i+1}]^{-1}\, x_i, \tag{10.17}$$

where $P_i$ is the solution of the Recurrent Riccati Equation (RRE), $i = N, \ldots, 1$, (see e.g. [CD91]) :

$$-P_{i-1} = C^T C + A^T P_i\, A - A^T P_i B (I + R^{-1} B^T P_i B)^{-1} R^{-1} B^T P_i A, \ \ P_N = S. \tag{10.18}$$

**Corollary 10.2.3 (Optimality conditions based on admissibility)** *The solution of the (standard) $\boldsymbol{LQ^N}$ problem is solution of the $\boldsymbol{LQ_+^N}$ problem for all $\hat{x}_0 \geq 0$ if and only if the $\boldsymbol{LQ^N}$-optimal state trajectories are admissible, i.e. nonnegative for all $\hat{x}_0 \geq 0$, or equivalently, one of the following equivalent conditions holds :*
*a) The standard closed-loop matrix $A + B\, K_i$ is nonnegative for all $i = 0, \ldots, N-1$, i.e.*

$$\forall\, k, l,\ \forall\, i = 0, \ldots, N-1,\ [B\, R^{-1} B^T P_i]_{kl} \leq a_{kl}. \tag{10.19}$$

*b) The matrix solution of the matrix recurrent Hamiltonian equation (10.16) is such that for all $i = 0, \ldots, N,\ \ X_i X_0^{-1} \geq 0$.*

**Proof :** Corollary 10.2.3 follows from Theorem 10.2.1 and Proposition 10.2.2 by applying the discrete time version of Theorem 2.1.1 (see also [Bea06]). In addition, the solution of the $\boldsymbol{LQ^N}$ problem is given as in Theorem 10.2.1 where the multiplier matrices $\Lambda_i$ are identically equal to zero. See the proof of Corollary 4.1.4.

□

### 10.2.3 Positivity Criteria

In this subsection, the $\boldsymbol{LQ^N}$ problem is studied with the aim of finding conditions on the problem data such that the standard closed-loop system is positive, i.e. such that the conditions of Corollary 10.2.3 hold. This can be interpreted as solving an inverse $\boldsymbol{LQ^N_+}$ problem. The criteria that are obtained here are specific to the discrete time case, except for Theorem 10.2.4, which is the discrete time version of Theorem 4.2.2.

**- Minimal energy control**

Consider the particular problem of minimal energy control with penalization of the final state, i.e. the $\boldsymbol{LQ^N}$ problem (10.1)-(10.4) where $C$ is equal to zero. By computing the expression of $P_i$ in terms of the matrix solution of the recurrent Hamiltonian equation, we obtain the following result, which is a direct adaptation of the continuous case, see Theorem 4.2.2.

**Theorem 10.2.4** *Consider the minimal energy* $\boldsymbol{LQ^N_+}$ *problem (10.1), (10.4)-(10.5), i.e. with* $C = 0$*. Let us denote* $\lambda_{\min}(R) := \min\{\lambda \,:\, \lambda \in \sigma(R)\}$*. Assume that* $A \gg 0$*. If the spectral radius* $\rho(S)$ *of the final state penalty matrix is sufficiently small such that*

$$\rho(S) = \max_{\mu_i \in \sigma(S)} \mu_i \quad < \quad \gamma := \begin{cases} \dfrac{\lambda_{\min}(R)\,(1-\sigma)}{\|B\|^2}, & \text{if } \sigma < 1 \\ \dfrac{\lambda_{\min}(R)\,(\sigma - 1)}{\|B\|^2\,\sigma^N}, & \text{if } \sigma > 1 \\ \dfrac{\lambda_{\min}(R)}{\|B\|^2 N}, & \text{if } \sigma = 1 \end{cases} \tag{10.20}$$

*where* $\sigma := \sigma_{\min}(A)\,\sigma_{\max}(A)$*, with* $\sigma_{\min}(A)$ *(*$\sigma_{\max}(A)$ *respectively) denoting the smallest (the largest respectively) singular value of* $A$*, then the* $\boldsymbol{LQ^N}$ *closed-loop system is positive and therefore the solution of the* $\boldsymbol{LQ^N}$ *problem is solution of the* $\boldsymbol{LQ^N_+}$ *problem.*

**Proof :** The positivity constraint on the closed-loop matrix can be written in terms of the solution $P_i$ of the RRE (see condition (10.19)), where $B \geq 0$. In addition, $P_i = Y_i\,X_i^{-1} = (A^T)^{N-i} S\;[I + G(N,i)\,S]^{-1}\,A^{N-i}$, where

$$G(N,\,i) := \sum_{k=0}^{N-i-1} (A^{-1})^{i-N+k+1} B\,R^{-1} B^T (A^T)^{N-i-k-1},$$

and, for $\sigma \neq 1$,

$$\|G(N,\,i)\,S\| \leq \frac{\|B\|^2}{\lambda_{\min}(R)}\,\rho(S)\,\Omega$$

where

$$\Omega := \sum_{k=0}^{N-i-1} \sigma^{N-i-k-1} = \frac{1-\sigma^{N-i}}{1-\sigma}.$$

Thus, if (10.20) holds, then

$$\|G(N,\, i)\, S\| \leq \frac{\rho(S)}{\gamma}$$

and

$$\|S\, [I + G(N,\, i)\, S]^{-1}\| \leq \frac{\rho(S)}{1 - \frac{\rho(S)}{\gamma}}.$$

Hence, by choosing $\rho(S)$ sufficiently small, condition (10.19) will hold, since $\forall\, k, l,\ a_{kl} > 0$ and the sequences $((A^T)^{N-i})_{i=0}^N$ and $(A^{N-i})_{i=0}^N$ are bounded.

□

**Remarks 10.2.3** *a) If $\sigma \geq 1$ and if the time horizon $N$ is increased, $\rho(S)$ has to be decreased accordingly for condition (10.20) to be satisfied with a fixed matrix $R$. As in the continuous time case, this reveals a tradeoff between positivity and stability of the closed-loop system in a receding horizon approach.*
*b) The minimal energy control problem with nonnegative controls and with a final state equality constraint is solved in [Ka 02, Subsection 3.4.1] for reachable systems. Here we use a penalization term in the cost instead of a final state constraint, it is not assumed that the system is reachable and it is not required that the input function $(u_i)_{i=0}^{N-1}$ be nonnegative.*

**- Nonnegative Hamiltonian matrix**

A positivity criterion based on the Hamiltonian matrix is stated. This result will be used in the next subsection.

**Theorem 10.2.5** *If the Hamiltonian matrix $H$ and the penalty matrix $S$ are nonnegative and if the solution of the matrix recurrent Hamiltonian equation is such that $X_0^{-1} \geq 0$, then the $\boldsymbol{LQ^N}$ closed-loop system is positive and therefore the solution of the $\boldsymbol{LQ^N}$ problem is solution of the $\boldsymbol{LQ_+^N}$ problem.*

**Proof :** Multiplying the matrix recurrent Hamiltonian equation (10.16) on the right by $X_0^{-1}$ gives

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} X_0^{-1} = H \begin{bmatrix} X_{i+1} \\ Y_{i+1} \end{bmatrix} X_0^{-1}$$

with

$$\begin{bmatrix} X_N\, X_0^{-1} \\ Y_N X_0^{-1} \end{bmatrix} = \begin{bmatrix} X_0^{-1} \\ S\, X_0^{-1} \end{bmatrix}.$$

It follows by induction that, for all $i = 0, \ldots, N,\ X_i\, X_0^{-1} \geq 0$. Then, by using Corollary 10.2.3 (b), one gets the conclusion.

□

**- Monomial systems**

**Definition 10.2.1** *A positive system $[A,\ B]$, described by (10.1), is said to be* **monomial** *if $A$ is a monomial matrix (see Appendix A.3) and $B$ is of the form* $B = \begin{bmatrix} \mathrm{diag}[b_i]_{i=1}^m \\ 0_{(n-m)\times m} \end{bmatrix}$.

For this particular class of systems, Theorem 10.2.5 leads to a positivity criteria for the $\boldsymbol{LQ_+^N}$ closed-loop system.

**Theorem 10.2.6** *Consider a monomial system described by (10.1) and the quadratic cost (10.4) where $C$, $R$ and $S$ are diagonal matrices. Then the $\boldsymbol{LQ^N}$ closed-loop system is positive and therefore the solution of the $\boldsymbol{LQ^N}$ problem is solution of the $\boldsymbol{LQ_+^N}$ problem.*

**Proof :** Let $A = D\,P$ with $D$ a positive definite diagonal matrix and $P$ a permutation matrix such that $P^T = P^{-1}$. Using Lemma A.3.1 with $C^TC$ a diagonal matrix and $P^{-1}$ a permutation matrix, one has $C^TC\,P^{-1} \overset{\mathbf{s}}{=} P^{-1}C^TC$ which implies that, by Definition A.3.2 and Remark A.3.1,

$$C^TC\,P^{-1} = P^{-1}C^TC\bar{D}$$

where $\bar{D}$ is a positive definite diagonal matrix. Now by using the explicit form of $H$ with $A = D\,P$, one gets :

$$\begin{aligned} H &= \begin{bmatrix} P^{-1}\,D^{-1} & P^{-1}\,D^{-1}\,B\,R^{-1}\,B^T \\ P^{-1}\,C^TC\,\bar{D}D^{-1} & P^{-1}\,D + P^{-1}\,C^TC\,\bar{D}\,D^{-1}\,B\,R^{-1}\,B^T \end{bmatrix} \\ &= \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix} \begin{bmatrix} D^{-1} & D^{-1}\,B\,R^{-1}\,B^T \\ C^TC\,\bar{D}D^{-1} & D + C^TC\,\bar{D}\,D^{-1}\,B\,R^{-1}\,B^T \end{bmatrix} \\ &= \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix} \begin{bmatrix} D_1 & D_2 \\ D_3 & D_4 \end{bmatrix}. \end{aligned}$$

Therefore $H$ is nonnegative since each term of the second matrix ($D_i$, $i = 1, \ldots 4$) is a diagonal and nonnegative matrix. Now, by using the matrix recurrent Hamiltonian equation with this expression of $H$, it can be shown by induction that, for all $i = 0, \ldots, N-1$, $X_i = (P^{-1})^{N-i}D_{X,i}$ and $Y_i = (P^{-1})^{N-i}D_{Y,i}$ where $D_{X,i}$ and $D_{Y,i}$ are positive definite diagonal matrices. Indeed, by inverse induction, first show that $X_{N-1}$ and $Y_{N-1}$ have this structure :

$$\begin{aligned} \begin{bmatrix} X_{N-1} \\ Y_{N-1} \end{bmatrix} &= H \begin{bmatrix} X_N \\ Y_N \end{bmatrix} = H \begin{bmatrix} I \\ S \end{bmatrix} \\ &= \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix} \begin{bmatrix} D_1 & D_2 \\ D_3 & D_4 \end{bmatrix} \begin{bmatrix} I \\ S \end{bmatrix} \\ &= \begin{bmatrix} P^{-1}\,(D_1 + D_2\,S) \\ P^{-1}\,(D_3 + D_4\,S) \end{bmatrix}. \end{aligned}$$

Then

$$\begin{cases} X_{N-1} = (P^{-1})^{N-(N-1)}\,(D_1 + D_2\,S) = (P^{-1})^{N-(N-1)}\,D_{X,N-1} \\ Y_{N-1} = (P^{-1})^{N-(N-1)}\,(D_3 + D_4\,S) = (P^{-1})^{N-(N-1)}\,D_{Y,N-1}. \end{cases}$$

Now assume that

$$\begin{cases} X_{i+1} &= (P^{-1})^{N-(i+1)}\, D_{X,\,i+1} \\ Y_{i+1} &= (P^{-1})^{N-(i+1)}\, D_{Y,\,i+1}. \end{cases}$$

And show that $X_i = (P^{-1})^{N-i} D_{X,i}$ and $Y_i = (P^{-1})^{N-i} D_{Y,i}$. Observe that

$$\begin{aligned} \begin{bmatrix} X_i \\ Y_i \end{bmatrix} &= H \begin{bmatrix} X_{i+1} \\ Y_{i+1} \end{bmatrix} = H \begin{bmatrix} (P^{-1})^{N-i-1}\, D_{X,\,i+1} \\ (P^{-1})^{N-i-1}\, D_{Y,\,i+1} \end{bmatrix} \\ &= \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix} \begin{bmatrix} D_1 & D_2 \\ D_3 & D_4 \end{bmatrix} \begin{bmatrix} (P^{-1})^{N-i-1}\, D_{X,\,i+1} \\ (P^{-1})^{N-i-1}\, D_{Y,\,i+1} \end{bmatrix} \\ &= \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix} \begin{bmatrix} D_1\,(P^{-1})^{N-i-1}\, D_{X,\,i+1} + D_2\,(P^{-1})^{N-i-1}\, D_{Y,\,i+1} \\ D_3\,(P^{-1})^{N-i-1}\, D_{X,\,i+1} + D_4\,(P^{-1})^{N-i-1}\, D_{Y,\,i+1} \end{bmatrix} \end{aligned}$$

with $D_i\,(P^{-1})^{N-i-1} \overset{\mathbf{s}}{=} (P^{-1})^{N-i-1}\, D_i$ for $i = 1, \ldots, 4$ by Lemma A.3.1. Then, applying Remark A.3.1 gives

$$D_i\,(P^{-1})^{N-i-1} = (P^{-1})^{N-i-1}\, D_i\, \bar{D}$$

where $\bar{D}$ is a positive definite diagonal matrix. Defining

$$\begin{cases} \widetilde{D}_X &:= \bar{D}\, D_{X,\,i+1}, \\ \widetilde{D}_Y &:= \bar{D}\, D_{Y,\,i+1}, \end{cases}$$

leads to the following identity :

$$\begin{aligned} \begin{bmatrix} X_i \\ Y_i \end{bmatrix} &= \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix} \begin{bmatrix} (P^{-1})^{N-i-1} & 0 \\ 0 & (P^{-1})^{N-i-1} \end{bmatrix} \begin{bmatrix} D_1\, \widetilde{D}_X + D_2\, \widetilde{D}_Y \\ D_3\, \widetilde{D}_X + D_4\, \widetilde{D}_Y \end{bmatrix} \\ &= \begin{bmatrix} (P^{-1})^{N-i} & 0 \\ 0 & (P^{-1})^{N-i} \end{bmatrix} \begin{bmatrix} D_1\, \widetilde{D}_X + D_2\, \widetilde{D}_Y \\ D_3\, \widetilde{D}_X + D_4\, \widetilde{D}_Y \end{bmatrix} \end{aligned}$$

Then

$$\begin{cases} X_i &= (P^{-1})^{N-i}\, D_{X,\,i} \\ Y_i &= (P^{-1})^{N-i}\, D_{Y,\,i} \end{cases}$$

where

$$\begin{cases} D_{X,\,i} &= D_1\, \widetilde{D}_X + D_2\, \widetilde{D}_Y \\ D_{Y,\,i} &= D_3\, \widetilde{D}_X + D_4\, \widetilde{D}_Y \end{cases}$$

are positive definite diagonal matrices. In particular, $X_0 = (P^{-1})^N\, D_{X,\,0}$ where $D_{X,\,0}$ is a positive definite diagonal matrix. Hence $X_0$ is a monomial matrix and $X_0^{-1} \geq 0$. It follows by Theorem 10.2.5 that the $\boldsymbol{LQ^N}$ closed-loop system is positive.

□

Note that if $A$ is a diagonal matrix, Theorem 10.2.6 obviously holds. Moreover, Theorem 10.2.6 can be readily extended to the infinite horizon problem, see [Bea06].

### 10.2.4 Numerical examples

Consider the $\boldsymbol{LQ_+^N}$ problem for the positive system described by

$$x_{i+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x_i + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_i, \quad i = 0, \ldots, N-1. \tag{10.21}$$

with the cost $J(\hat{x}_0,\ (u_i)_{i=0}^{N-1}) = \dfrac{1}{2}\left[\sum_{i=0}^{N-1} \|u_i\|^2 + x_N^T S\, x_N\right]$, under the constraints $\forall i = 0, \ldots, N$, $x_i^1 \geq 0$ and $x_i^2 \geq 0$, where $x_i^j$ denotes the $j^{\text{th}}$ component of $x_i$. Notice that the matrices defining the system (10.21) are the same as those used for the numerical example of the positive $\boldsymbol{LQ_+^{t_f}}$ problem in continuous time, see Section 4.3. Here, the numerical example is treated in its discrete time version. Let the final state penalty matrix be given by

$$S = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{10.22}$$

and $N = 20$. By computing the solution of the $\boldsymbol{LQ^N}$ problem by means of the recurrent Hamiltonian equation, we obtain that the optimal state trajectories are not nonnegative. This means that the solution of the $\boldsymbol{LQ^N}$ problem is not admissible for the $\boldsymbol{LQ_+^N}$ problem. See Figure 10.1 where the optimal state trajectories are drawn for the initial condition $\hat{x}_0 = [1\ \ 0]^T$.

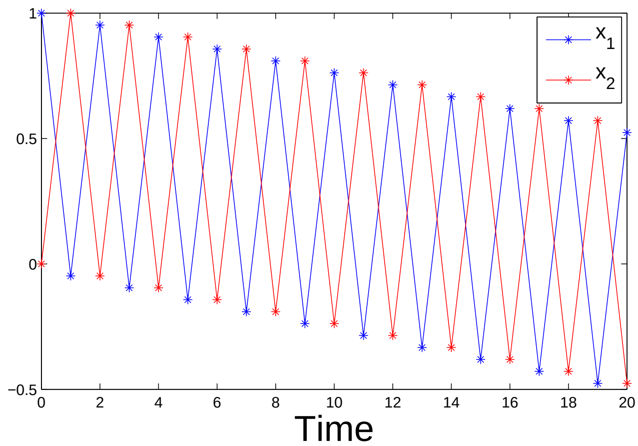



Figure 10.1: State trajectories for system (10.21) with $S$ given by (10.22).

The numerical solution of the $\boldsymbol{LQ_+^N}$ problem has been computed for the fixed initial condition $\hat{x}_0 = [1\ \ 0]^T$ by using `Matlab` and the particular function `quadprog`. This optimization algorithm leads to the optimal control depicted in Figure 10.2. The corresponding state $x_i(t)$ and multiplier $\lambda_i(t)$ trajectories are depicted in Figure 10.3.

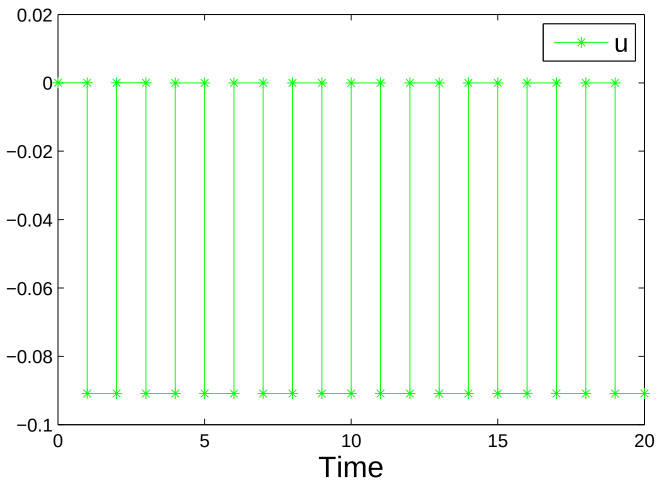



Figure 10.2: Optimal control for system (10.21) with $S$ given by (10.22).

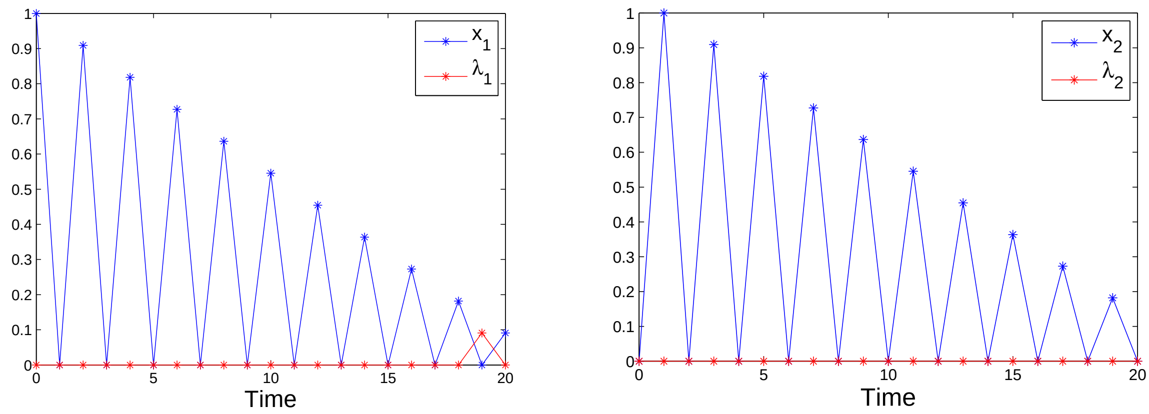



Figure 10.3: State trajectories and associated multipliers for system (10.21) with $S$ given by (10.22).

As was to be expected, the optimal state trajectories are numerically verified to be nonnegative for all time. In this case, we obtain a solution for a fixed initial condition.
To get the solution of the $\boldsymbol{LQ}_+^{\boldsymbol{N}}$ problem for any initial condition, with an appropriate choice of the penalty matrix $S$, we can use the positivity criterion of Subsection 10.2.3. Obviously, system (10.21) is monomial and the cost verifies the conditions of Theorem 10.2.6 with $S$ equal to any diagonal matrix. Hence, $S$ can be chosen in order to improve the stability of the closed-loop system while ensuring its positivity and the optimal control is of state-feedback form. Let us consider $S = I_2$ and compute the matrix solution $[X_i^T \quad Y_i^T]^T$ of the Hamiltonian equation (10.16). Then, computing the optimal control $u_i = K_i\, x_i$ (given by (10.17)) and the state trajectories $x_i = X_i\, X_0^{-1}\, \hat{x}_0$, leads to Figures 10.4 and 10.5, respectively, with the initial condition $\hat{x}_0 = [1 \;\; 0]^T$.

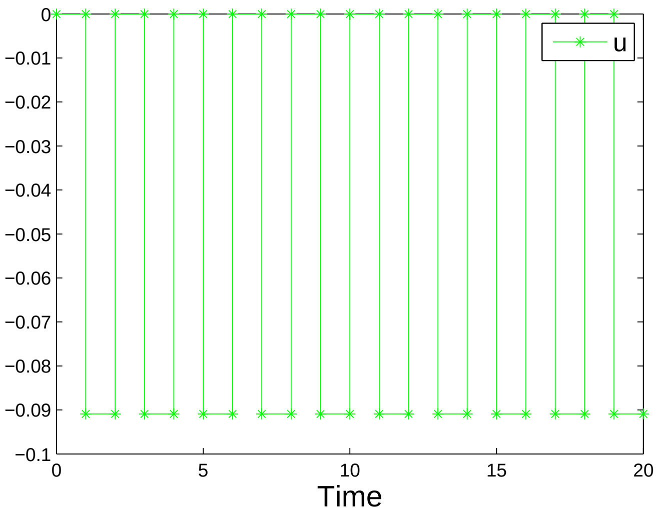



Figure 10.4: Optimal control for system (10.21) with $S = I$.

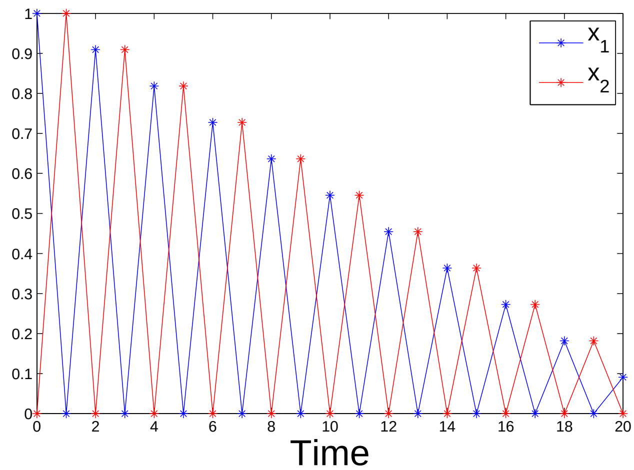



Figure 10.5: State trajectories for system (10.21) with $S = I$.

Let us mention that the feedback $K_i$ is given as follows :

| $K_0$ | $K_1$ | $K_2$ | $\cdots$ | $K_{18}$ | $K_{19}$ | $K_{20}$ |
|---|---|---|---|---|---|---|
| $[0 \quad -0.09]$ | $[0 \quad -0.09]$ | $[0 \quad -0.1]$ | $\cdots$ | $[0 \quad -0.333]$ | $[0 \quad -0.5]$ | $[0 \quad -0.5]$ |

Table 10.1: Feedback $K_i$ of system (10.21) with $S = I_2$.

Other illustrative examples on the discrete time problem have been done in [Bea06], wherein a receding approach is also studied for the infinite horizon positive LQ problem.

### 10.2.5 Computational method

– <u>MATRIX ALGORITHM :</u>

This subsection is devoted to the design of an algorithm which computes the solution $[X_i^T \quad Y_i^T \quad \Lambda_i^T]^T$ of the recurrent Hamiltonian equation satisfying conditions (10.13)-(10.15) by using the solution computed at the previous step only, that is $[X_{i+1}^T \quad Y_{i+1}^T \quad \Lambda_{i+1}^T]^T$. One way to proceed is to rewrite conditions (10.13) and (10.15). Observe that (10.15) can be developed as :

$$X_i \underbrace{X_{i-1}^{-1} X_{i-1} X_{i-2}^{-1} X_{i-2} \ \ldots \ X_1^{-1} X_1}_{=I_n} X_0^{-1} \geq 0$$

If it is assumed that

$$X_i X_{i-1}^{-1} \geq 0 \quad \text{for all } i = 0, \ldots, N \tag{10.23}$$

(instead of $X_i X_0^{-1} \geq 0$), then

$$\underbrace{X_i X_{i-1}^{-1}}_{\geq 0} \underbrace{X_{i-1} X_{i-2}^{-1}}_{\geq 0} X_{i-2} \ \ldots \ X_1^{-1} \underbrace{X_1 X_0^{-1}}_{\geq 0} \geq 0.$$

Now condition (10.13) can be written as

$$\Lambda_i \underbrace{X_{i-1}^{-1} X_{i-1} X_{i-2}^{-1} X_{i-2} \ \ldots \ X_1^{-1} X_1}_{=I_n} X_0^{-1} \geq 0.$$

If, in addition, it is assumed that

$$\Lambda_i X_{i-1}^{-1} \geq 0 \quad \text{for all } i = 0, \ldots, N \tag{10.24}$$

(instead of $\Lambda_i X_0^{-1} \geq 0$), then

$$\underbrace{\Lambda_i X_{i-1}^{-1}}_{\geq 0} \underbrace{X_{i-1} X_{i-2}^{-1}}_{\geq 0} X_{i-2} \ \ldots \ X_1^{-1} \underbrace{X_1 X_0^{-1}}_{\geq 0} \geq 0.$$

An algorithm using conditions (10.23) and (10.24) instead of (10.13) and (10.15) can then be described. See Algorithm 1, where the Hamiltonian matrix $H$ is decomposed as follows

$$H = \begin{bmatrix} A^{-1} & A^{-1} B R^{-1} B^T \\ C^T C A^{-1} & A^T + C^T C A^{-1} B R^{-1} B^T \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \tag{10.25}$$

**ALGORITHM 1**

1. `Init` :
   - Let $X_N = I_n$
   - Solve $\Lambda_N^T + \Lambda_N = 0$ such that $X_{N-1} = H_{11} + H_{12}\,(S - \Lambda_N)$ and
     (a) $X_{N-1}^{-1} \geq 0$
     (b) $\Lambda_N\, X_{N-1}^{-1} \geq 0$
   - Compute $Y_N = S - \Lambda_N$
2. `For` $i = N-1, \dots, 1$ :
   - Compute $X_i = H_{11}\, X_{i+1} + H_{12}\, Y_{i+1}$
   - Compute $M_i = H_{11}\, X_i + H_{12}\,(H_{21}\, X_{i+1} + H_{22}\, Y_{i+1})$
   - Solve $\Lambda_i^T X_i + X_i^T \Lambda_i = 0$ such that
     (a) $X_i\,[M_i - H_{12}\,\Lambda_i]^{-1} \geq 0 \quad \rightsquigarrow \quad X_i\, X_{i-1}^{-1} \geq 0$
     (b) $\Lambda_i\,[M_i - H_{12}\,\Lambda_i]^{-1} \geq 0 \quad \rightsquigarrow \quad \Lambda_i\, X_{i-1}^{-1} \geq 0$
   - Compute $Y_i = H_{21}\, X_{i+1} + H_{22}\, Y_{i+1} - \Lambda_i$
3. `End` :
   - Compute $X_0 = H_{11}\, X_1 + H_{12}\, Y_1$
   - Solve $\Lambda_0^T X_0 + X_0^T \Lambda_0 = 0$ such that $\Lambda_0\, X_0^{-1} \geq 0$
   - Compute $Y_0 = H_{21}\, X_1 + H_{22}\, Y_1 - \Lambda_0$

To illustrate this algorithm, consider the following positive system described by

$$x_{i+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} x_i + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u_i \tag{10.26}$$

and the cost (10.4) where

$$C = 0_2\,;\; S = I_2 \;\text{ and }\; R = 1. \tag{10.27}$$

Assume that $N = 4$. The successive steps of Algorithm 1 can be summarized as follows, with the Hamiltonian matrix given by

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{10.28}$$

1. <u>`Init`</u> :

   - $X_4 = I_2$.
   - Solve $\Lambda_4^T + \Lambda_4 = 0$.
     Then $\Lambda_4$ is a skew-symmetric matrix of the form :

     $$\Lambda_4 = \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}$$

     with an arbitrarily parameter $\alpha \in \mathbb{R}$ such that

     $$X_3 = H_{11}\, X_4 + H_{12}(-\Lambda_4) = \begin{pmatrix} 0 & 1 \\ 2 & -\alpha \end{pmatrix}$$

     verifies :

     (a) $X_3^{-1} = \begin{bmatrix} \frac{1}{2}\,\alpha & \frac{1}{2} \\ 1 & 0 \end{bmatrix} \geq 0 \;\Leftrightarrow\; \alpha \geq 0$

     (b) $\Lambda_4\, X_3^{-1} = \begin{bmatrix} \alpha & 0 \\ -\frac{1}{2}\,\alpha^2 & -\frac{1}{2}\,\alpha \end{bmatrix} \geq 0$ with $X_3^{-1} \geq 0 \;\Leftrightarrow\; \alpha = 0.$

     These conditions implies that the parameter $\alpha$ is equal to zero and then $\Lambda_4 = 0_2$ and $X_3 = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$.
   - Compute $Y_4 = S - \Lambda_4 = I_2$.

2. <u>`For` $i = 3$</u> :

   - $X_3 = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$.
   - Compute $M_3 = H_{11}\, X_3 + H_{12}\,(H_{21}\, X_4 + H_{22}\, Y_4) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.
   - Let $\Lambda_3 = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$. Solve $\Lambda_3^T X_3 + X_3^T \Lambda_3 = \begin{bmatrix} 4\,a_3 & a_1 + 2\,a_4 \\ a_1 + 2\,a_4 & 2\,a_2 \end{bmatrix} = 0$, that is $a_1 = -2\,a_4,\; a_2 = 0,\; a_3 = 0$ with $a_4 \in \mathbb{R}$ such that

     (a) $X_3\,[M_3 - H_{12}\,\Lambda_3]^{-1} = \begin{bmatrix} -\frac{1}{4}a_4 & \frac{1}{2} \\ 1 & 0 \end{bmatrix} \geq 0 \quad \rightsquigarrow \; a_4 \leq 0$

     (b) $\Lambda_3\,[M_3 - H_{12}\,\Lambda_3]^{-1} = \begin{bmatrix} -\frac{1}{2}a_4 & 0 \\ -\frac{1}{4}a_4^2 & \frac{1}{2}a_4 \end{bmatrix} \geq 0 \quad \rightsquigarrow \; a_4 = 0$

     Then $\Lambda_3 = 0_2$.
   - Compute $Y_3 = H_{21}\, X_4 + H_{22}\, Y_4 - \Lambda_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

3. <u>`For` $i = 2$</u>:

   - Compute $X_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.

   - Compute $M_2 = H_{11}\, X_2 + H_{12}\, (H_{21}\, X_3 + H_{22}\, Y_3) = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix}$.

   - Let $\Lambda_2 = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$. Solve $\Lambda_2^T X_2 + X_2^T \Lambda_2 = \begin{bmatrix} 4\, b_1 & 2\, b_3 + 2\, b_2 \\ 2\, b_3 + 2\, b_2 & 4\, b_4 \end{bmatrix} = 0$, that is $b_1 = 0,\ b_2 = -b_3,\ b_4 = 0$ with $b_3 \in \mathbb{R}$ such that

     (a) $X_2\, [M_2 - H_{12}\, \Lambda_2]^{-1} = \begin{bmatrix} -\frac{1}{3} b_3 & \frac{2}{3} \\ 1 & 0 \end{bmatrix} \geq 0 \quad \rightsquigarrow \ b_3 \leq 0$

     (b) $\Lambda_2\, [M_2 - H_{12}\, \Lambda_2]^{-1} = \begin{bmatrix} -\frac{1}{2}\, b_3 & 0 \\ -\frac{1}{6} b_3^2 & \frac{1}{3} b_3 \end{bmatrix} \geq 0 \quad \rightsquigarrow \ b_3 = 0$

     Then $\Lambda_2 = 0_2$.

   - Compute $Y_2 = H_{21}\, X_3 + H_{22}\, Y_3 - \Lambda_2 = I_2$.

4. <u>`For` $i = 1$</u>:

   - Compute $X_1 = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix}$.

   - Compute $M_1 = H_{11}\, X_1 + H_{12}\, (H_{21}\, X_2 + H_{22}\, Y_2) = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

   - Let $\Lambda_1 = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$. Solve $\Lambda_1^T X_1 + X_1^T \Lambda_1 = \begin{bmatrix} 6\, c_3 & 2\, c_1 + 3\, c_4 \\ 2\, c_1 + 3\, c_4 & 4\, c_2 \end{bmatrix} = 0_2$, that is $c_1 = -\frac{3}{2}\, c_4,\ c_2 = 0,\ c_3 = 0$ with $c_4 \in \mathbb{R}$ such that

     (a) $X_1\, [M_1 - H_{12}\, \Lambda_1]^{-1} = \begin{bmatrix} -\frac{1}{3} c_4 & \frac{2}{3} \\ 1 & 0 \end{bmatrix} \geq 0 \quad \rightsquigarrow \ c_4 \leq 0$

     (b) $\Lambda_1\, [M_1 - H_{12}\, \Lambda_1]^{-1} = \begin{bmatrix} -\frac{1}{2} c_4 & 0 \\ -\frac{1}{6} c_4^2 & \frac{1}{3} c_4 \end{bmatrix} \geq 0 \quad \rightsquigarrow \ c_4 = 0$

     Then $\Lambda_1 = 0_2$.

   - Compute $Y_1 = H_{21}\, X_2 + H_{22}\, Y_2 - \Lambda_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

5. <u>`End`</u>:

   - Compute $X_0 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

   - Let $\Lambda_0 = \begin{bmatrix} d_1 & d_2 \\ d_3 & d_4 \end{bmatrix}$. Solve $\Lambda_0^T X_0 + X_0^T \Lambda_0 = \begin{bmatrix} 6\, d_1 & 3\, d_3 + 3\, d_2 \\ 3\, d_3 + 3\, d_2 & 6\, d_4 \end{bmatrix} = 0_2$ that is $d_1 = 0,\ d_2 = -d_3,\ d_4 = 0$ with $d_3 \in \mathbb{R}$ such that
     $\Lambda_0\, X_0^{-1} = \begin{bmatrix} 0 & \frac{1}{3} d_2 \\ \frac{1}{3} d_3 & 0 \end{bmatrix} \geq 0 \quad \rightsquigarrow \ d_2,\, d_3 \geq 0$

Then $\Lambda_0 = \begin{bmatrix} 0 & d_2 \\ d_3 & 0 \end{bmatrix}$ with $d_2,\ d_3 \in \mathbb{R}$ such that $d_2,\ d_3 \geq 0$.

- Compute $Y_0 = H_{21} - \Lambda_0 = \begin{bmatrix} 1 & -d_2 \\ -d_3 & 1 \end{bmatrix}$.

Now let $x_0 = [1 \quad 0]^T$ and compute the state trajectories $x_i$ and the associated optimal control $u_i$ by using the link between the vector and matrix form of the recurrent Hamiltonian equation (obtained as in the continuous time case, see equation (3.20) in the proof of Theorem 3.2.1 b)) :

$$\begin{bmatrix} x_i \\ p_i \\ \lambda_i \end{bmatrix} := \begin{bmatrix} X_i \\ Y_i \\ \Lambda_i \end{bmatrix} X_0^{-1} x_0. \tag{10.29}$$

and

$$u_i = -R^{-1} B^T p_i.$$

The resulting solutions $x_i,\ \lambda_i,\ p_i$ and $u_i$ are given in Table 10.2 :

| | $i=0$ | $i=1$ | $i=2$ | $i=3$ | $i=4$ |
|---|---|---|---|---|---|
| $x_i$ | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} \frac{2}{3} \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ \frac{2}{3} \end{bmatrix}$ | $\begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix}$ |
| $\lambda_i$ | $\begin{bmatrix} 0 \\ \frac{1}{3}d_3 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ |
| $p_i$ | $\begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3}d_3 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ \frac{1}{3} \end{bmatrix}$ | $\begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ \frac{1}{3} \end{bmatrix}$ | $\begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix}$ |
| $u_i$ | $-\frac{1}{3}$ | $0$ | $-\frac{1}{3}$ | $0$ | $-\frac{1}{3}$ |

Table 10.2: Solutions of system (10.26) with (10.27) by applying Algorithm 1.

This example shows how the algorithm can be applied in order to obtain the optimal state and input trajectories. However, the algorithm is applied here with a very small value of $N$. In fact, increasing the final horizon $N$ gives, either calculations which are quickly very complicated to solve, or conditions which are incompatible, notably conditions (a) and (b) in the loop. Another algorithm is developed below by using the conditions of Theorem 10.2.1 and the vector recurrent Hamiltonian equation (10.6).

**Remark 10.2.4** *In order to check the results of Table 10.2, the solution of the* $\boldsymbol{LQ_+^N}$ *problem is computed by using the optimization algorithm* `quadprog` *in* `Matlab` *with* $N = 4$*. Moreover, let us remark that system (10.26), with* $C$*,* $S$ *and* $R$ *given by (10.27), is the monomial system (10.21) which is studied in Subsection 10.2.4. Therefore, the solution can be computed by solving the standard solution of the Hamiltonian equation (10.16) with* $N = 4$*. This two different ways to compute the solution lead to results given in Table 10.2, with the multipliers* $\lambda_i = 0$ *for all* $i$*.*

– VECTOR ALGORITHM :

Here, the solution of the $\boldsymbol{LQ_+^N}$ problem is computed by means of the vector recurrent Hamiltonian equation (10.6). An algorithm can be described as follows. See Algorithm 2, with the decomposition of $H$ as in (10.25).

ALGORITHM 2

1. `Init` :
   - Fix the final time $N$
   - Choose $x_N \geq 0$
   - Compute $\lambda_N = \begin{bmatrix} \alpha_N^1 \\ \vdots \\ \alpha_N^n \end{bmatrix}$ such that
     (a) $\lambda_N^T x_N = 0$
     (b) $\lambda_N \geq 0$
   - Compute $p_N = Sx_N - \lambda_N$
2. `For` $i = N-1, \ldots, 1$ :
   - Compute $x_i = H_{11}\, x_{i+1} + H_{12}\, p_{i+1}$
   - Compute $\lambda_i$ such that
     (a) $\lambda_i^T x_i = 0$
     (b) $\lambda_i \geq 0$
     (c) $x_i \geq 0$
   - Compute $p_i = H_{21}\, x_{i+1} + H_{22}\, p_{i+1} - \lambda_i$
3. `End` :
   - Compute $x_0 = H_{11}\, x_1 + H_{12}\, p_1$
   - Compute $\lambda_0$ such that
     (a) $\lambda_0^T x_0 = 0$
     (b) $\lambda_0 \geq 0$
     (c) $x_0 \geq 0$
   - Compute $p_0 = H_{21}\, x_1 + H_{22}\, p_1 - \lambda_0$
   - Check that $x_i$ and $\lambda_i$ are such that
     (a) $\lambda_i^T x_i = 0$
     (b) $\lambda_i \geq 0$ for $i = 0, \ldots, N$
     (c) $x_i \geq 0$

     with a good choice of parameters $\alpha_i^j$ for $i = 0, \ldots, N$ and $j = 1, \ldots, n$.

Consider again system (10.26) to illustrate this algorithm, i.e.

$$x_{i+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} x_i + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u_i,$$

and the cost (10.4) where

$$C = 0_2,\ R = 1,\ S = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \tag{10.30}$$

Let us apply Algorithm 2 to system (10.26), (10.30) :

1. `Init` :
   - Let $N = 4$ with $n = 2$.
   - Choose $x_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
   - Compute $\lambda_4 = \begin{bmatrix} \alpha_4^1 \\ \alpha_4^2 \end{bmatrix}$ such that

     (a) $\lambda_4^T x_4 = 0 \ \Leftrightarrow\ \alpha_4^1 = 0$

     (b) $\lambda_4 \geq 0 \ \Leftrightarrow\ \lambda_4 = \begin{bmatrix} 0 \\ \alpha_4^2 \end{bmatrix} \geq 0$ with $\alpha_4^2 \geq 0$.
   - Compute $p_4 = S x_4 - \lambda_4 = \begin{bmatrix} 1 \\ 1 - \alpha_4^2 \end{bmatrix}$.
2. `For` $i = 3$ :
   - Compute $x_3 = H_{11}\, x_4 + H_{12}\, p_4 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$.
   - Compute $\lambda_3 = \begin{bmatrix} \alpha_3^1 \\ \alpha_3^2 \end{bmatrix}$ such that

     (a) $\lambda_3^T x_3 = 0 \ \Leftrightarrow\ \alpha_3^2 = 0$

     (b) $\lambda_3 \geq 0 \ \Leftrightarrow\ \alpha_3^1 \geq 0$ i.e. $\lambda_3 = \begin{bmatrix} \alpha_3^1 \\ 0 \end{bmatrix}$

     (c) $x_3 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \geq 0$
   - Compute $p_3 = H_{21}\, x_4 + H_{22}\, p_4 - \lambda_3 = \begin{bmatrix} 1 - \alpha_4^2 - \alpha_3^1 \\ 1 \end{bmatrix}$.
3. `For` $i = 2$ :
   - Compute $x_2 = H_{11}\, x_3 + H_{12}\, p_3 = \begin{bmatrix} 2 \\ 1 - \alpha_4^2 - \alpha_3^1 \end{bmatrix}$.
   - Compute $\lambda_2 = \begin{bmatrix} \alpha_2^1 \\ \alpha_2^2 \end{bmatrix}$ such that

(a) $\lambda_2^T x_2 = 0 \;\Leftrightarrow\; \alpha_2^1 = \frac{1}{2}(-1 + \alpha_4^2 + \alpha_3^1)\,\alpha_2^2$

(b) $\lambda_2 \geq 0$ with $x_2 \geq 0$ gives $\alpha_2^2 = 0$. Then $\lambda_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

(c) $x_2 \geq 0 \;\Leftrightarrow\; 1 - \alpha_4^2 - \alpha_3^1 \geq 0$

- Compute $p_2 = H_{21}\,x_3 + H_{22}\,p_3 - \lambda_2 = \begin{bmatrix} 1 \\ 1 - \alpha_4^2 - \alpha_3^1 \end{bmatrix}$.

4. `For` $i = 1$ :

- Compute $x_1 = H_{11}\,x_2 + H_{12}\,p_2 = \begin{bmatrix} 1 - \alpha_4^2 - \alpha_3^1 \\ 3 \end{bmatrix}$.
- Compute $\lambda_1 = \begin{bmatrix} \alpha_1^1 \\ \alpha_1^2 \end{bmatrix}$ such that

(a) $\lambda_1^T x_1 = 0 \;\Leftrightarrow\; \alpha_1^1 = \frac{3\,\alpha_1^2}{-1+\alpha_4^2+\alpha_3^1}$

(b) $\lambda_1 \geq 0$ with $x_1 \geq 0$ gives $\alpha_1^2 = 0$. Then $\lambda_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

(c) $x_1 \geq 0 \;\Leftrightarrow\; 1 - \alpha_4^2 - \alpha_3^1 \geq 0$

- Compute $p_1 = H_{21}\,x_2 + H_{22}\,p_2 - \lambda_1 = \begin{bmatrix} 1 - \alpha_4^2 - \alpha_3^1 \\ 1 \end{bmatrix}$.

5. `End` :

- Compute $x_0 = H_{11}\,x_1 + H_{12}\,p_1 = \begin{bmatrix} 3 \\ 2 - 2\,\alpha_4^2 - 2\,\alpha_3^1 \end{bmatrix}$.
- Compute $\lambda_0 = \begin{bmatrix} \alpha_0^1 \\ \alpha_0^2 \end{bmatrix}$ such that

(a) $\lambda_0^T x_0 = 0 \;\Leftrightarrow\; \alpha_0^1 = \frac{2}{3}(-1 + \alpha_4^2 + \alpha_3^1)\,\alpha_0^2$

(b) $\lambda_0 \geq 0$ with $x_0 \geq 0$ gives $\alpha_0^2 = 0$. Then $\lambda_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

(c) $x_0 \geq 0 \;\Leftrightarrow\; 1 - \alpha_4^2 - \alpha_3^1 \geq 0$

- Compute $p_0 = H_{21}\,x_1 + H_{22}\,p_1 - \lambda_0 = \begin{bmatrix} 1 \\ 1 - \alpha_4^2 - \alpha_3^1 \end{bmatrix}$.

Finally, the resulting solutions $x_i,\ \lambda_i,\ p_i$ and $u_i = -R^{-1}B^T p_i$ are summarized in Table 10.3, where $\alpha,\ \beta \in \mathbb{R}$ such that $\alpha,\ \beta \geq 0$ and $1 - \alpha - \beta \geq 0$. These results have been numerically verified with several values of $\alpha$ and $\beta$ by using the optimization algorithm `quadprog` with $x_0$ given as in Table 10.3.

| | $i=4$ | $i=3$ | $i=2$ | $i=1$ | $i=0$ |
|---|---|---|---|---|---|
| $x_i$ | $\begin{bmatrix}1\\0\end{bmatrix}$ | $\begin{bmatrix}0\\2\end{bmatrix}$ | $\begin{bmatrix}2\\1-\alpha-\beta\end{bmatrix}$ | $\begin{bmatrix}1-\alpha-\beta\\3\end{bmatrix}$ | $\begin{bmatrix}3\\2-2\,\alpha-2\,\beta\end{bmatrix}$ |
| $\lambda_i$ | $\begin{bmatrix}0\\\alpha\end{bmatrix}$ | $\begin{bmatrix}\beta\\0\end{bmatrix}$ | $\begin{bmatrix}0\\0\end{bmatrix}$ | $\begin{bmatrix}0\\0\end{bmatrix}$ | $\begin{bmatrix}0\\0\end{bmatrix}$ |
| $p_i$ | $\begin{bmatrix}1\\1-\alpha\end{bmatrix}$ | $\begin{bmatrix}1-\alpha-\beta\\1\end{bmatrix}$ | $\begin{bmatrix}1\\1-\alpha-\beta\end{bmatrix}$ | $\begin{bmatrix}1-\alpha-\beta\\1\end{bmatrix}$ | $\begin{bmatrix}1\\1-\alpha-\beta\end{bmatrix}$ |
| $u_i$ | $-1$ | $-1+\alpha+\beta$ | $-1$ | $-1+\alpha+\beta$ | $-1$ |

Table 10.3: Solutions of system (10.26) with (10.30) by applying the Algorithm 2.

This example shows that the conditions of Theorem 10.2.1 are computable. However, these two algorithms also reveal that they can be applied easily with a small final time $N$. Obviously using a larger final time increases the difficulty of the calculations. A lot of studies have been done on several algorithms but currently, we have not managed to create a systematic algorithm for computing solutions of the $\boldsymbol{LQ_+^N}$ problem by using the recurrent Hamiltonian equation (under its vector form or matrix form). The idea is to use, if it is possible, conditions (10.13)-(10.15) of Proposition 10.2.2 to find a solution of the $\boldsymbol{LQ_+^N}$ problem for any nonnegative initial condition. If it is not possible, the class of initial conditions for which it works should be found as well as an (matrix) algorithm for this class of initial conditions.

# Conclusions and Further Research Perspectives

The purpose of the research work described in this thesis was the study of the linear quadratic optimal control problem for input/state-invariant linear systems and its application to the particular locally positively invariant nonlinear system described by the chemostat model. We have developed theoretical results and numerical methodologies in order to solve the problem of co-existence of species in a chemostat. In the following paragraphs, we summarize our results and suggest some possible directions for future research.

In the first part of this thesis, important results have been developed to describe the input/state-invariance of time-varying and time-invariant linear systems. This input and/or state-invariance has been characterized by the matrices which describe the dynamics. The well-known particular case of positive systems has also been briefly described. Here, the input-invariance has been studied in the particular case of state-feedback control. It could be interesting to develop similar results for the general case of systems $[A(\cdot)\, B(\cdot)]$ or $[A,\, B]$ with any input $u$.

The main part of this thesis was devoted to the study of the linear quadratic problem in finite and infinite horizon, for input/state-invariant systems and for positive systems. In both cases, in finite horizon, necessary and sufficient optimality conditions were obtained by using the maximum principle with state and input constraints. Moreover, optimality conditions were established which were based on the admissibility of the solution of the standard $\boldsymbol{LQ^{t_f}}$ problem. In addition, for the positive $\boldsymbol{LQ_+^{t_f}}$ problem, sufficient conditions were stated in terms of the matrix solution of the RDE and the particular problem of minimal energy control with penalization of the final state was studied. Moreover, analytical and numerical studies of trajectories were performed on examples.

The LQ problem with constraints has already been studied, often with nonnegative constraints only, either on the state or on the input. The specific feature of the approach that is followed here is to describe the LQ problem for general state and input constraints, by using the admissibility of the solution of the standard $\boldsymbol{LQ^{t_f}}$ problem, with the objective of applying the results and the methodologies to a biological application. Obviously, another research possibility is to solve the input/state-invariant or the positive LQ problem for itself, by computing the multipliers associated to the constraints and by computing the matrix solutions of the Hamiltonian differential equation. In the numerical examples, a standard optimization algorithm was used in order to solve the problems. It would be interesting to find an ad hoc control algorithm adapted to this problem.

A receding horizon approach was developed in order to obtain a solution of the infinite horizon input/state-invariant $\boldsymbol{LQ}^{\infty}_{\bar{u},\bar{x}}$ problem. For the positive $\boldsymbol{LQ}^{\infty}_{+}$ problem, criteria were established, by using a Newton-like iterative scheme, an Hamiltonian approach and the study of a diagonal solution of the ARE. For the unstable case, the iterative scheme was revealed hard to implement. Thus a perspective for this work is to analyze the unstable case either with another iterative scheme or with another methodology that would be more adapted to this special case. Finally, in the last chapter of this part, the inverse input/state-invariant $\boldsymbol{LQ}^{\text{inv}}_{\bar{u},\bar{x}}$ problem was described and solved by using LMIs or BMIs. Further study of the LMIs and BMIs would be attractive in order to implement a systematic method to solve the inverse input/state-invariant or the inverse positive LQ problem. In addition, it would be interesting to study the proof of the convergence of the heuristic iterative process described in Chapter 7.

The last part was devoted to the application of the input/state-invariant LQ problem in order to solve the problem of coexistence of species in competition in a chemostat. The methodology that was used is to guarantee the local positive input/state-invariance of the nonlinear system (which describes the chemostat model) by ensuring the input/state-invariance of its linear approximation around an equilibrium, with the application of an appropriate LQ-optimal control. An interesting perspective in this framework can be, of course, the study of the input/state-invariance of nonlinear systems. An advanced study of nonlinear systems can also be relevant in order to obtain a global coexistence of species. The main objective of this study was to apply the theory of the input/state-invariant LQ problem on an attractive biological application. Other methodologies could be more adapted to solve the problem of coexistence by considering the nonlinear model itself, instead of performing the analysis on the linearized system.

Finally, the last chapter was devoted to the discrete time case, for which the well-known results on positive systems have been recalled before considering the positive $\boldsymbol{LQ}^{\boldsymbol{N}}_{+}$ problem in finite horizon. The analysis was similar to the one of the continuous time case and therefore most of the results were merely adaptations of the continuous time problem. Furthermore, the study of monomial systems (which include the well-known class of compartmental systems, which are significant in applications) was really specific to the discrete time case. By describing a new concept for monomial matrices (namely structural similarity), we have proved that the positive $\boldsymbol{LQ}^{\boldsymbol{N}}_{+}$ problem had a solution for this particular class of positively invertible systems. Future possible work in this framework is to adapt the results of input/state-invariant continuous time systems in order to obtain similar results for the input/state-invariant LQ problem in discrete time. A receding horizon approach can also be developed for the infinite horizon LQ problem. It was briefly described in [Bea06] for the positive LQ problem, notably with some numerical examples. Another perspective for the discrete time case is to develop an adapted iterative scheme which converges to the solution of the recurrent Riccati equation, as in the continuous time positive $\boldsymbol{LQ}^{\infty}_{+}$ problem. Finally, an inverse LQ problem can also be described for discrete time systems.

In conclusion, the research work reported in this thesis yields some methods for solving the LQ-optimal control problem for input/state-invariant linear systems. Several perspectives have been proposed in order to improve them. Finally, we hope that this work will be useful, in some way, for further research.

# Summary of Contributions

Our main contribution is the study of the LQ-optimal control problem with state and input constraints and the application to the problem of coexistence of species in a chemostat model. Our contributions are summarized as follows :

- characterizations of input and/or state-invariant time-varying systems in terms of the matrices defining the dynamics (Chapter 1) ;
- a proof of optimality conditions for the input/state-invariant $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem based on the maximum principle (Theorem 3.2.1) ;
- positivity criteria for the positive $\boldsymbol{LQ}^{\boldsymbol{t_f}}_{+}$ problem in finite horizon (in terms of an upper bound of the solution of the RDE (Theorem 4.2.1), via the study of the minimal energy control problem (Theorems 4.2.2 and 10.2.4) ;
- a definition and a characterization of an equivalence relation (structural similarity) for the set of monomial matrices (Definition A.3.2 and Lemma A.3.1) ;
- a positivity criterion for the positive $\boldsymbol{LQ}^{\boldsymbol{N}}_{+}$ problem in discrete time for monomial systems (Theorem 10.2.6) ;
- the study of the infinite horizon positive $\boldsymbol{LQ}^{\boldsymbol{\infty}}_{+}$ problem by means of a Newton type iterative scheme (Section 6.1) ;
- the study of the inverse input/state-invariant $\boldsymbol{LQ}^{\mathrm{inv}}_{\bar{\boldsymbol{u}},\bar{\boldsymbol{x}}}$ problem by means of LMIs and BMIs (Chapter 7) ;
- a criterion of local positive invariance of a nonlinear system by means of the input/state-invariance of its linear approximation around an equilibrium (Theorem 8.2.1) ;
- the resolution of the problem of coexistence of species in a chemostat model by applying an appropriate LQ-optimal control (Chapter 9, Sections 9.3 and 9.4) ;
- the development of specific numerical examples.

Up to our knowledge, the proofs which are detailed in this thesis are part of its contributions. If a cited result is already available in the literature, we only mention a reference without giving a proof. Finally, part of this thesis (especially Chapters 4, 6 and 10) is based on the following publications :

## Journal paper

[BW10] – 2010 : Ch. Beauthier and J. J. Winkin, *LQ-optimal control of positive linear systems*, Optimal control : Applications and Methods, Vol. 31, No. 6, pp. 547-566, 2010.

## Conference proceedings (with review process)

- [LWB06] – 2006 : M. Laabissi and J. J. Winkin and Ch. Beauthier, *On the positive LQ-problem for linear continuous-time systems*, Proceedings of the second Multidisciplinary International Symposium on Positive Systems : Theory and Applications (POSTA 2006), Grenoble, France, in Lecture Notes in Control and Information Sciences, Vol. 341, pp. 295-302, 2006.
- [BW08] – 2008 : Ch. Beauthier and J. J. Winkin, *Finite horizon LQ-optimal control for continuous time positive systems*, Proceedings of the Eighteenth International symposium on Mathematical Theory of Networks and Systems (MTNS2008), Virginia Tech, Blacksburg, Virginia, USA, CD-ROM Paper Nr 054, 2008.
- [BW09] – 2009 : Ch. Beauthier and J. J. Winkin, *On the positive LQ-problem for linear discrete time systems*, Proceedings of the Third Multidisciplinary International Symposium on Positive Systems: Theory and Applications (POSTA 2009), in Positive Systems, pp. 45-53, 2009.

## Master's thesis (DEA)

[Bea06] – 2006 : Ch. Beauthier, *Le problème linéaire quadratique positif*, Mémoire de DEA, Facultés Universitaires Notre-Dame de la Paix (FUNDP), Namur, 2006.

# Main Notations and Abbreviations

| | | |
|---|---|---|
| General Abbreviations | | |
| | e.g. | for example |
| | i.e. | that is |
| | viz. | namely |
| | w.r.t | with respect to |
| | w.l.g. | without loss of generality |
| | psd | positive semidefinite |
| | pd | positive definite |
| | LTI | linear time invariant |
| | LTV | linear time varying |
| | LQ | linear quadratic |
| | LMI | linear matrix inequality |
| | BMI | bilinear matrix inequality |
| | LMIs | linear matrix inequalities |
| | BMIs | bilinear matrix inequalities |
| Riccati equation | | |
| | RDE | Riccati differential equation |
| | ARE | algebraic Riccati equation |
| | RRE | recurrent Riccati equation |

| Linear Quadratic problem | | |
|---|---|---|
| Finite horizon | | |
| | $\boldsymbol{LQ^{t_f}}$ | standard LQ problem |
| | $\boldsymbol{LQ^{t_f}_{+}}$ | positive LQ problem |
| | $\boldsymbol{LQ^{t_f}_{\bar{x}}}$ | state-invariant LQ problem |
| | $\boldsymbol{LQ^{t_f}_{\bar{u}}}$ | input-invariant LQ problem |
| | $\boldsymbol{LQ^{t_f}_{\bar{u},\bar{x}}}$ | input/state-invariant LQ problem |
| | $\boldsymbol{LQ^{N}}$ | standard LQ problem (in discrete time) |
| | $\boldsymbol{LQ^{N}_{+}}$ | positive LQ problem (in discrete time) |
| Infinite horizon | | |
| | $\boldsymbol{LQ^{\infty}}$ | standard LQ problem |
| | $\boldsymbol{LQ^{\infty}_{+}}$ | positive LQ problem |
| | $\boldsymbol{LQ^{\infty}_{\bar{u},\bar{x}}}$ | input/state-invariant LQ problem |
| | $\boldsymbol{LQ^{\infty}_{\mathrm{inv}}}$ | inverse standard LQ problem |
| | $\boldsymbol{LQ^{\mathrm{inv}}_{+}}$ | inverse positive LQ problem |
| | $\boldsymbol{LQ^{\mathrm{inv}}_{\bar{x}}}$ | inverse state-invariant LQ problem |
| | $\boldsymbol{LQ^{\mathrm{inv}}_{\bar{u},\bar{x}}}$ | inverse input/state-invariant LQ problem |

| Algebra | |
|---|---|
| $\mathbb{R}$ | set of real numbers |
| $\mathbb{C}$ | set of complex numbers |
| $\mathbb{R}_+$ | nonnegative orthant |
| $\mathcal{R}e(z)$ | real part of a complex $z$ |
| $\lvert \cdot \rvert$ | absolute value of a scalar |
| $\lVert \cdot \rVert$ | vector or matrix norm (Euclidean unless otherwise specified) |
| $e_i$ | $i^{\text{th}}$ vector of the canonical basis |
| $I_m$ | identity matrix of dimension $m$ |
| $0_m$ | zero matrix of dimension $m$ |
| $\sigma(A)$ | set of eigenvalues of $A$ |
| $\rho(A)$ | spectral radius of $A$ |
| $\mathcal{N}(A)$ | the null space of $A$ |
| $\lambda_{\min}(A)$ | minimum eigenvalue of $A$ |
| $\mathcal{L}^-(A)$ | stable subspace, i.e. $A$-invariant subspace spanned by the (generalized) eigenvectors corresponding to eigenvalues with negative real parts |
| $\mathcal{L}^0(A)$ | critical subspace, i.e. $A$-invariant subspace spanned by the (generalized) eigenvectors corresponding to eigenvalues with zero real parts |
| $\mathcal{L}^{0+}(A)$ | unstable subspace, i.e. $A$-invariant subspace spanned by the (generalized) eigenvectors corresponding to eigenvalues with nonnegative real parts |
| $\lambda_F$ | Frobenius eigenvalue of $A$ |
| $A \geq 0$ | $A$ is a nonnegative matrix, i.e. every entries of $A$ are nonnegative |
| $A > 0$ | $A$ is a positive matrix, i.e. every entries of $A$ are nonnegative and at least one entry is (strictly) positive |
| $A \gg 0$ | $A$ is a strictly positive matrix, i.e. every entries of $A$ are (strictly) positive |
| $A \overset{\mathbf{s}}{=} B$ | $A$ and $B$ are structurally similar |
| $A_l^+$ | left pseudo-inverse of $A$ |
| $A_r^+$ | right pseudo-inverse of $A$ |

| System theory | |
|---|---|
| $x(t)$ | state trajectory |
| $u(t)$ | input trajectory |
| $\mathcal{U}$ | set of piecewise continuous input functions |
| $R = [A(\cdot),\, B(\cdot)]$ | system description $\dot{x}(t) = A(t)\, x(t) + B(t)\, u(t)$ |
| | denoted by $R = [A,\, B]$ in the time-invariant case |
| $R = [A(\cdot),\, 0]$ | system description $\dot{x}(t) = A(t)\, x(t)$ |
| | denoted by $R = [A,\, 0]$ in the time-invariant case |
| $\Phi(t,\, t_0)$ | fundamental matrix of system $[A(\cdot),\, 0]$ which satisfies $\begin{cases} \frac{\partial \Phi}{\partial t}(t,\, t_0) = A(t)\, \Phi(t,\, t_0), \quad \forall\, t \in [t_0,\, t_f] \\ \Phi(t_0,\, t_0) = I_n \end{cases}$ |
| $\Phi_K(t,\, t_0)$ | fundamental matrix of system $[A + B\, K(\cdot),\, 0]$ which satisfies $\begin{cases} \frac{\partial}{\partial t}\Phi_K(t,\, t_0) = (A + B\, K(t))\, \Phi_K(t,\, t_0), \quad \forall\, t \in [t_0,\, t_f] \\ \Phi_K(t_0,\, t_0) = I_n \end{cases}$ |
| $NO(C,\, A)$ | unobservable subspace |
| $ND(C,\, A)$ | undetectable subspace |

# Part IV

# Appendix

# Appendix A

# Particular Matrices

This appendix is devoted to particular matrices as nonnegative and Metzler matrices which play an important role in system theory. There exist a large class of references devoted to these particular matrices, e.g. those used here are the following : [BR97, Chapter 1], [HJ85, Chapter 8], [Ava00, Chapters 2-3], [LT85, Chapter 15], [BP94, Chapter 2], [Min88, Chapters 1 and 4], [BNS89, Chapter 2] and [HCH10, Chapter 2]. Definitions and main properties (as the spectral properties) are given for such matrices.

Then the general theory of Z-matrices and M-matrices is briefly described, based on the following references : [BR97, Chapter 1], [BP94, Chapter 6], [Min88, Chapter 6], [Guo01] and [BNS89, Chapter 2]. In the $\boldsymbol{LQ_+^N}$ problem in discrete time, see Chapter 10.2, we use the class of monomial matrices or positively invertible matrices. Their properties are studied in this chapter, based on [PC72], [BP74] and [BP94, Chapter 5]. Moreover, an equivalence relation on the set of such matrices is defined.

Finally, at the end of the chapter, the definition of the Kronecker product and its use for the Lyapunov equation are recalled, see [HJ91, Chapter 4] and [LT85, Chapter 12]. It is used in Chapter 6 for the study of the positive $\boldsymbol{LQ_+^\infty}$ problem in infinite horizon.

## A.1 Nonnegative and Metzler matrices

### A.1.1 Definitions

**Definition A.1.1**

- *A matrix $A \in \mathbb{R}^{n\times n}$ is said to be* **nonnegative***, denoted $A \geq 0$, if for all $i,\ j = 1,\dots,n$, $a_{ij} \geq 0$, i.e. every entries of $A$ are nonnegative.*

- *A matrix $A \in \mathbb{R}^{n\times n}$ is said to be* **positive***, denoted $A > 0$, if $A \geq 0$ and there exist $i,\ j = 1,\dots,n$, $a_{ij} > 0$, i.e. every entries of $A$ are nonnegative and at least one entry is (strictly) positive.*

- *A matrix $A \in \mathbb{R}^{n\times n}$ is said to be* **strictly positive***, denoted $A \gg 0$, if for all $i,\ j = 1,\dots,n$, $a_{ij} > 0$, i.e. every entries of $A$ are (strictly) positive.*

*In particular, these notations and definitions obviously apply to vectors $x \in \mathbb{R}^n$. However for the scalar case, "strictly positive" coincides with "positive".*

- *A matrix $A \in \mathbb{R}^{n\times n}$ is said to be a* **Metzler matrix***, if for all $i, j = 1, \ldots, n$, $i \neq j$, $a_{ij} \geq 0$, i.e. every off-diagonal entries of $A$ are nonnegative.*

The following result obviously holds :

**Proposition A.1.1** *A matrix $A$ is a Metzler matrix if and only if there exists $\alpha \in \mathbb{R}$, $\alpha > 0$ such that $A + \alpha\, I_n \geq 0$.*

Therefore we can move easily from nonnegative matrices to Metzler matrices and conversely. By applying the definition of a nonnegative matrix, one has the following result :

**Proposition A.1.2** *$A$ is a nonnegative matrix if and only if the positive orthant $\mathbb{R}^n_+$ is $A$-invariant, i.e.*

$$\forall\, x \in \mathbb{R}^n_+,\ A\, x \in \mathbb{R}^n_+.$$

One has a similar result for Metzler matrix :

**Proposition A.1.3** *$A$ is a Metzler matrix if and only if*

$$\forall\, t \geq 0\ :\ e^{A\,t} \geq 0, \tag{A.1}$$

*or equivalently, $\forall\, t \geq 0$, the positive orthant $\mathbb{R}^n_+$ is $e^{A\,t}$-invariant, i.e.*

$$\forall\, t \geq 0,\, \forall\, x \in \mathbb{R}^n_+,\ e^{A\,t}x \in \mathbb{R}^n_+.$$

**Proof :**
Necessity : Since $A$ is a Metzler matrix, by Proposition A.1.1, there exists $\alpha > 0$ such that $A + \alpha\, I_n \geq 0$. Then

$$\forall\, t \geq 0,\ e^{A+\alpha\, I_n} = \sum_{k=0}^{\infty} \frac{(A + \alpha\, I_n)^k}{k!}\, t^k \ \geq 0.$$

Hence, with $e^{(A+\alpha\, I)\,t} = e^{A\,t}\, e^{\alpha\,t} \geq 0$, one has $\forall\, t \geq 0,\ e^{A\,t} \geq 0$ since $e^{\alpha\,t}$ is a positive scalar.

Sufficiency : Since $A = \dfrac{\mathrm{d}}{\mathrm{dt}}(e^{A\,t})\Big|_{t=0} = \lim\limits_{t\to 0+} \dfrac{e^{A\,t} - I_n}{t}$, we obtain, with $e_j$ denoted the $j^{\text{th}}$ vector of the canonical basis, for $i \neq j$ :

$$\begin{aligned} a_{ij} &= \lim_{t\to 0+} \frac{< e^{A\,t}\, e_j - e_j,\, e_i >}{t} \\ &= \lim_{t\to 0+} \left\{ \frac{< e^{A\,t}\, e_j,\, e_i >}{t} - \frac{< e_j,\, e_i >}{t} \right\} \\ &= \lim_{t\to 0+} \frac{< e^{A\,t}\, e_j,\, e_i >}{t} \geq 0. \end{aligned}$$

Hence $a_{ij} \geq 0$ for $i \neq j$.

$\square$

**Remarks A.1.1**

- *The $A$-invariance of $\mathbb{R}_+^n$ implies the $e^{At}$-invariance of $\mathbb{R}_+^n$.*

  *Indeed, if $\forall\, x \in \mathbb{R}_+^n,\ A\,x \in \mathbb{R}_+^n$, then, with $x \geq 0$ and $t \geq 0$, one has :*

$$e^{A\,t}\,x = \sum_{k=0}^{\infty} \frac{(A\,t)^k}{k!}\,x = \sum_{k=0}^{\infty} \frac{A^k\,x}{k!}\,t^k.$$

  *Then since $A\,x \geq 0$, by recurrency for all $k$, $A^k\,x \geq 0$. Hence $e^{A\,t}\,x \geq 0$.*

- *Conversely, the $e^{At}$-invariance of $\mathbb{R}_+^n$ does not imply necessarily its $A$-invariance.*

  *Indeed, consider the case where $n = 1$. Then $\mathbb{R}_+$ is $e^{At}$-invariant since for all $t \geq 0$, $e^{A\,t} > 0$. However $\mathbb{R}_+$ is not $A$-invariant if $A < 0$.*

## A.1.2 Spectral properties

First we recall the definitions of spectrum and spectral radius and the concept of dominant eigenvalue and eigenvector associated to nonnegative and Metzler matrix.

**Definition A.1.2**

- *The* **spectrum** *of a matrix $A$, denoted $\sigma(A)$, is the set of its eigenvalues, i.e.*

$$\sigma(A) := \{\lambda \in \mathbb{C} \ :\ A\,x = \lambda\,x,\, x \neq 0\}.$$

- *The* **spectral radius** *of a matrix $A$, denoted $\rho(A)$, is defined as :*

$$\rho(A) := \max\{|\lambda| \ :\ \lambda \in \sigma(A)\}.$$

- *A* **dominant eigenvalue**, *$\lambda_d$,* **of a nonnegative matrix** *$A$, is defined as follows :*

$$\forall\, \lambda \in \sigma(A)\ :\ |\lambda_d| \geq |\lambda|,\ \textit{ i.e. }\ |\lambda_d| = \rho(A).$$

- *A* **dominant eigenvalue**, *$\lambda_d$,* **of a Metzler matrix** *$A$, is defined as follows :*

$$\forall\, \lambda \in \sigma(A)\ :\ \mathcal{R}e(\lambda_d) \geq \mathcal{R}e(\lambda).$$

- *A* **dominant eigenvector**, *$v_d$, of a matrix $A$ is an eigenvector associated to a dominant eigenvalue, i.e. $A\,v_d = \lambda\,v_d$.*

We also use the notations $\mathcal{L}^-(A),\ \mathcal{L}^0(A),\ \mathcal{L}^+(A)$ which denote the $A$-invariant subspaces spanned by a basis of (generalized) eigenvectors corresponding to eigenvalues with negative, zero and positive real parts. For a matrix $A$, $\mathcal{N}(A)$ denotes the null space.

The following results on the spectral properties of nonnegative and Metzler matrices are well-known, see e.g. [Ava00, Chapter 3], [HJ85, Chapter 8] and [HCH10, Chapter 2].

**Theorem A.1.4 (Perron-Frobenius for nonnegative matrices)** *Let $A$ be a nonnegative matrix. Then $\rho(A)$ is an eigenvalue of $A$, called the Frobenius eigenvalue, and there exists a positive eigenvector $x$, associated to $\rho(A)$, which is called the Frobenius eigenvector, such that $A\,x = \rho(A)\,x$ and $\forall\ \lambda\ \in\ \sigma(A),\ |\lambda| \leq \rho(A)$.*

**Theorem A.1.5 (Perron-Frobenius for Metzler matrices)** *Let $A$ be a Metzler matrix. Then there exists a real eigenvalue $\lambda_F$ of $A$, which is called the Frobenius eigenvalue, such that there exists a positive eigenvector $x$, associated to $\lambda_F$, which is called the Frobenius eigenvector, such that $A\,x = \lambda_F\,x$ and $\forall\ \lambda \in \sigma(A),\ \mathcal{R}e(\lambda) \leq \lambda_F$.*

## A.2 Z-matrices and M-matrices

**Definition A.2.1**

- *A real $n \times n$-matrix $D$ is said to be a $Z$-**matrix** if $-D$ is a Metzler matrix.*
- *A real $n \times n$-matrix $D$ is said to be a $M$-**matrix** if $D = s\,I_n - \tilde{D}$ for some matrix $\tilde{D} \geq 0$ and for some real number $s \geq \rho(\tilde{D})$.*

It can be easily seen that any $Z$-matrix $D$ is of the form $D = s\,I_n - \tilde{D}$ for some real number $s$ and some matrix $\tilde{D} \geq 0$, and that any $M$-matrix is a $Z$-matrix. Moreover, by the Perron-Frobenius theorem, one has the following result :

**Proposition A.2.1** *A $M$-matrix $D = s\,I_n - \tilde{D}$ is nonsingular if and only if $s > \rho(\tilde{D})$.*

The following result can be found e.g. in [BP94] and [HJ91, Theorem 2.5.3, pp. 114-115] ; see also [GL00b]. Observe that, in those references, a $M$-matrix is assumed to be nonsingular by definition. In the present context, it is useful to consider $M$-matrices which might possibly be singular. See also [HCH10, Chapter 2].

**Theorem A.2.2** *For any $Z$-matrix $D$, the following assertions are equivalent :*
*(i) $D$ is nonsingular and $D^{-1} \geq 0$.*
*(ii) $Dx \gg 0$ for some vector $x \gg 0$.*
*(iii) All eigenvalues of $D$ have positive real parts.*
*Moreover any of these assertions characterizes the fact that $D$ is a nonsingular $M$-matrix, i.e. such that $D = s\,I_n - \tilde{D}$, where $\tilde{D} \geq 0$ and $s > \rho(\tilde{D})$.*

Another useful result is the following proposition :

**Proposition A.2.3** *If $A$ is a nonsingular $M$-matrix, then the solution $x$ of $A\,x = q$ with $q \gg 0$ is such that $x \gg 0$.*

**Proof :** Since $A$ is a nonsingular $M$-matrix, $A^{-1} \geq 0$ and $x = A^{-1}q \geq 0$, by Theorem A.2.2. Then $x \gg 0$. By contradiction, if there exists $x_i = 0$, then, with $A^{-1} := B$, one has

$$x_i = \sum_{j=1}^{n} b_{ij} q_j = 0$$

with $q_j \gg 0$ and $b_{ij} \geq 0$. That is $b_{ij} = 0$, for all $j$ which is in contradiction with the fact that $A$ is nonsingular. Then $x$ is strictly positive.

□

Finally, the following proposition gives a criterion on the stability of a $M$-matrix, see [HCH10, Theorem 2.10] :

**Proposition A.2.4** *If $A$ is a nonsingular $M$-matrix, then $-A$ is asymptotically stable, i.e.* $\forall\, \lambda \in \sigma(-A), \quad \mathcal{R}e(\lambda) < 0$.

## A.3 Monomial matrices

**Definition A.3.1** *A nonnegative matrix $M$ is said to be* **monomial** *if $M$ is a diagonal matrix up to a permutation, i.e. $M = D\,P = \mathrm{diag}[m_i]_{i=1}^{n}\,P$, where $D$ is a positive definite diagonal matrix and $P$ is a permutation matrix, or equivalently $M^{-1} \geq 0$, see e.g. [BP74] and [PC72].*

**Definition A.3.2** *Let $L$ and $M$ be monomial matrices. $L$ and $M$ are said to be* **structurally similar**, *denoted by $L \overset{\mathbf{s}}{=} M$, if and only if there exist positive definite diagonal matrices $D_1$ and $D_2$ such that $L = D_1\,M\,D_2$.*

It is easy to check that "$\overset{\mathbf{s}}{=}$" is an equivalence relation on the set of monomial matrices, see [Bea06, Theorem 5.2.6].

**Remark A.3.1** *In Definition A.3.2, one of the diagonal matrices $D_i$ can be chosen as the identity matrix, such that one has $L \overset{\mathbf{s}}{=} M$ with $L = D_1\,M$ or with $L = M\,D_2$.*

The following straightforward result is needed in the study of the $\boldsymbol{LQ_{+}^{N}}$ problem for positive discrete time systems, see Subsection 10.2.3.

**Lemma A.3.1** *Let $L$ and $M$ be monomial matrices. Let $P$ be a permutation matrix such that $L = D_1\,P$ and $M = P\,D_2$ where $D_1$ and $D_2$ are positive definite diagonal matrices. Then $L$ and $M$ are structurally similar.*

**Proof :** By assumption $L \overset{\mathbf{s}}{=} P$ and $M \overset{\mathbf{s}}{=} P$. Then, by the transitivity and symmetry properties of the equivalence relation $\overset{\mathbf{s}}{=}$, it follows that $L \overset{\mathbf{s}}{=} M$.

## A.4 Kronecker product

We also need the Kronecker product defined as follows :

**Definition A.4.1** *The* **Kronecker product** *of matrices* $C = (c_{ij}) \in \mathbb{R}^{k\times l}$ *and* $D \in \mathbb{R}^{m\times n}$ *of any size, is the matrix* $C \otimes D \in \mathbb{R}^{km\times ln}$ *given by :*

$$C \otimes D := \begin{bmatrix} c_{11}D & \dots & c_{1l}D \\ . & \dots & . \\ c_{k1}D & \dots & c_{kl}D \end{bmatrix}. \tag{A.2}$$

The stack operator maps an $m \times n$-matrix into an $mn$-vector. More precisely, the stack of an $m \times n$-matrix $C$, denoted by $\mathrm{vect}(C)$, is the $mn$-vector formed by stacking the columns of $C$.

With the definition of the Kronecker product and the vector $\mathrm{vect}(C)$ defined for a matrix $C$, well-known matrix equations can be rewritten, as the Sylvester equation and the Lyapunov equation, see e.g. [HJ91, Chapter 4] :

**Proposition A.4.1** *For given matrices* $C, D, Q$ *and* $X$ *with compatible sizes, the Sylvester equation*

$$C\,X + X\,D = Q$$

*is equivalent to the linear algebraic equation*

$$(I \otimes C + D^T \otimes I)\ \mathrm{vect}(X) = \mathrm{vect}(Q),$$

*and the Lyapunov equation*

$$C^T X + X\,C = -Q$$

*is equivalent to*

$$(I \otimes C^T + C^T \otimes I)\ \mathrm{vect}(X) = -\,\mathrm{vect}(Q).$$

In addition, the following result can be easily shown, see [HJ91, pp. 268-269] or [LT85, pp. 411-412],

**Proposition A.4.2** *If* $C$ *and* $D$ *are* $M$*-matrices, then so is for* $(I \otimes C + D^T \otimes I)$.

# Appendix B

# Maximum Principle with State and Input Constraints

In this chapter, the maximum principle with state and input constraints is presented, as it is developed in [HSV95], with the same notations. These results are used in Part II and Subsection 9.4.2.

## B.1 Problem statement

Consider the following optimal control problem : maximize the following cost

$$J = \int_0^T F(x(t),\, u(t),\, t)\, \mathrm{dt} + S(x(T),\, T) \tag{B.1}$$

for the following system dynamics

$$\dot{x}(t) = f(x(t),\, u(t),\, t),\; x(0) = x_0 \tag{B.2}$$

under the constraints

$$\begin{aligned} g(x(t),\, u(t), t) &\geq 0 && \text{(B.3a)} \\ h(x(t),\, t) &\geq 0 && \text{(B.3b)} \\ a(x(T),\, T) &\geq 0 && \text{(B.3c)} \\ b(x(T),\, T) &= 0 && \text{(B.3d)} \end{aligned}$$

Assume that the functions $F$ from $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ into $\mathbb{R}$, $S$ from $\mathbb{R}^n \times \mathbb{R}$ into $\mathbb{R}$, $f$ from $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ into $\mathbb{R}^n$, $g$ from $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ into $\mathbb{R}^s$, $h$ from $\mathbb{R}^n \times \mathbb{R}$ into $\mathbb{R}^q$, and $a,\, b$ from $\mathbb{R}^n \times \mathbb{R}$ into $\mathbb{R}^l$ and $\mathbb{R}^{l'}$ respectively, are continuous differentiable with respect to all their arguments.

In the sequel, the following constraint qualification is assumed to hold for all possible values of $x(T)$ and $T$ :

$$\operatorname{rank}\begin{bmatrix} \frac{\partial a}{\partial x} & \operatorname{diag}(a) \\ \frac{\partial b}{\partial x} & 0 \end{bmatrix} = l + l' \tag{B.4}$$

where $\operatorname{diag}(a) := \operatorname{diag}([a_1(x(T), T), \ldots, a_l(x(T), T)])$ denotes the diagonal matrix containing the components of $a(x(T), T))$ on its diagonal. This full rank condition means that the gradients w.r.t. $x$ of the equality constraints and of the active inequality constraints must be linearly independent. In order to distinguish between the mixed constraints (B.3a) and the pure state constraints (B.3b), we assume that each component of the function $g$ depends explicitly on the control $u$. More precisely, we impose the following full rank condition :

$$\operatorname{rank}\left[\frac{\partial g}{\partial u} \quad \operatorname{diag}(g)\right] = s \tag{B.5}$$

for all arguments $x(t),\ u(t),\ t$ that could arise along an optimal solution. The *constraint qualification* (B.5) means that the gradients w.r.t. $u$ of all the active constraints $g \geq 0$ must be linearly independent.

## B.2 Direct adjoining approach

In this method, the Hamiltonian $H$ and Lagrangian $L$ are defined as follows :

$$H(x,\, u,\, \lambda_0,\, \lambda,\, t) = \lambda_0\, F(x,\, u,\, t) + \lambda\, f(x,\, u,\, t) \tag{B.6}$$

$$L(x,\, u,\, \lambda_0,\, \lambda,\, \mu,\, \nu,\, t) = H(x,\, u,\, \lambda_0,\, \lambda,\, t) + \mu\, g(x,\, u,\, t) + \nu\, h(x,\, t), \tag{B.7}$$

where $\lambda_0 \geq 0$ is a constant, $\lambda \in \mathbb{R}^n$ is the adjoint variable, and $\mu \in \mathbb{R}^s$ and $\nu \in \mathbb{R}^q$ are multipliers.

**Theorem B.2.1** *Let $(x^*(\cdot),\, u^*(\cdot))$ be an optimal pair for the problem over a fixed interval $[t_0,\, T]$ (i.e. $(x^*(\cdot),\, u^*(\cdot))$ globally maximizes (B.1) where $x^*(\cdot)$ is the state trajectory corresponding to $u^*(\cdot)$ and conditions (B.3) are satisfied), such that $u^*(\cdot)$ is right-continuous with left-hand limits and the constraint qualification (B.5) holds for every triple $\{t,\, x^*(t),\, u\}$, $t \in [t_0,\, T]$ with $u \in \Omega(x^*(t),\, t) := \{u \in \mathbb{R}^m \mid g(x,\, u,\, t) \geq 0\} \subset \mathbb{R}^m$.*
*Assume that $x^*(\cdot)$ has only finitely many junction (i.e. switching) times.*
*Then there exist a constant $\lambda_0 \geq 0$, a piecewise absolutely continuous[1] costate trajectory $\lambda(\cdot)$ mapping $[t_0,\, T]$ into $\mathbb{R}^n$, a piecewise continuous multiplier functions $\mu(\cdot)$ and $\nu(\cdot)$ mapping $[t_0,\, T]$ into $\mathbb{R}^s$ and $\mathbb{R}^q$, respectively, a vector $\eta(\tau_i) \in \mathbb{R}^q$ for each point $\tau_i$ of discontinuity of $\lambda(\cdot)$, and $\alpha \in \mathbb{R}^l$, $\beta \in \mathbb{R}^{l'}$, $\gamma \in \mathbb{R}^q$ such that $(\lambda_0,\, \lambda(t),\, \mu(t),\, \nu(t),\, \alpha,\, \beta,\, \gamma,\, \eta(\tau_1),\, \eta(\tau_2), \ldots) \neq 0$*

[1]A piecewise absolutely continuous function is a piecewise continuous function whose continuous segments are absolutely continuous

*for every $t$ and the following conditions hold almost everywhere :*

$$\begin{aligned} u^*(t) &= \arg\max_{u\in\Omega} H(x^*(t),\, u,\, \lambda_0,\, \lambda(t),\, t) && \text{(B.8a)}\\ L_u^*(t) &= H_u^*(t) + \mu\, g_u^*(t) = 0 && \text{(B.8b)}\\ \dot{\lambda}(t) &= -L_x^*(t) && \text{(B.8c)}\\ \mu(t) &\geq 0, \quad \mu(t)\, g^*(t) = 0 && \text{(B.8d)}\\ \nu(t) &\geq 0, \quad \nu(t)\, h^*(t) = 0 && \text{(B.8e)}\\ dH^*(t)/\,\mathbf{dt} &= dL^*(t)/\,\mathbf{dt} = L_t^*(t) = \partial L^*(t)/\partial t. && \text{(B.8f)} \end{aligned}$$

*At the terminal time $T$, the following transversality conditions hold :*

$$\begin{aligned} \lambda(T) &= \lambda_0\, S_x^*(T) + \alpha\, a_x^*(T) + \beta\, b_x^*(T) + \gamma\, h_x^*(T) && \text{(B.9a)}\\ \alpha &\geq 0,\, \gamma \geq 0, && \text{(B.9b)}\\ \alpha\, a^*(T) &= \gamma\, h^*(T) = 0. && \text{(B.9c)} \end{aligned}$$

*For any time $\tau$ in a boundary interval and for any contact time $\tau$, the costate trajectory $\lambda$ may have a discontinuity given by the following jump conditions :*

$$\begin{aligned} \lambda(\tau^-) = \lambda(\tau^+) + \eta(\tau)\, h_x^*(\tau) && \text{(B.10a)}\\ H^*(\tau^-) = H^*(\tau^+) - \eta(\tau)\, h_t^*(\tau) && \text{(B.10b)}\\ \eta(\tau) \geq 0,\, \eta(\tau)\, h^*(\tau) = 0, && \text{(B.10c)} \end{aligned}$$

*where $\tau^+$ and $\tau^-$ denote the left-hand side and the right-hand side limits, respectively.*

**Remark B.2.1** *The condition*

$$(\lambda_0,\, \lambda(t),\, \mu(t),\, \nu(t),\, \alpha,\, \beta,\, \gamma,\, \eta(\tau_1),\, \eta(\tau_2), \ldots) \neq 0 \text{ for every } t$$

*can play an important role in distinguishing the normal case ($\lambda_0$ = 1) from the abnormal case ($\lambda_0 = 0$). In fact, this condition implies that $\lambda_0 = 1$ in the examples analyzed in [HSV95, Section 9].*

# Appendix C

# Discretization of a Linear Continuous Time System

In this chapter, the discretization of a linear continuous time system is described, and the associated discrete cost is also given.

## C.1 Discretization of a linear continuous time system

Consider the following LTI continuous time system

$$\dot{x}(t) = A\,x(t) + B\,u(t),\; x(0) = x_0, \quad t \in [t_0,\, t_f] \tag{C.1}$$

In order to transform this system into a discrete time system, consider $t_0 = i_0\,h$, $t_f = N\,h$, $x(t) = x(i\,h) =: x_i$, $u(t) = u(i\,h) =: u_i$, where $h$ is the sampling time and $t \in [i\,h,\,(i+1)\,h[$, for $i = i_0, \ldots, N-1$. We are looking for

$$x_{i+1} = f(x_i,\, u_i).$$

First integrate the homogeneous part of (C.1), i.e. $\dot{x}(t) = A\,x(t)$, that gives $x(t) = e^{A\,t}\,\alpha$, where $\alpha$ is a constant vector. Now by applying the constant variation method $\alpha \rightsquigarrow \alpha(t)$, we look for a solution of the form $x(t) = e^{A\,t}\,\alpha(t)$. Then

$$\begin{array}{rrcl} & \dot{x}(t) & = & A\,e^{A\,t}\,\alpha(t) + e^{A\,t}\,\dot{\alpha}(t) \\ & & = & A\,e^{A\,t}\,\alpha(t) + B\,u(t) \\ \Leftrightarrow & \dot{\alpha}(t) & = & e^{-A\,t} B\,u(t) \end{array}$$

Integrating from $t_0$ to $t_f$ gives : $\alpha(t_f) = \alpha(t_0) + \displaystyle\int_{t_0}^{t_f} e^{-A\tau} B\,u(\tau)\,\mathrm{d}\tau$. Then

$$x(t_f) = e^{A\,t_f}\,\alpha(t_f) = e^{A\,t_f}\alpha(t_0) + \int_{t_0}^{t_f} e^{A(t_f-\tau)} B\,u(\tau)\,\mathrm{d}\tau$$

Hence, with $t_f = (i+1)\,h$ and $t_0 = i\,h$, we obtain

$$x((i+1)\,h) = x_{i+1} = e^{A\,h} x_i + \int_{ih}^{(i+1)h} e^{A((i+1)\,h-\tau)} B\,\mathrm{d}\tau\, u_i$$

Therefore

$$x_{i+1} = \mathcal{A}x_i + \mathcal{B}u_i$$

with

$$\mathcal{A} = e^{Ah} \text{ and } \mathcal{B} = \int_0^h e^{A\,s} B \,\mathrm{ds}$$

## C.2 Discretization of a linear quadratic cost

Consider the following linear quadratic continuous cost

$$\frac{1}{2}\left(\int_0^{t_f} (\|R^{1/2}u\|^2 + \|C\,x\|^2)\,\mathrm{dt} + x(t_f)^T S\,x(t_f)\right) \tag{C.2}$$

Considering, for $i = i_0, \ldots, N-1$, $x(t) = x(i\,h) =: x_i$, $u(t) = u(i\,h) =: u_i$, $t \in [i\,h,\ (i+1)\,h[$, the cost (C.2) becomes :

$$\begin{array}{rl} & \displaystyle\frac{1}{2}\left(\sum_{i=i_0}^{N-1}\int_{ih}^{(i+1)\,h} (\|R^{1/2}u\|^2 + \|C\,x\|^2)\,\mathrm{dt} + x_{(i+1)h}^T S\,x_{(i+1)h}\right) \\ = & \displaystyle\frac{1}{2}\left(\sum_{i=i_0}^{N-1}(\|R^{1/2}u_i\|^2 + \|C\,x_i\|^2)\,((i+1)h - ih) + x_N^T S\,x_N\right) \end{array}$$

Then we obtain the following discrete time cost :

$$\frac{1}{2}\left(\sum_{i=i_0}^{N-1} h\,(\|R^{1/2}u_i\|^2 + \|C\,x_i\|^2) + x_N^T S\,x_N\right) \tag{C.3}$$

## C.3 Discretization of the adjoint equation and associated multipliers

Consider the following adjoint equation associated to a minimal energy problem :

$$\dot{p}(t) = -A^T p(t) + \lambda(t) \tag{C.4}$$

In order to transform this system into a discrete time system, consider $t_0 = i_0\,h$, $t_f = N\,h$, $p(t) = p(i\,h) =: p_i$, $\lambda(t) = \lambda(i\,h) =: \lambda_i$, where $h$ is the sampling time and $t \in [i\,h,\ (i+1)\,h[$, for $i = i_0, \ldots, N-1$. First integrate the homogeneous part of (C.4), i.e. $\dot{p}(t) = -A^T\,p(t)$, that gives $p(t) = e^{-A^T\,t}\,\beta$, where $\beta$ is a constant vector. Now by applying the constant variation method $\beta \rightsquigarrow \beta(t)$, $p(t)$ is of the form $p(t) = e^{-A^T\,t}\,\beta(t)$, such that

$$\begin{array}{rcl} \dot{p}(t) & = & -A^T\,e^{-A^T\,t}\,\beta(t) + e^{-A^T\,t}\,\dot{\beta}(t) \\ & = & -A^T\,e^{-A^T\,t}\,\beta(t) + \lambda(t) \\ \Leftrightarrow\quad \dot{\beta}(t) & = & e^{A^T\,t}\lambda(t) \end{array}$$

Integrating from $t_0$ to $t_f$ gives : $\beta(t_f) = \beta(t_0) + \int_{t_0}^{t_f} e^{A^T \tau} \lambda(\tau)\, \mathrm{d}\tau$. Then

$$p(t_f) = e^{-A^T t_f}\, \beta(t_f) = e^{-A^T t_f}\, \underbrace{\beta(t_0)}_{=p_0} + \int_{t_0}^{t_f} e^{-A^T (t_f - \tau)} \lambda(\tau)\, \mathrm{d}\tau$$

Hence, with $t_f = (i+1)\, h$ and $t_0 = i\, h$, we obtain

$$p((i+1)\, h) = p_{i+1} = e^{-A^T h} p_i + \int_{ih}^{(i+1)h} e^{-A^T ((i+1)\, h - \tau)} \lambda(\tau)\, \mathrm{d}\tau$$

Therefore

$$p_{i+1} = e^{-A^T h} p_i + \int_0^h e^{-A^T s}\, \mathrm{d}s\, \lambda_i$$

and

$$p_i = \mathcal{A}^T p_{i+1} - \int_0^h e^{-A^T s}\, \mathrm{d}s\, \lambda_i \quad \text{with } \mathcal{A} = e^{Ah}.$$

Now, the adjoint equation in discrete time is given by using the recurrent Hamiltonian equation (10.6), see Section 10.2,

$$p_i^d = A^T p_{i+1}^d - \lambda_i^d$$

Hence

$$\lambda_i^d = \int_0^h e^{-A^T s}\, \mathrm{d}s\, \lambda_i^c \tag{C.5}$$

where $\lambda_i^d$ and $\lambda_i^c$ denote the multiplier $\lambda_i$ in discrete time and the multiplier coming from the discretization, respectively.

## C.4 Application to the numerical example of Section 4.3

Consider the numerical example of subsection 4.3.2 with $t_0 = 0$, and

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ i.e. } e^{At} = \begin{bmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{bmatrix}.$$

Hence, one has

$$\mathcal{A} = \begin{bmatrix} \cosh(h) & \sinh(h) \\ \sinh(h) & \cosh(h) \end{bmatrix} \text{ and } \mathcal{B} = \begin{bmatrix} \cosh(h) - 1 \\ \sinh(h) \end{bmatrix}$$

and (C.5) becomes :

$$\begin{aligned} \lambda_i^d &= \int_0^h \begin{bmatrix} \cosh(s) & -\sinh(s) \\ -\sinh(s) & \cosh(s) \end{bmatrix} \mathrm{d}s\, \lambda_i^c \\ &= \begin{bmatrix} \sinh(s) & 1 - \cosh(s) \\ 1 - \cosh(s) & \sinh(s) \end{bmatrix} \begin{bmatrix} 0 \\ \lambda_2^c \end{bmatrix} \end{aligned}$$

where $\lambda_i^c = [0 \quad \lambda_2^c]^T$, for $x_0 = [1 \quad 0]^T$ and $x_0 = [0 \quad 1]^T$, whatever the interval considered. Then

$$\begin{bmatrix} \lambda_1^d \\ \lambda_2^d \end{bmatrix} = \begin{bmatrix} (1 - \cosh(h))\,\lambda_2^c \\ \sinh(h)\,\lambda_2^c \end{bmatrix} \simeq \begin{bmatrix} 0 \\ h\,\lambda_2^c \end{bmatrix}$$

with $1 - \cosh(h) \simeq 0$ and $\sinh(h) \simeq h$. Therefore,

$$\lambda^d \simeq h\,\lambda^c \tag{C.6}$$

Hence, one can compare the discrete time multipliers with the discretized multipliers by considering the scaling factor $h$ between them.

# Bibliography


[AM90] B. D. O. ANDERSON and J. B. MOORE. *Optimal Control - Linear Quadratics Methods*. Prentice Hall, 1990.

[AS03] D. ANGELI and E. D. SONTAG. *Monotone control systems*. IEEE Transactions on Automatic Control, vol. 48, no. 10, pp. 1684–1698, 2003.

[Ava00] N. AVAERT. *Système dynamique linéaire non négatif*. Master's thesis, Mémoire de deuxième licence en sciences mathématiques, Facultés Universitaires Notre-Dame de la Paix, Namur, 1999-2000.

[AZ00] F. ALLGÖWER and T. ZHENG, editors. *Nonlinear Model Predictive Control*. Birkhäuser Verlag, 2000.

[BCR$^+$02] R. BRU, L. CACCETTA, S. ROMERO, V. RUMCHEV and E. SANCHEZ. *Recent development in reachability and controllability of positive linear systems*. In *Proceedings of the 15th IFAC World Congress, Barcelona, Spain*. 2002.

[BD90] G. BASTIN and D. DOCHAIN. *On-Line Estimation and Adaptive Control of Bioreactors*. Elsevier, 1990.

[Bea06] C. BEAUTHIER. *Le problème linéaire quadratique positif*. Master's thesis, Mémoire de DEA, FUNDP, Namur, 2006.

[BEFB94] S. BOYD, L. EL GHAOUI, E. FERON and V. BALAKRISHNAN. *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics, Philadelphia, 1994.

[BF02] L. BENVENUTI and L. FARINA. *Positive and compartmental systems*. IEEE Transactions on Automatic Control, vol. 47, no. 2, 2002.

[BF04] L. BENVENUTI and L. FARINA. *A tutorial on the positive realization problem*. IEEE Transactions on Automatic Control, vol. 49, pp. 651–664, 2004.

[BHW85] G. BUTLER, S. HSU and P. WALTMAN. *A mathematical model of the chemostat with periodic washout rate*. SIAM Journal on applied mathematics, vol. 45, pp. 435–449, 1985.

[Bix92] R. E. BIXBY. *Implementation of the simplex method : the initial basis*. ORSA Journal on Computing, vol. 4, no. 3, 1992.

[Bla99] F. BLANCHINI. *Set invariance in control – a survey*. Automatica, vol. 35, no. 11, pp. 1747–1768, 1999.

[BNS89] A. BERMAN, N. NEUMANN and R. J. STERN. *Nonnegative Matrices in Dynamic Systems*. J. Wiley and Sons, 1989.

[BP74] A. BERMAN and R. J. PLEMMONS. *Inverses of nonnegative matrices*. Linear and Multilinear Algebra, vol. 2, pp. 161–172, 1974.

[BP94] A. BERMAN and R. J. PLEMMONS. *Nonnegative matrices in the mathematical sciences*. SIAM Classics in applied mathematics, vol. 9, 1994.

[BR97] R. B. BAPAT and T. E. S. RAGHAVAN. *Nonnegative Matrices and Applications*. Encyclopedia of mathematics and its applications 64, Cambridge, 1997.

[BW85] G. J. BUTLER and G. S. K. WOLKOWICZ. *A mathematical model of the chemostat with a general class of functions describing nutrient uptake*. SIAM Journal on applied mathematics, vol. 45, pp. 138–151, 1985.

[BW08] C. BEAUTHIER and J. J. WINKIN. *Finite horizon LQ-optimal control for continuous time positive systems*. In *Proceedings of the Eighteenth International symposium on Mathematical Theory of Networks and Systems (MTNS2008), Virginia Tech, Blacksburg, Virginia, USA, CD-ROM Paper Nr 054*. 2008.

[BW09] C. BEAUTHIER and J. J. WINKIN. *On the positive LQ-problem for linear discrete time systems*. In *Positive Systems, Proceedings of the Third Multidisciplinary International Symposium on Positive Systems: Theory and Applications (POSTA 2009)*, pp. 45–53. 2009.

[BW10] C. BEAUTHIER and J. J. WINKIN. *LQ-optimal control of positive linear systems*. Optimal control : Applications and Methods, vol. 31, no. 6, pp. 547–566, 2010.

[CBHB09] V. CHELLABOINA, S. BHAT, W. M. HADDAD and D. S. BERNSTEIN. *Modeling and analysis of mass-action kinetics*. IEEE Control Systems Magazine, vol. 29, no. 4, pp. 60–78, 2009.

[CD91] F. M. CALLIER and C. A. DESOER. *Linear System Theory*. Springer-Verlag, New York, 1991.

[CH93] E. B. CASTELAN and J. C. HENNET. *On invariant polyhedra of continuous-time linear systems*. IEEE Transactions on Automatic Control, vol. 38, no. 11, pp. 1680–1685, 1993.

[CJ89] R. CASTELEIN and A. JOHNSON. *Constrained optimal control*. IEEE Transactions on Automatic Control, vol. 34, no. 1, pp. 122–126, 1989.

[CW81] F. M. CALLIER and J. L. WILLEMS. *Criterion for the convergence of the solution of the Riccati differential equation*. IEEE Transactions on Automatic Control, vol. 26, no. 6, pp. 1232–1242, 1981.

[CW95] F. M. CALLIER and J. J. WINKIN. *Convergence of the time-invariant Riccati differential equation towards its strong solution for stabilizable systems*. Journal of mathematical analysis and applications, vol. 192, pp. 230–257, 1995.

[CW96] F. M. CALLIER and J. J. WINKIN. *Asymptotic behavior of the solution of the projection Riccati differential equation*. IEEE Transactions on Automatic Control, vol. 41, no. 5, pp. 646–659, 1996.

[CWW94] F. M. CALLIER, J. J. WINKIN and J. L. WILLEMS. *Convergence of the time-invariant Riccati differential equation and LQ-problem : mechanisms of attraction*. International journal of control, vol. 59, no. 4, pp. 983–1000, 1994.

[DAS06] P. DE LEENHEER, D. ANGELI and E. D. SONTAG. *Crowding effects promote coexistence in the chemostat*. J. Math. Anal. Appl., vol. 319, pp. 48–60, 2006.

[DL04] P. DE LEENHEER. *Notes on monotone systems*, Oct. 2004. Course whithin the framework of the Graduate School in systems and control, Louvain-la-neuve, UCL.

[DS02] P. DE LEENHEER and H. SMITH. *Feedback control for a chemostat with two organisms*. Cédérom proceedings of the Symposium on Mathematical Theory of Networks and Systems (MTNS), University of Notre-Dame, 2002.

[DS03] P. DE LEENHEER and H. SMITH. *Feedback control for chemostat models*. Journal of Mathematical Biology, vol. 46, pp. 48–70, 2003.

[FR00] L. FARINA and S. RINALDI. *Positive Linear Systems : Theory and Applications*. John Wiley, New-York, 2000.

[GL00a] C. GUO and A. J. LAUB. *On a Newton-like method for solving algebraic Riccati equations*. SIAM J. Matrix Anal. Appl., vol. 21, no. 2, pp. 694–698, 2000.

[GL00b] C. GUO and A. J. LAUB. *On the iterative solution of a class of nonsymmetric algebraic Riccati equations*. SIAM J. Matrix Anal. Appl., vol. 22, no. 2, pp. 376–391, 2000.

[God83] K. GODFREY. *Compartmental Models and Their Applications*. Academic Press, London, 1983.

[Goe10] R. GOEBEL. *The value function for the linear-quadratic regulator with conical control constraints*. In *Proceedings of the 49th IEEE Conference on Decision and Control, Atlanta, USA*. December 2010.

[Guo01] C. GUO. *Nonsymmetric algebraic Riccati equations and Wiener-Hopf factorization for M-matrices*. SIAM J. Matrix Anal. Appl., vol. 23, no. 1, pp. 225–242, 2001.

[HCH10] W. M. HADDAD, V. CHELLABOINA and Q. HUI. *Nonnegative and Compartmental Dynamical Systems*. Princeton University Press, 2010.

[HCPH10] L. HAN, M. K. CAMLIBEL, J. S. PANG and W. P. M. H. HEEMELS. *Convergence of discrete-time approximations of constrained linear-quadratic optimal control problems*. In *Proceedings of the 49th IEEE Conference on Decision and Control, Atlanta, USA*. December 2010.

[HJ85] R. A. HORN and C. R. JOHNSON. *Matrix Analysis*. Cambridge University Press, 1985.

[HJ91] R. A. HORN and C. R. JOHNSON. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[HS83] J. HALE and A. S. SOMOLINAS. *Competition for fluctuating nutrient*. J. Math. Biol., vol. 18, 1983.

[Hsu80] S. B. HSU. *A competition model for a seasonally fluctuating nutrient*. J. Math. Biol., vol. 9, pp. 115–132, 1980.

[HSV95] R. F. HARTL, S. P. SETHI and R. G. VICKSON. *A survey of the maximum principles for optimal control problems with state constraints*. SIAM Review, vol. 37, no. 2, pp. 181–218, 1995.

[Hut61] G. E. HUTCHINSON. *The paradox of the plankton*. The American Naturalist, vol. 95, no. 882, pp. 137–145, 1961.

[HVS98] W. P. M. H. HEEMELS, S. J. L. VAN EIJNDHOVEN and A. A. STOORVOGEL. *Linear quadratic regulator problem with positive controls*. Int. J. Control, vol. 70, no. 4, pp. 551–578, 1998.

[Jac09] A. C. JACQUES. *Le problème linéaire-quadratique positif inverse : analyse et résolution par les inéquations matricielles linéaires*. Master's thesis, Mémoire de deuxième master en sciences mathématiques, Facultés Universitaires Notre-Dame de la Paix, Namur, 2008-2009.

[JK04] G. JANK and D. KREMER. *Open loop nash games and positive systems - solvability conditions for nonsymmetric Riccati equations*. Lehrstuhl II für Mathematik, RWTH Aachen, Germany, 2004.

[Joh94] A. JOHNSON. *LQ state-constrained control*. In *Proceedings of the IEEE/IFAC Joint Symposium on Computer-Aided Control System Design, Tucson, Arizon, USA*, pp. 423–428. 1994.

[Ka 01] T. KAČZOREK. *Externally and internally positive time-varying linear systems*. Int. J. Appl. Math. Comput. Sci., vol. 11, no. 4, pp. 957–964, 2001.

[Ka 02] T. KAČZOREK. *Positive 1D and 2D Systems*. Springer, London, 2002.

[Ka 03] T. KAČZOREK. *Locally positive nonlinear systems*. Int. J. Appl. Math. Comput. Sci., vol. 13, no. 4, pp. 505–509, 2003.

[Kha02] H. K. KHALIL. *Nonlinear Systems, Third Edition*. Prentice Hall, 2002.

[KS72] H. KWAKERNAAK and P. SIVAN. *Linear Optimal Control Systems*. Wiley-Interscience, New-York, 1972.

[LH06] C. LOBRY and J. HARMAND. *A new hypothesis to explain the coexistence of $n$ species in the presence of a single resource*. C. R. Biologies, vol. 329, pp. 40–46, 2006.

[Loc01] A. LOCATELLI. *Optimal Control : An Introduction*. Birkhäuser Verlag, Basel, 2001.

[LT85] P. LANCASTER and M. TISMENETSKY. *The Theory of Matrices*. Second edition with applications, Computer Science and Applied Mathematics, Academic Press, 1985.

[Lue79] D. G. LUENBERG. *Introduction to Dynamic Systems : Theory, Models and Applications*. John Wiley and Sons, Inc. New York, 1979.

[LWB06] M. LAABISSI, J. J. WINKIN and C. BEAUTHIER. *On the positive LQ-problem for linear continuous-time systems*. In *Lecture Notes in Control and Information Sciences, Proceedings of the second Multidisciplinary International Symposium on Positive Systems : Theory and Applications (POSTA 06), Grenoble, France*, vol. 341, pp. 295–302. 2006.

[Min88] H. MINC. *Nonnegative Matrices*. Wiley Interscience Series in Discrete Time Mathematics and Optimization, John Wiley and Sons, 1988.

[MPS90] K. N. MURTY, K. R. PRASAD and M. A. S. SRINIVAS. *Upper and lower bounds for the solution of the general matrix Riccati differential equation*. Journal of mathematical analysis and applications, vol. 147, pp. 12–21, 1990.

[Nag86] R. NAGEL (ED.). *One-parameter Semigroups of Positive Operators*, vol. 1184. Lecture Notes in Mathematics, 1986.

[PC72] R. J. PLEMMONS and R. E. CLINE. *The generalized inverse of a nonnegative matrix*. In *Proceedings of the american mathematical society*, vol. 31. 1972.

[RD09] B. ROSZAK and E. J. DAVISON. *Necessary and sufficient conditions for stabilizability of positive LTI systems*. Systems and Control Letters, vol. 58, pp. 474–481, 2009.

[RDH09] A. RAPAPORT, D. DOCHAIN and J. HARMAND. *Long run coexistence in the chemostat with multiple species*. Journal of Theoritical Biology, vol. 257, pp. 252–259, 2009.

[RHM07] A. RAPAPORT, J. HARMAND and F. MAZENC. *Coexistence in the design of a series of two chemostats*. Nonlinear Analysis : Real World Applications, 2007.

[SFA79] G. STEPHANOPOULOS, A. G. FREDRICKSON and R. ASIS. *The growth of competing microbial populations in CSTR with periodically varying inputs*. Amer. Instit. Chem. Eng. J., vol. 25, pp. 863–872, 1979.

[SGL94] M. G. SAFONOV, K. C. GOH and J. LY. *Control system synthesis via bilinear matrix inequalities*. In I. PRESS, editor, *Proceedings of the American Control Conference, Baltimore*, pp. 45–49. 1994.

[Smi81] H. L. SMITH. *Competitive coexistence in an oscillating chemostat*. SIAM J. Appl. Math, vol. 40, pp. 498–522, 1981.

[Smi95] H. L. SMITH. *Monotone Dynamical Systems : An Introduction to the Theory of Competitive and Cooperative Systems*. American Mathematical Society, Providence, 1995.

[SW95] H. L. SMITH and P. WALTMAN. *The Theory of the Chemostat : Dynamics of Microbial Competition, Cambridge Studies in Mathematical Biology*, vol. 13. 1995.

[SW05] C. SCHERER and S. WEILAND. *Linear matrix inequalities in control*. Delft University of Technology, Netherlands, 2005.

[Van97] J. M. VAN DEN HOF. *Realization of continuous-time positive linear systems*. Systems and control letters, vol. 31, pp. 243–253, 1997.

[Van98] J. M. VAN DEN HOF. *Positive linear observers for linear compartmental systems*. SIAM J. Control Optimization, vol. 36, no. 2, 1998.

[Van07] J. H. VAN SCHUPPEN. *Control and System Theory of Positive Systems*. Lecture Notes, 2007.

[Van08] J. H. VAN SCHUPPEN. *Problems of control and system theory of biochemical reaction systems*. In *Proceedings of the Eighteenth International symposium on Mathematical Theory of Networks and Systems (MTNS2008), Virginia Tech, Blacksburg, Virginia, USA*. 2008.

[VB00] J. G. VANANTWERP and R. D. BRAATZ. *A tutorial on linear and bilinear matrix inequalities*. Journal of Process Control, vol. 10, pp. 363–385, 2000.

[WC83] J. L. WILLEMS and F. M. CALLIER. *Large finite horizon and infinite horizon LQ-optimal control problems*. Optimal control : Applications and Methods, vol. 4, no. 1, pp. 31–45, 1983.

[Yal] `http://users.isy.liu.se/johanl/yalmip/`.

[Zhe10] T. ZHENG, editor. *Model Predictive Control*. Sciyo, 2010.