Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

Corso di Laurea Magistrale in Bioingegneria

An Introduction to Positive Switched Systems and their Application to HIV Treatment Modeling

Laureanda:
Irene Zorzan

Relatore:
Prof.ssa Maria Elena Valcher

Contro-relatore:
Prof. Ettore Fornasini

14 Luglio 2014

Anno Accademico 2013-2014

An Introduction to Positive Switched Systems and their Application to HIV Treatment Modeling

«Ho soprattutto arrampicato con la fantasia. Col pensiero sono sempre stato un po' più avanti rispetto alle mie capacità, e ho scalato pareti sempre più ripide, finché a un certo punto nessuna via, per quanto ardita, mi è parsa impossibile.»

(*"La mia vita al limite"*, Reinhold Messner)

Premessa

In questo lavoro di tesi il mondo della Teoria dei Sistemi e della Teoria dei Controlli si fonde all'ambito biomedico: mentre il secondo ha sempre suscitato in me un certo fascino, come è normale che sia per uno studente di Bioingegneria, l'interesse verso il primo, un mondo rigoroso, preciso, in cui nulla sembra essere lasciato al caso, è stata una scoperta, qualcosa nato un po' alla volta tra un corso universitario e l'altro, grazie a chi queste materie me le ha sapute insegnare. Questo è il motivo per cui ho scelto di seguire i corsi di Teoria dei Sistemi e Sistemi Multivariabili, di svolgere una tesi su un argomento poco comune per uno studente di Bioingegneria, ed è anche il motivo per cui spero in futuro di poter lavorare in quest'ambito.

Irene Zorzan

Contents

Summary

In the present work an introduction to positive switched systems is provided, along with an interesting application of this kind of systems to the biomedical area. Reflecting this twofold objective, the thesis is divided into two parts.

In the first part classical theoretical aspects, namely stability and stabilizability, concerning positive switched systems are addressed.

In Chapter 1 positive switched systems are introduced, both in the discrete time case and in the continuous time case. A brief overview on the main approaches to the study of stability and stabilizability issues is also presented, focusing in particular on the Lyapunov function techniques.

Chapter 2 is devoted to the problem of stability under arbitrary switching: a number of results from the literature dealing with conditions ensuring the existence of certain types of Lyapunov functions are collected.

In Chapter 3 stabilizability of discrete-time positive switched systems is addressed. In that respect, some recent works appeared in the literature concerning the design of stabilizing switching sequences by means of different types of Lyapunov functions are presented.

The second part of the work aims at showing the importance of the previously presented theory with an application from the biomedical area, specifically to the problem of drug scheduling in HIV infection treatment.

Chapter 4 provides some essential notions concerning the human immune system and the Human Immunodeficiency Virus, focusing in particular on antiretroviral therapies commonly in use and the problem of the development of drug-resistant viral variants.

In Chapter 5 two different models of viral mutation treatment are presented: the first one deals 4 viral variants and 2 drug therapies that can be administered, while the second one provides a more extensive description of HIV dynamics taking into account 64 viral variants and 3 drug combinations.

Chapter 6 is focused on the problem of treatment scheduling in HIV infection: interesting simulation studies performed by Hernandez-Vargas and coauthors using the above mentioned models are presented. All these results establish the importance of proactive switching among drug regimens in order to manage viral mutation and limit, or delay, the emergence of drug-resistant variants.

Finally, in the Appendix some essential notions on positive systems along with some basic properties of positive matrices and Metzler matrices are provided.

List of symbols

$\mathbb{Z}_+$	set of non-negative integers
$\mathbb{R}$	field of real numbers
$\mathbb{R}_+$	set of non-negative real numbers
$\mathbb{R}^n$	set of n-dimensional real vectors
$\mathbb{R}^{m\times n}$	set of $m \times n$ real matrices
$\mathbb{R}^n_+$	set of n-dimensional vectors with non-negative entries
$\mathbb{R}^{m\times n}_+$	set of $m \times n$ matrices with non-negative entries
$\mathbf{x}^T$	transpose of vector $\mathbf{x}$
$\mathbf{e}_i$	i-th canonical vector
$\mathbf{1}_n$	n-dimensional vector with all entries equal to 1
$[\mathbf{x}_i]_j$	j-th entry of vector $\mathbf{x}_i$
$\lVert\mathbf{x}\rVert$	Euclidean norm of vector $\mathbf{x}$
$\mathbf{x} > 0$	non-negative vector
$\mathbf{x} \gg 0$	strictly positive vector
$\mathbf{x} < 0$	non-positive vector
$\mathbf{x} \ll 0$	strictly negative vector
A^T	transpose of matrix A
I_n	identity matrix of order n
$\mathbb{1}_{n\times n}$	matrix in $\mathbb{R}^{n\times n}$ with all entries equal to 1
A^{-1}	inverse of matrix A
$\det(A)$	determinant of matrix A
$A^{(i)}$	i-th column of matrix A
$\ker_+(A)$	positive kernel of matrix A
$[A_i]_{hk}$	(h,k) entry of matrix A_i
$A > 0$	positive (namely, non-negative) matrix
$A \gg 0$	strictly positive matrix
$A < 0$	negative (namely, non-positive) matrix
$A \ll 0$	strictly negative matrix
$A \succ 0$	positive definite matrix
$A \succeq 0$	positive semidefinite matrix
$A \prec 0$	definite negative matrix
$A \preceq 0$	negative semidefinite matrix
$\mathrm{Cone}(A)$	cone whose generating vectors are the columns of A
$\mathcal{K}^*$	dual cone of cone $\mathcal{K}$
$\lvert a\rvert$	absolute value of the number a
$\Re\{a\}$	real part of complex number a
$[1,p]$	set of integers {1,2,. . . ,p}

$\min_i$	minimum with respect to i
$\max_i$	maximum with respect to i
$\min S$	minimum element of set S
$\max S$	maximum element of set S
$\mathrm{argmin} S$	index of minimum element of ordered set S
$\mathrm{argmax} S$	index of maximum element of ordered set S
$\inf S$	infimum of set S
$\sup S$	supremum of set S

Part I

Stability Theory of Positive Switched Systems

Chapter 1

Positive Switched Systems

In the present chapter we introduce positive switched systems and three different approaches to the investigation of their stability and stabilizability properties. In particular, we focus on Lyapunov functions techniques, which represent the most popular approach.

In what follows, basic notions of positive systems are taken for granted. In that respect, we refer the reader to the Appendix, which provides some interesting properties of positive matrices and Metzler matrices along with a brief characterization of discrete-time and continuous-time positive systems.

1.1 Introduction

By a positive switched system we mean a dynamic system consisting of a family of positive state-space models and a switching law, specifying when and how the switching among the various subsystems takes place. Switching among different models naturally arises as a way to formalize the fact that the behavior of a system changes under different operating conditions, and is therefore represented by different mathematical structures. On the other hand, the positivity constraint is pervasive in engineering as well as in chemical, biological and economic modeling: pressures, absolute temperatures, concentration of substances, population levels, any type of resource measured by a quantity and probabilities are all examples of variables that are confined to be positive or non-negative. While both switched systems and positive systems have attracted a great deal of attention over the past decades, the interest in positive switched systems is relatively recent and strongly motivated by the fact that this class of systems is frequently encountered in many application fields. Typical applications can be found in consensus and synchronization problems [17], wireless power control [33], transmission control problems and congestion control [35] and so on. In addition, a major motivation for studying such positive switched systems comes from the possibility of employing them in system biology and pharmacokinetic: this is the case, for instance, when describing the viral mutation dynamics under drug treatment [11] [10], as we will see in Chapter 5.

From a theoretical point of view, although the main properties of both switched systems and positive systems have been well understood during the past decades, many basic problems concerning positive switched systems remain unanswered and that is why these systems still represent quite a challenge. As far as properties like reachability and controllability are concerned, they cannot be investigated as special cases of the analogous properties for standard switched systems, as positive switched systems are defined on cones rather than on linear spaces (and this is an obvious consequence of the positivity constraint on system matrices and on the soliciting inputs). On the other hand, properties like stability and stabilizability, meanwhile inheriting the general results derived for non-positive switched systems, offer new tools and new testing criteria that find no equivalent in the general case. In this sense, studying switched positive systems is more challenging than studying general switched systems, because one has to combine the features of switched systems and positive systems to obtain more "elegant" results.

1.2 Discrete-time Positive Switched Systems

A *discrete-time positive switched system* (DPSS) is described by the following equation:

$$\mathbf{x}(t+1) = A_{\sigma(t)}\mathbf{x}(t), \quad t \in \mathbb{Z}_+, \tag{1.1}$$

where $\mathbf{x}(t) \in \mathbb{R}^n_+$ denotes the n-dimensional state variable at time t, σ is an arbitrary switching sequence, taking values in the set $[1,p] := 1, 2, \ldots, p$, and for each $i \in [1,p]$ the matrix A_i is an $n \times n$ positive matrix.

The state at any time instant $t \in \mathbb{Z}_+$, starting from the initial condition $\mathbf{x}(0)$ and under the effect of the switching sequence $\sigma(0)$, $\sigma(1)$, $\ldots$, $\sigma(t-1)$, can be expressed as follows:

$$\mathbf{x}(t) = A_{\sigma(t-1)} \ldots A_{\sigma(1)} A_{\sigma(0)} \mathbf{x}(0).$$

The index $i = \sigma(t)$ in (1.1) is called the active mode at the time instant t. In general, the active mode at t may depend not only on the time instant t, but also on the current state $\mathbf{x}(t)$ and/or previous active mode $\sigma(\tau)$ for $\tau < t$. Accordingly, the switching law is usually classified as time-dependent (if it depends on the time t only), state-dependent (if it depends on state $\mathbf{x}(t)$ as well), and with memory (if it also depends on the history of active modes).

1.3 Continuous-time Positive Switched Systems

Even if in the sequel only the stability and stabilizability of discrete-time positive switched systems will be considered, it is convenient to introduce continuous-time positive switched systems, as some results derived for the latter find a straightforward extension to the former.

Analogously to its discrete-time counterpart just introduced, a *continuous-time positive switched system*(CPSS) is described by the following equation:

$$\dot{\mathbf{x}}(t) = A_{\sigma(t)}\mathbf{x}(t), \quad t \in \mathbb{R}_+, \tag{1.2}$$

where $\mathbf{x}(t) \in \mathbb{R}^n_+$ denotes the value of the n-dimensional state variable at time t, σ is an arbitrary switching sequence, taking values in the set $[1,p]$, and for each $i \in [1,p]$ the matrix A_i is an $n \times n$ Metzler matrix. We assume that the switching sequence is piecewise continuous, and hence in every time interval $[0,t]$

there is a finite number of discontinuities, which correspond to a finite number of switching instants $0 = t_0 < t_1 < \cdots < t_k < t$. This actually corresponds to the no-chattering requirement for the continuous-time switched systems (note that this is not an issue in the discrete-time case). Also, we assume that, at the switching time t_l, σ is right continuous.

Given a time interval $[0, t]$, corresponding to a set of switching instants $\{t_0, t_1, \ldots, t_k\}$ satisfying $0 = t_0 < t_1 < \cdots < t_k < t$, the state at the time instant t, starting from the initial condition $\mathbf{x}(0)$, can be expressed as follows:

$$\mathbf{x}(t) = e^{A_{i_k}(t-t_k)} \ldots e^{A_{i_0}(t_1-t_0)}\mathbf{x}(0),$$

where $i_l = \sigma(t_l)$, $l = 0, 1 \ldots, k$.

As seen for the discrete-time case, the switching law $\sigma(t)$ can be classified as time-dependent, state-dependent or with memory.

1.4 Stability and stabilizability

Stability and stabilizability issues for positive switched systems include several interesting phenomena, all of them inherited from general switched systems. In this regard, we immediately point out two remarkable facts: switching between individually stable subsystems may cause instability and conversely, switching between unstable subsystems may yield a stable switched system. As these examples suggest, the stability of (positive) switched systems depends not only on the dynamics of each subsystems but also on the properties of the switching signals. Therefore, when stability analysis for autonomous positive switched systems is considered, many questions arise. First of all, we may look for conditions ensuring the asymptotic stability of the system under arbitrary switching rules. Secondly, if all individual subsystems are stable, we may want to calculate a lower bound on the dwell time so as to guarantee the convergence to zero of the state trajectory. If we need to be less restrictive, we might consider an average dwell time, which allows the possibility of switching fast when necessary and then compensating for it by switching sufficiently slowly later. Furthermore, we may study stability under constrained switching when each subsystem is associated with a closed convex region of the positive orthant and can only be active for states within that region. Finally, we may want to specify a time-dependent or a state-dependent switching rule that makes the resulting system asymptotically stable.

Generally speaking, three different approaches to the investigation of stability and stabilizability of (positive) switched systems can be found in the literature. The first one is based on the evaluation of the joint spectral radius of a finite set of matrices $\mathcal{A} := \{A_i, i \in [1, p]\}$, namely the evaluation of $\rho(\mathcal{A}) := \limsup_{k\to+\infty} \max\{\rho(A_{i_1}A_{i_2} \ldots A_{i_k})^{1/k} \colon A_{i_l} \in \mathcal{A}\}$: the asymptotic stability of the DPSS switching into the set of matrices $\mathcal{A}$ is thus equivalent to the fact that $\rho(\mathcal{A})$ is smaller than 1. An alternative line of research, the so-called variational approach, is based on identifying the most critical switching sequence and investigating the resulting system behavior. The basic idea is rather intuitive: if the worst case trajectory, namely the "most unstable" trajectory, is stable, then the whole system should be stable as well. Finally, the most popular approach is undoubtedly the one based on Lyapunov functions, which lays its foundations on Lyapunov stability theory[1] and has a rather significant advantage: it captures the very nature of positive systems, namely the fact that their states are always non-negative. As it will be clearer later, this method consists in finding conditions which guarantee the existence of certain types of Lyapunov functions.

We have so far presented a brief overview on stability and stabilizability problems and the main approaches to their solution which can be found in the literature. Owing to the great extent of the subject, in what follows we limit us to stability under arbitrary switching and state-feedback stabilization. Both topics are investigated by resorting to Lyapunov functions techniques, which are the object of the following section.

1.4.1 Lyapunov functions

The main purpose of this section is to introduce those types of Lyapunov functions we will encounter in the following chapters. This is far away from providing a full and deep characterization of Lyapunov functions, which goes beyond the objectives of the present work.

In general, a Lyapunov function for a system $\mathbf{x}(t+1) = A_i\mathbf{x}(t)$ or $\dot{\mathbf{x}}(t) = A_i\mathbf{x}(t)$ is a positive definite function $V(\mathbf{x})\colon \mathbb{R}^n \to \mathbb{R}$ having the property that it is decreasing along the system trajectories. This amounts to saying that its difference $\Delta V_i(\mathbf{x}) := V(A_i\mathbf{x}) - V(\mathbf{x})$ (for discrete-time systems) or derivative $\dot{V}_i(\mathbf{x})$ (for continuous-time systems) taken along the system's trajectories is negative. When positive systems are considered, the need to evaluate the asymptotic system evolution only on the positive orthant allows to employ a larger class of Lyapunov functions, namely *copositive* Lyapunov functions, which means that $V(\mathbf{x})$ is strictly positive for every $\mathbf{x} > 0$ and becomes zero at the origin. We now introduce three classes of copositive functions:

- *linear copositive functions*:
$$V(\mathbf{x}) = \mathbf{v}_i^T\mathbf{x}, \quad \text{with } \mathbf{v}_i \in \mathbb{R}^n \text{ and } \mathbf{v}_i \gg 0;$$
- *quadratic copositive functions*:
$$V(\mathbf{x}) = \mathbf{x}^T P_i \mathbf{x}, \quad \text{with } P_i = P_i^T \in \mathbb{R}^{n\times n} \text{ such that } \mathbf{x}^T P_i \mathbf{x} > 0 \ \forall \mathbf{x} > 0;$$
- *quadratic positive definite functions*:
$$V(\mathbf{x}) = \mathbf{x}^T P_i \mathbf{x}, \quad \text{with } P_i = P_i^T \succ 0.$$

When dealing with positive switched systems, namely when the index i takes values in the set $[1,p]$, finding a copositive Lyapunov function (belonging to any of the previous three classes) such that $\Delta V_i(\mathbf{x})$ (or $\dot{V}_i(\mathbf{x})$) is negative for every $\mathbf{x} > 0$ and for each $i \in [1,p]$ means ensuring the asymptotic stability of the system under arbitrary switching signals. The Lyapunov function $V(\mathbf{x})$ is a

[1] We recall that Lyapunov's theory is inspired by the concept of energy and energy-dissipation (or preservation): the main idea is based on the fact that if an equilibrium point of a dynamical system is the local minimum of an energy function and the system is dissipative, then the equilibrium is (locally) stable.

common copositive Lyapunov function for the positive switched system if the vector $\mathbf{v}_i$ (or the matrix P_i) does not depend on the index i, which amounts to saying that $V(\mathbf{x})$ is a copositive Lyapunov function for each subsystem. Conversely, when the vector $\mathbf{v}_i$ (or the matrix P_i) depends on the active mode, $V(\mathbf{x})$ is a *switched* copositive Lyapunov function. Basically, if there exists a copositive Lyapunov function for each i-th subsystem, then these functions are patched together based on the switching signal $\sigma(t)$ to construct a global copositive Lyapunov function. For instance, a Switched Linear Copositive Lyapunov Function takes the following form:

$$V(t, \mathbf{x}(t)) = \left(\sum_{i=1}^{p} \eta_i(t) \mathbf{v}_i \right)^T \mathbf{x}(t),$$

where $\eta_i(t) = \begin{cases} 1, & \text{if } \sigma(t) = i \\ 0, & \text{otherwise.} \end{cases}$

A slightly different reasoning, but based on exactly the same principle, allows to determine stabilizing switching signals. In this case, indeed, we look for copositive Lyapunov functions (either linear, quadratic or quadratic positive definite) satisfying the following condition: for every $\mathbf{x} > 0$ there exists $i \in [1, p]$ such that $\Delta V_i(\mathbf{x})$ (or $\dot{V}_i(\mathbf{x})$) is negative. Alternatively, we consider *piecewise* copositive Lyapunov functions, which are the minimum of p linear or quadratic terms, i.e. the minimum of p individual Lyapunov functions. Thus, we have a Piecewise Linear Copositive Lyapunov Function of the form:

$$V(\mathbf{x}) = \min_{i=1,\dots,p} \mathbf{v}_i^T \mathbf{x},$$

or a Piecewise Quadratic Copositive Lyapunov Function of the form:

$$V(\mathbf{x}) = \min_{i=1,\dots,p} \mathbf{x}^T P_i \mathbf{x}.$$

As already mentioned, this is not an exhaustive presentation of Lyapunov functions and many other types of Lyapunov functions can be found in the literature, such as multiple Lyapunov functions, control Lyapunov functions, Lyapunov-like functions and so on.

Before concluding, it is worth making a remark concerning computation tractability of conditions for the existence of a Lyapunov function belonging to any of the previous types. Obviously, we are interested in the derivation of testable conditions, namely compact and easily verifiable conditions, but this goal cannot always be accomplished. Indeed, if these conditions are presented in the form of a set of Linear Matrix Inequalities (LMIs), these can be solved efficiently by means of standard numerical software (such as the LMI Control Toolbox of Matlab). On the contrary, the numerical determination of a solution of a set of non-linear inequalities is not a simple task and looking for an alternative, although certainly more conservative, reformulation might be convenient (this is the case, for instance, when dealing with piecewise copositive Lyapunov functions, whose computation requires the solution of a set of Lyapunov-Metzler inequalities). Alternatively, there have been various attempts to derive algebraic conditions on the subsystems' state matrices for the existence of a certain type of Lyapunov function, since this kind of conditions should be easier to verify and may give us valuable insights in the stability problem.

Chapter 2

Stability Under Arbitrary Switching

In this chapter our main concern is to understand conditions that guarantee the stability of the DPSS when there is no restriction on the switching signals. This problem, which is not a trivial one, is relevant when the switching mechanism is either unknown or too complicated to be useful in the stability analysis. As we will see, even if it has attracted a great deal of attention and a number of interesting results have been obtained, the topic still deserves more investigation as we are far away from a full comprehension of the mechanisms that can lead to instability.

2.1 Preliminary considerations

When dealing with switched systems, the possibility of ensuring a good asymptotic behavior to the system trajectories independently of the switching law requires all individual subsystems to be asymptotically stable. Indeed, if the i-th subsystem is unstable for some $i \in [1, p]$, then the switched system is unstable for $\sigma(t) \equiv i$. Therefore, throughout this chapter it will be assumed that all individual subsystems are asymptotically stable, which amounts to saying that all matrices A_i, $i \in [1, p]$, are positive Schur matrices if we are dealing with DPSSs or Metzler Hurwitz matrices if we are dealing with CPSSs.

This assumption, in general, is not sufficient to guarantee stability under arbitrary switching. However, owing to the fact that all results obtained for general switched systems hold true for positive switched systems, some special cases can be identified. More precisely [24],[25], the above subsystems' stability assumption ensures the existence of a common quadratic Lyapunov function and, hence, the stability of the (positive) switched system, when one at least of the following conditions is satisfied:

- the subsystems' state matrices are pairwise commutative, i.e. $A_i A_j = A_j A_i$ for all $i, j \in [1, p]$;

- all the subsystems' state matrices are symmetric, i.e. $A_i = A_i^T$ for all $i \in [1, p]$;

- all the subsystems' state matrices are normal, i.e. $A_i A_i^T = A_i^T A_i$ for all $i \in [1, p]$.

When none of the previous conditions holds, unconstrained switching may destabilize a switched system even if all individual subsystems are stable. An illustrating example for the discrete-time case is given in Example 2.1 below, while [29] provides an example for the continuous-time case.

Example 2.1. Consider the pair of positive, Schur matrices:

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

It is clear that the matrix:

$$A_1 A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

is not Schur, and hence the state trajectory corresponding to the periodic switching sequence:

$$\sigma(t) = \begin{cases} 2, & \text{if } t \text{ is even;} \\ 0, & \text{if } t \text{ is odd,} \end{cases}$$

does not converge to zero corresponding to every positive $\mathbf{x}(0)$.

Leaving aside the aforementioned special cases, in what follows we are interested in determining what additional requirements must be imposed on the DPSS (1.1) to ensure stability under all possible switching signals. To this aim, common Copositive Lyapunov Functions (CLFs) will be first considered, and, later on, the switched linear Copositive Lyapunov Function method will be presented.

2.2 Common Copositive Lyapunov Functions

It is well known [7],[20],[27],[28],[30] that the existence of a common Copositive Lyapunov Function, either linear, quadratic or quadratic positive definite, represents a sufficient condition for asymptotic system stability. While studying the existence of such Lyapunov functions for switched systems is certainly conservative, establishing conditions under which such functions exist is a natural place to begin the study of the stability of positive switched systems.

Before proceeding, it is worth mentioning one of the first works dealing with discrete-time switched systems (note that most of the results obtained so far have been derived in the continuous-time case). In [27] Mason and Shorten propose a necessary and sufficient condition for the existence of a Common Quadratic Lyapunov Function when a general second-order switched system with $p = 2$ modes is considered. The result is based on the stability of the matrix pencil formed by the pair of subsystems' state matrices. Given two matrices $M, N \in \mathbb{R}^{n \times n}$, the matrix pencil $\Gamma[M, N]$ is defined as the one-parameter family of matrices $\Gamma[M, N] = \{M + \gamma N \colon \gamma \in [0, \infty)\}$. Formally, the result can be summarized by the following proposition.

Proposition 2.1. Let A_1, A_2 be Schur matrices in $\mathbb{R}^{2 \times 2}$ and let $C(A_i)$ indicate the matrix product $C(A_i) := (A_i - I_n)(A_i + I_n)^{-1}$ for $i = 1, 2$. A necessary and sufficient condition for the existence of a Common Quadratic Lyapunov Function is that the pencils $\Gamma[C(A_1), C(A_2)]$ and $\Gamma[C(A_1), C(A_2)^{-1}]$ consist entirely of Hurwitz matrices.

Of course, generalizing the above algebraic condition to higher dimensional systems or to the case with $p > 2$ modes turns out to be difficult. Thus, research efforts have taken alternative directions and quite interesting results have been obtained.

2.2.1 Common Linear CLFs

In this section we quote necessary and sufficient conditions for the existence of a Common Linear CLF. Notice, first, that a linear copositive function $V(\mathbf{x}) = \mathbf{v}^T\mathbf{x}$, with $\mathbf{v} \gg 0$, is a Common Linear CLF for the DPSS (1.1) if and only if:

$$\Delta V_i(\mathbf{x}) = V(A_i\mathbf{x}) - V(\mathbf{x}) = \mathbf{v}^T A_i \mathbf{x} - \mathbf{v}^T \mathbf{x} \ll 0, \quad \forall i \in [1, p] \text{ and } \forall \mathbf{x} > 0,$$

which amounts to saying that:

$$\mathbf{v}^T (A_i - I_n) \ll 0, \quad \forall i \in [1, p]. \tag{2.1}$$

When dealing with CPSSs, $V(\mathbf{x}) = \mathbf{v}^T\mathbf{x}$ is a Common Linear CLF for the system (1.2) if and only if:

$$\dot{V}_i(\mathbf{x}) = \mathbf{v}^T \dot{\mathbf{x}} = \mathbf{v}^T A_i \mathbf{x} \ll 0, \quad \forall i \in [1, p] \text{ and } \forall \mathbf{x} > 0,$$

which amounts to saying that:

$$\mathbf{v}^T A_i \ll 0, \quad \forall i \in [1, p]. \tag{2.2}$$

Recalling that if A_i is a positive and Schur matrix, then $\bar{A}_i = A_i - I_n$ is a Metzler and Hurwitz matrix, it is straightforward to see that every result derived in the continuous-time case can be applied to DPSSs by substituting the Metzler, Hurwitz matrix A_i in (2.2) with $\bar{A}_i$, which is still a Metzler, Hurwitz matrix.

The interest in the existence of a Common Linear CLF is motivated by the fact that checking whether there exists a strictly positive vector such that (2.1) (or (2.2)) holds just amounts to solve a family of LMIs. This justifies our concern in characterizing, within the class of asymptotically stable positive switched systems, those admitting a Common Linear CLF.

A first characterization is provided in [30], where Mason and Shorten propose a necessary and sufficient condition for a pair of asymptotically stable continuous-time positive systems to have a Common Linear CLF. This condition is given in Proposition 2.2 below and its derivation is based on the following preliminary lemma, whose proof can be found in [28].

Lemma 2.1. Let $A_1, A_2 \in \mathbb{R}^{n\times n}$ be Metzler, Hurwitz matrices such that there exists no non-zero vector $\mathbf{v} > 0$ with $A_i^T\mathbf{v} < 0$ for $i = 1, 2$. Then, there exist $\mathbf{w}_1 \gg 0$, $\mathbf{w}_2 \gg 0$ in $\mathbb{R}^n$ such that:

$$A_1\mathbf{w}_1 + A_2\mathbf{w}_2 = 0.$$

Before stating the foretold result, we need to introduce some notation. Given $A \in \mathbb{R}^{n\times n}$ and an integer i with $1 \le i \le n$, $A^{(i)}$ denotes the i-th column of A. Thus, $A^{(i)}$ denotes the vector in $\mathbb{R}^n$ whose j-th entry is a_{ij} for $1 \le j \le n$. For a positive integer n, we denote the set of all mappings $\pi\colon [1,n] \to [1,2]$ by $\mathcal{C}_{n,2}$. Given two matrices $A_1, A_2 \in \mathbb{R}^{n\times n}$ and a mapping $\pi \in \mathcal{C}_{n,2}$, $A_\pi(A_1, A_2)$ denotes the matrix:

$$A_\pi(A_1, A_2) := \begin{bmatrix} A^{(1)}_{\pi(1)} & A^{(2)}_{\pi(2)} & \dots & A^{(n)}_{\pi(n)} \end{bmatrix}.$$

Thus, $A_\pi(A_1, A_2)$ is the matrix in $\mathbb{R}^{n\times n}$ whose i-th column is the i-th column of $A_{\pi(i)}$ for $1 \le i \le n$. We shall denote the set of all matrices that can be formed in this way by $\mathcal{S}(A_1, A_2)$:

$$\mathcal{S}(A_1, A_2) = \{A_\pi(A_1, A_2) : \pi \in \mathcal{C}_{n,2}\}.$$

Proposition 2.2. Let $A_1, A_2 \in \mathbb{R}^{n\times n}$ be Metzler, Hurwitz matrices. Then, the following statements are equivalent:

(i) $\exists \mathbf{v} \gg 0$ such that $V(\mathbf{x}) = \mathbf{v}^T\mathbf{x}$ is a Common Linear CLF for A_1, A_2;

(ii) The finite set $\mathcal{S}(A_1, A_2)$ consists entirely of Hurwitz matrices.

Proof. (i) $\Rightarrow$ (ii): Assuming there exists a Common Linear CLF amounts to saying there exists some vector $\mathbf{v} \gg 0$ in $\mathbb{R}^n$ such that $\mathbf{v}^T A_i \ll 0$ for $i = 1, 2$. This immediately implies that $\mathbf{v}^T A_i^{(j)} \ll 0$ for $i = 1, 2$ and for every $j \in [1, n]$, and hence:

$$\mathbf{v}^T A \ll 0 \quad \forall A \in \mathcal{S}(A_1, A_2). \tag{2.3}$$

Now note that as A_1, A_2 are Metzler, all matrices belonging to the set $\mathcal{S}(A_1, A_2)$ are also Metzler. Recalling that a Metzler matrix A is Hurwitz if and only of there is some vector $\mathbf{v} \gg 0$ such that $A\mathbf{v} \ll 0$, it follows immediately from (2.3) that each matrix in $\mathcal{S}(A_1, A_2)$ must be Hurwitz.

(ii) $\Rightarrow$ (i): We shall show that if a Common Linear CLF for the matrices A_1, A_2 does not exist, then at least one matrix belonging to the set $\mathcal{S}(A_1, A_2)$ must be non-Hurwitz.

First of all, suppose that there is no non-zero vector $\mathbf{v} > 0$ with $\mathbf{v}^T A_i < 0$ for $i = 1, 2$ (note that this is a stronger assumption than the non-existence of a strictly positive vector $\mathbf{v}$, as stated in (i); we will relax this assumption below). It follows from Lemma 2.1 that there are vectors $\mathbf{w}_1, \mathbf{w}_2$ such that $\mathbf{w}_1 \gg 0$, $\mathbf{w}_2 \gg 0$ and:

$$A_1\mathbf{w}_1 + A_2\mathbf{w}_2 = 0. \tag{2.4}$$

As $\mathbf{w}_1 \gg 0$, $\mathbf{w}_2 \gg 0$, there is some positive definite diagonal matrix $D = \mathrm{diag}(d_1, \dots, d_n)$ in $\mathbb{R}^{n\times n}$ with $\mathbf{w}_2 = D\mathbf{w}_1$. It follows from (2.4) that, for this D, we have:

$$A_1\mathbf{w}_1 + A_2 D\mathbf{w}_1 = (A_1 + A_2 D)\mathbf{w}_1 = 0$$

and hence:

$$\det(A_1 + A_2 D) = 0. \tag{2.5}$$

Now, define for each mapping $\pi \in \mathcal{C}_{n,2}$ the following product: $d_\pi := \prod_{i=1}^n d_i^{\pi(i)-1}$ (note that $d_\pi > 0$ for all $\pi \in \mathcal{C}_{n,2}$ since $d_i > 0$ for all i). In terms of this notation and recalling that the determinant of a matrix is multilinear in the columns, we can express $\det(A_1 + A_2 D)$ as:

$$\begin{aligned}\det(A_1 + A_2 D) &= \det\left[A_1^{(1)} + d_1 A_2^{(1)} \quad A_1^{(2)} + d_2 A_2^{(2)} \quad \dots \quad A_1^{(n)} + d_n A_2^{(n)}\right] = \\ &= \sum_{\pi \in \mathcal{C}_{n,2}} \det(A_\pi(A_1, A_2)) d_\pi.\end{aligned}$$

If all matrices in the set $\mathcal{S}(A_1, A_2)$ were Hurwitz, then $\det(A_\pi(A_1, A_2)) > 0$ for all $\pi \in \mathcal{C}_{n,2}$ if n is even and $\det(A_\pi(A_1, A_2)) < 0$ for all $\pi \in \mathcal{C}_{n,2}$ if n is odd. In either case, this would contradict (2.5) which implies that there are positive real numbers $d_1, \dots, d_n$ for which:

$$\sum_{\pi \in \mathcal{C}_{n,2}} \det(A_\pi(A_1, A_2)) d_\pi = 0.$$

Hence, there must exist at least one $\pi \in \mathcal{C}_{n,2}$ for which $A_\pi(A_1, A_2)$ is non-Hurwitz.

So far, we have shown that if there is no non-zero $\mathbf{v} > 0$ such that $\mathbf{v}^T A_i < 0$ for $i = 1, 2$, then at least one of the matrices $A_\pi(A_1, A_2)$ has to be non-Hurwitz. However, in order to finish the proof we need to extend this result to strictly positive vectors $\mathbf{v}$, as stated in the proposition. So, let us assume that there is no common $\mathbf{v} \gg 0$ such that $\mathbf{v}^T A_i \ll 0$ for $i = 1, 2$. If, additionally, there was no $\mathbf{v} > 0$ such that $\mathbf{v}^T A_i < 0$ for $i = 1, 2$, the result follows from the above discussion. Conversely, if there was such a $\mathbf{v} > 0$, an additional argument is needed.

Denote by $\mathbb{1}_{n \times n}$ the matrix in $\mathbb{R}^{n \times n}$ consisting entirely of ones ($\mathbb{1}_{n \times n}(i, j) = 1$ for all $1 \le i, j \le n$) and by $A_i(\varepsilon)$ the matrix $A_i(\varepsilon) := A_i + \varepsilon \mathbb{1}_{n \times n}$ for $\varepsilon > 0$ and $i = 1, 2$. It then follows that there cannot be a non-zero $\mathbf{v} > 0$ achieving $\mathbf{v}^T A_i(\varepsilon) < 0$ for $i = 1, 2$. This can be shown by argument of contradiction: assume there was a vector $\mathbf{v} > 0$ such that $\mathbf{v}^T A_i(\varepsilon) < 0$, then:

$$\begin{aligned}\mathbf{v}^T A_i(\varepsilon) = \mathbf{v}^T (A_i + \varepsilon \mathbb{1}_{n \times n}) &< 0 \\ \mathbf{v}^T A_i &< 0 - \varepsilon \mathbf{v}^T \mathbb{1}_{n \times n} \\ \mathbf{v}^T A_i &< 0\end{aligned}$$

for $\varepsilon > 0$ and $i = 1, 2$, which contradicts the first assumption. Thus, there is no non-zero $\mathbf{v} > 0$ such that $\mathbf{v}^T A_i(\varepsilon) < 0$ for $i = 1, 2$.

Now, if we choose any $\varepsilon > 0$ sufficiently small to ensure that $A_1(\varepsilon)$ and $A_2(\varepsilon)$ are Hurwitz and Metzler matrices, it follows from the above argument that there must be at least one non-Hurwitz matrix in the set $\mathcal{S}(A_1(\varepsilon), A_2(\varepsilon))$.

Finally, a limiting argument is needed. Consider a sequence of (ε_k) such that $\varepsilon_k \to 0$ as $k \to \infty$ and all ε_k's are small enough so that all $A_i(\varepsilon_k)$ are still Hurwitz and Metzler matrices. Since these matrices, and thus all $A_\pi(A_1(\varepsilon_k), A_2(\varepsilon_k))$, depend continuously on ε_k, it follows for all $\pi \in \mathcal{C}_{n,2}$ that:

$$A_\pi(A_1(\varepsilon_k), A_2(\varepsilon_k)) \to A_\pi(A_1, A_2) \quad \text{as } \varepsilon_k \to 0.$$

Since there is at least one $\pi \in \mathcal{C}_{n,2}$ for which $A_\pi(A_1(\varepsilon_k), A_2(\varepsilon_k))$ is non-Hurwitz, this will also be the case for $A_\pi(A_1, A_2)$. This shows that at least one matrix in the set $\mathcal{S}(A_1, A_2)$ is non-Hurwitz and completes the proof of the proposition. $\square$

The result provided by Proposition 2.2 is generalized to the case of a finite set of continuous-time positive subsystems in [20]. Of course, we need to extend the notation previously introduced. For positive integers n and p, we denote the set of all mappings $\pi\colon [1,n] \to [1,p]$ by $\mathcal{C}_{n,p}$. Given a family of p matrices in $\mathbb{R}^{n\times n}$, $\mathcal{A} = \{A_1, \dots A_p\}$, and a mapping $\pi \in \mathcal{C}_{n,p}$, $A_\pi(A_1, \dots, A_p)$ denotes the matrix:

$$A_\pi(A_1, \dots, A_p) := \begin{bmatrix} A^{(1)}_{\pi(1)} & A^{(2)}_{\pi(2)} & \dots & A^{(n)}_{\pi(n)} \end{bmatrix}.$$

We shall denote by $\mathcal{S}(A_1, \dots, A_p)$ the set of all matrices one obtains by selecting the 1^{st} column among the 1^{st} columns of the matrices in $\mathcal{A}$, the 2^{nd} column among the 2^{nd} columns of the matrices in $\mathcal{A}$ and so on:

$$\mathcal{S}(A_1, \dots, A_p) = \{A_\pi(A_1, \dots, A_p) : \pi \in \mathcal{C}_{n,p}\}.$$

We can now state Theorem 2.1, whose proof is omitted (but can be found in [20]), as it follows closely the lines of that of Proposition 2.2, including a generalization of Lemma 2.1.

Theorem 2.1. Let $A_1, \dots, A_p \in \mathbb{R}^{n\times n}$ be Metzler, Hurwitz matrices. Then, the following statements are equivalent:

(i) $\exists \mathbf{v} \gg 0$ such that $V(\mathbf{x}) = \mathbf{v}^T\mathbf{x}$ is a Common Linear CLF for $A_1, \dots, A_p$;

(ii) The finite set $\mathcal{S}(A_1, \dots, A_p)$ consists entirely of Hurwitz matrices.

This theorem states that p continuous-time positive systems share a linear CLF if and only if each of the p^n Metzler matrices belonging to $\mathcal{S}(A_1, \dots, A_p)$ are Hurwitz matrices. In that case, the switched system formed by these subsystems is asymptotically stable under arbitrary switching. To illustrate this result, we include the following numerical example from [20].

Example 2.2. Consider three Metzler and Hurwitz matrices:

$$A_1 = \begin{bmatrix} -12 & 6 & 6 \\ 1 & -10 & 2 \\ 5 & 3 & -10 \end{bmatrix} \quad A_2 = \begin{bmatrix} -12 & 4 & 0 \\ 6 & -10 & 9 \\ 4 & 3 & -13 \end{bmatrix} \quad A_3 = \begin{bmatrix} -9 & 2 & 8 \\ 6 & -10 & 4 \\ 3 & 0 & -11 \end{bmatrix}$$

It turns out that the matrix $A_\pi(A_1, A_2, A_3)$ is Hurwitz for any $\pi \in \mathcal{C}_{3,3}$ and hence a CPSS with these matrices will be asymptotically stable under arbitrary switching. If, however, the $(3,1)$-element of A_3 is changed from 3 to 5 (note that after change A_3 is still a Metzler and Hurwitz matrix), then the matrix $A_{(3,1,3)} = \begin{bmatrix} A_3^{(1)} & A_1^{(2)} & A_3^{(3)} \end{bmatrix}$ will have an eigenvalue $\lambda \approx 0.06$ which violates the Hurwitz condition.

A complete characterization for the existence of a Common Linear CLF based on the investigation of a new geometric object, the convex hull generated by the columns of the subsystems matrices, is provided in [7]. Presenting these results requires some basic notions on cones we now introduce.

Let us first recall that the *convex hull* of a given family of vectors $\mathbf{w}_1, \dots, \mathbf{w}_s$ is the set of vectors:

$$\mathcal{W}_\mathbf{w} := \left\{ \sum_{i=1}^{s} \alpha_i \mathbf{w}_i \colon \alpha_i \geq 0, \sum_{i=1}^{s} \alpha_i = 1 \right\}.$$

A set $\mathcal{K} \subset \mathbb{R}^n$ is a *cone* if $\alpha\mathcal{K} \subseteq \mathcal{K}$ for all $\alpha \geq 0$. A cone $\mathcal{K}$ is said to be:

- *convex* if it contains, with any two points, the line segment between them;
- *solid* if it is convex and it includes at least one interior point;
- *pointed* if it is convex and $\mathcal{K} \cap \{-\mathcal{K}\} = \{0\}$;
- *polyhedral* if it can be expressed as the set of non-negative linear combinations of a finite set of vectors, called *generating vectors*; if the generating vectors are the columns of a matrix A, we adopt the notation $\mathcal{K} = \text{Cone}(A)$.

The *dual cone* of a cone $\mathcal{K} \subset \mathbb{R}^n$ is:

$$\mathcal{K}^* := \Big\{\mathbf{v} \in \mathbb{R}^n : \mathbf{x}^T\mathbf{v} \geq 0,\ \forall \mathbf{x} \in \mathcal{K}\Big\}.$$

Finally, the following relations between a closed convex cone $\mathcal{K}$ and its dual cone $\mathcal{K}^*$ hold:

- $\mathcal{K}$ is pointed if and only if $\mathcal{K}^*$ is solid;
- $\mathcal{K}$ is solid if and only if $\mathcal{K}^*$ is pointed;
- $\mathcal{K}$ is polyhedral if and only if $\mathcal{K}^*$ is polyhedral.

In order to state the foretold results, we need a technical lemma, whose proof is provided in [7]. In the following, for the sake of simplicity, we denote the orthant of $\mathbb{R}^n$ including vectors with all non-negative entries except for the j-th, which is negative, as $\mathcal{O}_{j-}$.

Lemma 2.2. Let $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s$ be a family of $s \leq n+1$ vectors in $\mathbb{R}^n$, each of them belonging to some orthant $\mathcal{O}_{j-}$, $j \in [1, n]$. Suppose there exists a positive convex combination:

$$\mathbf{y} = \sum_{j=1}^{s} \mathbf{w}_j c_j, \quad c_j > 0,\ \sum_{j=1}^{s} c_j = 1,$$

such that $\mathbf{y}$ is a non-negative vector. If (at least) two vectors of the family, say $\mathbf{w}_1$ and $\mathbf{w}_2$, belong to the same orthant $\mathcal{O}_{j-}$, then a non-negative vector, possibly different from $\mathbf{y}$, can be obtained as a convex combination of a subfamily of $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s$, where either $\mathbf{w}_1$ or $\mathbf{w}_2$ has been removed.

Theorem 2.2. Let $A_1, \ldots, A_p \in \mathbb{R}^{n\times n}$ be Metzler, Hurwitz matrices and let W denote the matrix:

$$W := \begin{bmatrix} I_n & -A_1 & -A_2 & \ldots & -A_p \end{bmatrix} \in \mathbb{R}^{n\times(p+1)n}.$$

Then, the following statements are equivalent:

(i) $\exists \mathbf{v} \gg 0$ such that $V(\mathbf{x}) = \mathbf{v}^T\mathbf{x}$ is a Common Linear CLF for $A_1, \ldots, A_p$;

(ii) $\ker_+(W) = \{0\}$;

(iii) The convex hull of the vector family $\mathcal{W}_A := \{A_i^{(j)} : j \in [1, n], i \in [1, p]\}$ does not intersect the positive orthant $\mathbb{R}^n_+$;

(iv) For every map $\pi \in \mathcal{C}_{n,p}$, the convex hull of the vector family $\mathcal{W}_\pi := \{A^{(j)}_{\pi(j)} : j \in [1,n]\}$ does not intersect the positive orthant $\mathbb{R}^n_+$.

Proof. (i) $\Leftrightarrow$ (ii): Consider the polyhedral cone whose generating vectors are the columns of W:

$$\mathcal{K} := \text{Cone}(W) = \left\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = W\boldsymbol{\lambda},\ \boldsymbol{\lambda} \in \mathbb{R}^{(p+1)n}_+\right\}$$

and its dual cone:

$$\mathcal{K}^* := \left\{\mathbf{v} \in \mathbb{R}^n : \boldsymbol{\lambda}^T W^T \mathbf{v} \geq 0,\ \forall \boldsymbol{\lambda} \in \mathbb{R}^{(p+1)n}_+\right\} =$$
$$= \left\{\mathbf{v} \in \mathbb{R}^n_+ : \begin{bmatrix} I_n \\ -A_1^T \\ \vdots \\ -A_p^T \end{bmatrix} \mathbf{v} \geq 0 \right\}.$$

Notice that a strictly positive vector $\mathbf{v} \in \mathbb{R}^n$ defines a Common Linear CLF for $A_1, \dots, A_p$ if and only if it belongs to the interior of the closed convex cone $\mathcal{K}^*$. Hence, statement (i) is satisfied if and only if $\mathcal{K}^*$ is solid or, equivalently, $\mathcal{K}$ is pointed. However, as W is devoid of zero columns, it is easily seen that $\mathcal{K}$ is pointed if and only if the only non-negative vector in the kernel of W is the zero vector. Indeed, assume $\ker_+(W) = \{0\}$ and let $\mathbf{v} \in \mathcal{K} \cap \{-\mathcal{K}\}$, which amounts to saying $\mathbf{v} = W\boldsymbol{\lambda}_1$ and $-\mathbf{v} = W\boldsymbol{\lambda}_2$ for some $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}^{(p+1)n}_+$. Then $0 = W(\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2)$ implies $\mathbf{v} = W\boldsymbol{\lambda}_1 = 0$ since $\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 \in \ker_+(W)$. Conversely, any non zero vector $\boldsymbol{\lambda} \in \mathbb{R}^{(p+1)n}_+$ belonging to $\ker_+(W)$ can be written as $\boldsymbol{\lambda} = (\boldsymbol{\lambda} - \alpha\mathbf{e}_i) + \alpha\mathbf{e}_i$ where $\mathbf{e}_i$ is a canonical vector and $\alpha > 0$ is small enough to ensure $\boldsymbol{\lambda} - \alpha\mathbf{e}_i$ is still positive. Now, $W(\boldsymbol{\lambda} - \alpha\mathbf{e}_i + \alpha\mathbf{e}_i) = 0$ implies the existence of a non zero vector $\mathbf{v} = W(\boldsymbol{\lambda} - \alpha\mathbf{e}_i)$ belonging to $\mathcal{K}$, whose opposite $-\mathbf{v} = W\alpha\mathbf{e}_i$ belongs to $-\mathcal{K}$. So, we have proved that (i) and (ii) are equivalent statements.

(ii) $\Leftrightarrow$ (iii): There exists a positive vector in $\ker_+(W)$ if and only if there exist non-negative vectors $\mathbf{y}, \mathbf{x}_1, \dots, \mathbf{x}_p$ in $\mathbb{R}^n$ not all of them equal to zero, such that:

$$\begin{bmatrix} I_n & -A_1 & \dots & -A_p \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_p \end{bmatrix} = \mathbf{y} - A_1\mathbf{x}_1 - \dots - A_p\mathbf{x}_p = 0$$

and hence:

$$\mathbf{y} = \sum_{i=1}^p A_i \mathbf{x}_i = \sum_{i=1}^p \sum_{j=1}^n A_i^{(j)} [\mathbf{x}_i]_j.$$

Possibly rescaling $\mathbf{y}$ and the various non-negative coefficients $[\mathbf{x}_i]_j$, we can assume $\sum_{i=1}^p \sum_{j=1}^n [\mathbf{x}_i]_j = 1$, which amounts to saying that the convex hull of the family of vectors $\mathcal{W}_A$ includes a non-negative vector. Therefore, also (ii) and (iii) are equivalent.

(iii) $\Rightarrow$ (iv): The proof is obvious, since for each mapping $\pi \in \mathcal{C}_{n,p}$ the vector family $\mathcal{W}_\pi$ is a subset of the vector family $\mathcal{W}_A$.

(iv) $\Rightarrow$ (iii): Notice, first, that each vector of the families $\mathcal{W}_A$ and $\mathcal{W}_\pi$, being a column of a Metzler Hurwitz matrix, belongs to some orthant $\mathcal{O}_{j-}$, for some $j \in [1, n]$. We now proceed by showing that $(\bar{\text{iii}}) \Rightarrow (\bar{\text{iv}})$. Consider a non-negative vector $\mathbf{y} \in \mathbb{R}^n_+$ obtained as the convex combination of the vectors of $\mathcal{W}_A$. By the Caratheodory's theorem [1], there exist $s \leq n+1$ vectors, say $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_s$ in $\mathbb{R}^n$, such that:

$$\mathbf{y} = \sum_{j=1}^{s} \mathbf{w}_j c_j, \quad c_j > 0, \ \sum_{j=1}^{s} c_j = 1.$$

Starting from the above combination and repeatedly applying Lemma 2.2, we reduce ourselves to the situation when we have vectors, say $\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \ldots, \hat{\mathbf{w}}_r$, with $r \leq \min\{s, n\}$, endowed with the following properties:

- each of them belongs to $\mathcal{W}_A$;
- for every pair of distinct indices $i, j \in [1, r]$, $\hat{\mathbf{w}}_i$ and $\hat{\mathbf{w}}_j$ belong to distinct orthants;
- there exists a convex combination of the vectors $\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \ldots, \hat{\mathbf{w}}_r$ that gives a non-negative vector in $\mathbb{R}^n_+$.

If $r < n$, we complete the r-tuple above by introducing $n - r$ vectors of $\mathcal{W}_A$, each of them belonging to one of the orthants which are not represented by $\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \ldots, \hat{\mathbf{w}}_r$. So, in any case, we end up with an n-tuple of columns, $\{A^{(j)}_{\pi(j)}, j \in [1, n]\}$, that corresponds to a suitable map π, and produces, via convex combination, a non-negative vector in $\mathbb{R}^n_+$. This contradicts (iv). $\square$

Theorem 2.2 relates the existence of a Common Linear CLF to the structure of the convex hull generated by the columns of the subsystems matrices. Again, we provide an illustrating example.

Example 2.3. Consider the pair of Metzler and Hurwitz matrices:

$$A_1 = \begin{bmatrix} -1 & 1 \\ \frac{2}{3} & -\frac{29}{30} \end{bmatrix} \quad A_2 = \begin{bmatrix} -\frac{1}{2} & 10 \\ 0 & -\frac{2}{3} \end{bmatrix}.$$

It turns out that the convex hull of the vector family $\mathcal{W}_A$ intersects the positive orthant $\mathbb{R}^n_+$, as:

$$\mathbf{y} = \frac{2}{3} A_1^{(1)} + \frac{1}{3} A_2^{(2)} = \begin{bmatrix} \frac{8}{3} \\ \\ \frac{2}{9} \end{bmatrix} \in \mathbb{R}^n_+$$

and hence, by Theorem 2.2, there does not exist a Common Linear CLF for A_1 and A_2. Note that this does not imply the existence of some divergent trajectories for a CPSS with these matrices.

Remark 2.1. Both Theorem 2.1 and Theorem 2.2 present necessary and sufficient conditions for the existence of a Common Linear CLF, but the underlying approach is slightly different. On the one hand, Knorn, Mason and Shorten, resorting to positive diagonal transformations and to determinantal properties, provide a characterization in terms of the Hurwitz property of a certain finite

[1] Caratheodory's theorem states that if a vector $\mathbf{x}$ of $\mathbb{R}^n$ belongs to the convex hull of a set $\mathcal{P}$, there is a subset $\bar{\mathcal{P}}$ of $\mathcal{P}$ consisting of $s \leq n+1$ vectors such that $\mathbf{x}$ belongs to the convex hull of $\bar{\mathcal{P}}$.

set of matrices. On the other hand, Fornasini and Valcher exploit the theory of cones to provide conditions involving the convex hulls of certain vector families and their intersection with the positive orthant.

In this regard, it is worth pointing out that the equivalence between statement (ii) in Theorem 2.1 and statement (iv) in Theorem 2.2 directly follows from standard properties of Metzler matrices and the following lemma.

Lemma 2.3. Let U be a real matrix in $\mathbb{R}^{m\times n}$. Then, one and only one of the following alternatives holds:

(a) $\exists \mathbf{w} > 0$ such that $\mathbf{w}^T U \ll 0$;

(b) $\exists \mathbf{z} > 0$ such that $U\mathbf{z} > 0$.

Indeed, if all matrices in the set $\mathcal{S}(A_1,\ldots,A_p)$ are Hurwitz, condition (b) of Lemma 2.3 cannot be verified, and consequently for every map $\pi \in \mathcal{C}_{n,p}$ no convex combination of the columns of $A_\pi(A_1,\ldots,A_p)$ intersects the positive orthant $\mathbb{R}^n_+$. Conversely, if there exists a map $\pi \in \mathcal{C}_{n,p}$ such that the convex hull of $A_\pi(A_1,\ldots,A_p)$ intersects the positive orthant, namely condition (b) of Lemma 2.3 holds true, there does not exist a vector $\mathbf{w} > 0$ such that $\mathbf{w}^T A_\pi(A_1,\ldots,A_p) \ll 0$ and hence $A_\pi(A_1,\ldots,A_p)$ is not Hurwitz.

All results presented in Theorem 2.1 and Theorem 2.2 have been derived for CPSSs. The following corollary puts together such results and restates them for the discrete-time case.

Corollary 2.1. Let $A_1,\ldots,A_p \in \mathbb{R}^{n\times n}$ be positive, Schur matrices and let W denote the matrix:

$$W := \begin{bmatrix} I_n & -(A_1 - I_n) & -(A_2 - I_n) & \ldots & -(A_p - I_n) \end{bmatrix} \in \mathbb{R}^{n\times(p+1)n}.$$

Then, the following statements are equivalent:

(i) $\exists \mathbf{v} \gg 0$ such that $V(\mathbf{x}) = \mathbf{v}^T\mathbf{x}$ is a Common Linear CLF for $A_1,\ldots,A_p$;

(ii) The finite set $\mathcal{S}(A_1,\ldots,A_p)$ consists entirely of Schur matrices;

(iii) $\ker_+(W) = \{0\}$;

(iv) The convex hull of the vector family $\mathcal{W}_{\bar{A}} := \{(A_i - I_n)^{(j)} : j \in [1,n], i \in [1,p]\}$ does not intersect the positive orthant $\mathbb{R}^n_+$;

(v) For every map $\pi \in \mathcal{C}_{n,p}$, the convex hull of the vector family $\mathcal{W}_\pi := \{(A_{\pi(j)} - I_n)^{(j)} : j \in [1,n]\}$ does not intersect the positive orthant $\mathbb{R}^n_+$.

An alternative proof of the equivalence (i) $\Leftrightarrow$ (iv) of Corollary 2.1 can be found in [8], where DPSSs are addressed.

Before concluding, we briefly mention a recent work by Doan and coauthors centered around Collatz-Wielandt sets. In [4] the link between linear Lyapunov functions for DPSSs with irreducible matrices and corresponding Collatz-Wielandt sets is established and this leads to an algorithm to compute a Common Linear CLF whenever it exists. This approach differs considerably from what we have seen so far and its deeper investigation goes beyond this discussion.

2.2.2 Common Quadratic CLFs and Common Quadratic Positive Definite CLFs

A quadratic copositive function $V(\mathbf{x}) = \mathbf{x}^T P\mathbf{x}$, with $P_i = P_i^T \in \mathbb{R}^{n\times n}$, is a Common Quadratic CLF for the DPSS (1.1) if and only if:

$$\begin{aligned}\Delta V_i(\mathbf{x}) &= V(A_i\mathbf{x}) - V(\mathbf{x}) = \\ &= \mathbf{x}^T A_i^T P A_i \mathbf{x} - \mathbf{x}^T P\mathbf{x} = \\ &= \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < 0, \quad \forall i \in [1, p] \text{ and } \forall \mathbf{x} > 0.\end{aligned} \tag{2.6}$$

If, in addition, the matrix P is positive definite, $V(\mathbf{x})$ is said to be a Common Quadratic Positive Definite CLF.

When dealing with CPSSs, $V(\mathbf{x}) = \mathbf{x}^T P\mathbf{x}$ is a Common Quadratic CLF for the system (1.2) if and only if:

$$\begin{aligned}\dot{V}_i(\mathbf{x}) &= \dot{\mathbf{x}}^T P\mathbf{x} + \mathbf{x}^T P\dot{\mathbf{x}} = \\ &= (A_i\mathbf{x})^T P\mathbf{x} + \mathbf{x}^T P(A_i\mathbf{x}) = \\ &= \mathbf{x}^T (A_i^T P + P A_i)\mathbf{x} < 0, \quad \forall i \in [1, p] \text{ and } \forall \mathbf{x} > 0.\end{aligned} \tag{2.7}$$

Comparing (2.6) and (2.7) we see that, unlike linear Lyapunov functions previously considered, conditions concerned with the existence of quadratic (or quadratic positive definite) common CLFs for CPSSs cannot be extended to DPSSs.

Unfortunately, while a number of results have been derived in the continuous-time case, only few works can be found in the literature dealing with quadratic CLF for DPSSs. In that regard, in [7] necessary and sufficient conditions for the existence of Common CLFs belonging to any of the three classes are mutually related, thus proving that if a Common Linear CLF can be found, then a Common Quadratic CLF can be found, too, and this latter, in turn, ensures the existence of a Common Quadratic Positive Definite CLF. Formally, the result can be summarized by the following theorem.

Theorem 2.3. Let $A_1, \dots, A_p \in \mathbb{R}^{n\times n}$ be positive, Schur matrices. Then, the following statements are equivalent:

(i) $\exists \mathbf{v} \gg 0$ such that $V(\mathbf{x}) = \mathbf{v}^T\mathbf{x}$ is a Common Linear CLF for $A_1, \dots, A_p$;

(ii) $\exists P = P^T$ of rank 1 such that $V(\mathbf{x}) = \mathbf{x}^T P\mathbf{x}$ is a Common Quadratic CLF for $A_1, \dots, A_p$.

If (i)-(ii) hold, then the following condition holds:

(iii) $\exists \tilde{P} = \tilde{P}^T \succ 0$ such that $V(\mathbf{x}) = \mathbf{x}^T \tilde{P}\mathbf{x}$ is a Common Quadratic Positive Definite CLF for $A_1, \dots, A_p$.

If (iii) holds, then:

(iv) $\exists P = P^T$ such that $V(\mathbf{x}) = \mathbf{x}^T P\mathbf{x}$ is a Common Quadratic CLF for $A_1, \dots, A_p$.

Proof. (i) $\Rightarrow$ (ii): Let $\mathbf{v}$ be a strictly positive vector such that $\mathbf{v}^T A_i \mathbf{x} < \mathbf{v}^T \mathbf{x}$ for every $i \in [1,p]$ and every $\mathbf{x} > 0$. As all quantities involved are non-negative, for every $i \in [1,p]$ and every $\mathbf{x} > 0$ we have:

$$(\mathbf{v}^T A_i \mathbf{x})^2 < (\mathbf{v}^T \mathbf{x})^2$$
$$\mathbf{x}^T A_i^T \mathbf{v}\mathbf{v}^T A_i \mathbf{x} < \mathbf{x}^T \mathbf{v}\mathbf{v}^T \mathbf{x},$$

and hence (ii) is satisfied for $P = \mathbf{v}\mathbf{v}^T$.

(ii) $\Rightarrow$ (i): Notice, first, that any symmetric matrix P of rank 1 can be expressed as $P = \mathbf{v}\mathbf{v}^T$ for some vector $\mathbf{v}$. Moreover, as $\mathbf{x}^T P \mathbf{x} = \mathbf{x}^T \mathbf{v}\mathbf{v}^T \mathbf{x} = (\mathbf{v}^T \mathbf{x})^2 > 0$ for every $\mathbf{x} > 0$, all entries of $\mathbf{v}$ are nonzero and of the same sign, and it entails no loss of generality assuming that they are all positive. On the other hand, statement (ii) means that for every $i \in [1,p]$ and every $\mathbf{x} > 0$ we have:

$$\mathbf{x}^T A_i^T P A_i \mathbf{x} < \mathbf{x}^T P \mathbf{x}.$$

The previous condition can be rewritten as:

$$\mathbf{x}^T A_i^T \mathbf{v}\mathbf{v}^T A_i \mathbf{x} < \mathbf{x}^T \mathbf{v}\mathbf{v}^T \mathbf{x}$$
$$(\mathbf{v}^T A_i \mathbf{x})^2 < (\mathbf{v}^T \mathbf{x})^2.$$

Now, from the non-negativity of both $\mathbf{v}^T A_i \mathbf{x}$ and $\mathbf{v}^T \mathbf{x}$, we get statement (i), namely:

$$\mathbf{v}^T A_i \mathbf{x} < \mathbf{v}^T \mathbf{x}, \quad \forall i \in [1,p] \text{ and } \forall \mathbf{x} > 0.$$

(ii) $\Rightarrow$ (iii): If P is a symmetric matrix of rank 1 such that $\mathbf{x}^T P \mathbf{x} > 0$ in every point of the positive orthant, except for the origin, then, as shown in (ii) $\Rightarrow$ (i), $P = \mathbf{v}\mathbf{v}^T$ for some $\mathbf{v} \gg 0$. This implies that P is also positive semidefinite. Now, we set $\tilde{P} := P + \varepsilon I_n$, with $\varepsilon > 0$, and we prove that there exists ε such that $V(\mathbf{x}) = \mathbf{x}^T \tilde{P} \mathbf{x}$ is a Common Quadratic Positive Definite CLF for $A_1, \dots, A_p$. Clearly $\tilde{P}$ is positive definite, indeed:

$$\mathbf{x}^T \tilde{P} \mathbf{x} = \mathbf{x}^T P \mathbf{x} + \varepsilon \mathbf{x}^T \mathbf{x} > 0 \quad \forall \mathbf{x} \neq \mathbf{0}.$$

Now, consider the two functions:

$$f(\mathbf{x}) := \max_{i=1,\dots,p} |\mathbf{x}^T (A_i^T A_i - I_n) \mathbf{x}|,$$
$$g(\mathbf{x}) := \max_{i=1,\dots,p} \mathbf{x}^T (A_i^T P A_i - P) \mathbf{x}.$$

Both $f(\mathbf{x})$ and $g(\mathbf{x})$ are continuous in the compact set $\mathcal{E} := \mathbb{R}^n_+ \cap \{\mathbf{x} \in \mathbb{R}^n_+ : \|\mathbf{x}\|_2 = 1\}$ and hence, by Weierstrass' theorem and assumption (ii), we have:

$$\max_{\mathbf{x}\in\mathcal{E}} f(\mathbf{x}) = \max_{\mathbf{x}\in\mathcal{E}} \max_{i=1,\dots,p} |\mathbf{x}^T (A_i^T A_i - I_n) \mathbf{x}| = M \geq 0,$$
$$\max_{\mathbf{x}\in\mathcal{E}} g(\mathbf{x}) = \max_{\mathbf{x}\in\mathcal{E}} \max_{i=1,\dots,p} \mathbf{x}^T (A_i^T P A_i - P) \mathbf{x} = -\delta < 0.$$

Let ε be any positive number such that $\varepsilon M < \delta$, namely $-\delta + \varepsilon M < 0$. Then, for every $\mathbf{x} \in \mathcal{E}$, we have:

$$\begin{aligned}
&\max_{i=1,\dots,p} \mathbf{x}^T(A_i^T \tilde{P} A_i - \tilde{P})\mathbf{x} = \\
&= \max_{i=1,\dots,p} [\mathbf{x}^T(A_i^T P A_i - P)\mathbf{x} + \varepsilon(\mathbf{x}^T(A_i^T A_i - I_n)\mathbf{x})] \\
&\leq \max_{\mathbf{x}\in\mathcal{E}} \max_{i=1,\dots,p} [\mathbf{x}^T(A_i^T P A_i - P)\mathbf{x}] + \varepsilon \cdot \max_{\mathbf{x}\in\mathcal{E}} \max_{i=1,\dots,p} [|\mathbf{x}^T(A_i^T A_i - I_n)\mathbf{x}|] = \\
&= -\delta + \varepsilon M < 0.
\end{aligned}$$

By the homogeneity of $V(\mathbf{x})$, the result holds for every $\mathbf{x} > 0$.

(iii) $\Rightarrow$ (iv): The proof is obvious. $\square$

Remark 2.2. While the existence of a Common Linear CLF implies the existence of a Common Quadratic Positive Definite CLF, the converse is not true. Consider the pair of positive Schur matrices:

$$A_1 = \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{30} \end{bmatrix} \quad A_2 = \begin{bmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{3} \end{bmatrix}.$$

It is easy to see that the matrix:

$$A_{(1,2)} = \begin{bmatrix} A_1^{(1)} & A_2^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -\frac{2}{3}$. So, it is not a Schur matrix and, by Corrolary 2.1, a Common Linear CLF for A_1 and A_2 does not exist. However, it is a matter of simple calculation to show that the matrix:

$$\tilde{P} = \begin{bmatrix} 1 & \frac{4}{5} \\ \frac{4}{5} & 2 \end{bmatrix} = \tilde{P}^T \succ 0$$

makes $V(\mathbf{x}) = \mathbf{x}^T \tilde{P} \mathbf{x}$ a Common Quadratic Positive Definite CLF for A_1 and A_2.

2.3 Switched Linear Copositive Lyapunov Functions

As already pointed out in our earlier discussion, stability criteria obtained by means of common CLF (either linear, quadratic or quadratic positive definite) are overconservative. An alternative line of research, only partially explored to date, aims at reducing such conservatism by resorting to switched CLF. This method has been first applied to discrete-time general switched systems and then extended to discrete-time positive switched systems. In this regard, Liu provides [26] a necessary and sufficient condition for the existence of a Switched Linear CLF and formulates such condition both as a set of Linear Programming[2]

[2] A Linear Programming problem may be defined as the problem of maximazing or minimazing a linear function subject to linear constraint (which may be either equalities or inequalities). The interest in an LP formulation lies on the fact that there exist efficient algorithms to solve these kind of problems.

(LP) problems and LMI problems. The result is provided by Theorem 2.4 below, but first let us recall that a Switched Linear CLF takes the following form:

$$V(t, \mathbf{x}(t)) = \left(\sum_{i=1}^{p} \eta_i(t) \mathbf{v}_i \right)^T \mathbf{x}(t), \tag{2.8}$$

where $\eta_i(t) = \begin{cases} 1, & \text{if } \sigma(t) = i \\ 0, & \text{otherwise.} \end{cases}$

Theorem 2.4. Let $A_1, \dots, A_p \in \mathbb{R}^{n \times n}$ be positive, Schur matrices. Then, the following statements are equivalent:

(i) There exists a Switched Linear CLF of the form (2.8) for $A_1, \dots, A_p$;

(ii) (LP problem) There exist p vectors $\mathbf{v}_i \gg 0$, $i \in [1, p]$, such that:

$$\boldsymbol{\varphi}_{ij} = A_i^T \mathbf{v}_j - \mathbf{v}_i \ll 0, \quad \forall (i, j) \in [1, p] \times [1, p];$$

(iii) (LMI problem) There exist p diagonal positive definite matrices $P_i = \operatorname{diag}(p_{i1}, p_{i2}, \dots, p_{in})$, $i \in [1, p]$, such that:

$$\Theta_{ij} = \operatorname{diag}(\theta_{ij1}, \dots, \theta_{ijn}) \prec 0, \quad \forall (i, j) \in [1, p] \times [1, p],$$

where $\theta_{ijk} = (A_i^{(k)})^T \mathbf{p}_j - p_{ik}$ with $\mathbf{p}_j = \begin{bmatrix} p_{j1} & p_{j2} & \dots & p_{jn} \end{bmatrix}^T$.

It is straightforward to see that results obtained by Theorem 2.4 are less conservative than those derived by means of Common Linear CLFs. Indeed, if there exists a Common Linear CLF $V(\mathbf{x}) = \mathbf{v}_1^T \mathbf{x}$, then a Switched Linear CLF of the form (2.8) can be found too, as we only need to choose $\eta_1(t) \equiv 1$ and $\eta_i(t) \equiv 0$ for $i \neq 1$ (equivalently, notice that p identical vectors $\mathbf{v}_1 = \mathbf{v}_2 = \dots = \mathbf{v}_p$ satisfy condition (ii) in Theorem 2.4). However, it is also clear that such reduction of conservatism is at the price of increasing computational effort.

To summarize, throughout this chapter we have focused on the search for conditions ensuring asymptotic system stability and a number of interesting results have been presented. First of all, investigation of the existence of a Common Linear CLF has led to deeper insights into the properties the subsystems family must be endowed with. In particular, Theorem 2.1 provides a characterization referring to the Hurwitz property of a certain finite set of matrices, while Theorem 2.2 provides conditions involving the convex hulls of certain vector families and their intersection with the positive orthant. Secondly, Theorem 2.3 relates the existence of a Common Linear CLF to that of a Common Quadratic CLF and a Common Quadratic Positive Definite CLF. Finally, in Theorem 2.4, by resorting to Switched Linear CLFs, less conservative conditions, although more computationally demanding, have been provided.

Chapter 3

State-feedback Stabilizability

We have seen in the previous chapter conditions ensuring a good asymptotic behavior to the system trajectories, independently of the switching law. On the other hand, there are systems that are not stable under arbitrary switching (this is the case, for instance, when modeling viral mutation and escape in patients affected by HIV [11],[10], as we will see in Chapter 5). In addition, unstable subsystems are commonly encountered in many engineering processes, because of disturbances, unmodelled dynamics or possible faults. For such systems we are interested in determining, if possible, switching strategies that ensure the convergence to zero of the state trajectories. Despite the fact that both practical applications and theoretical reasons make the investigation of stabilizability property appealing to an increasing number of researchers, it is still a topic only partially explored and it offers a wide range of interesting open problems.

3.1 Definitions and preliminaries

We have so far referred to stabilizability in an intuitive way, as the possibility of making the state evolutions converge to zero by means of suitable switching signals. A more rigorous formulation of the concept is provided by the following definitions.

Definition 3.1. The DPSS (1.1) is *stabilizable* if for every positive initial state $\mathbf{x}(0)$ there exists a switching sequence $\sigma\colon \mathbb{Z}_+ \to [1,p]$ such that the state trajectory $\mathbf{x}(t)$, $t \in \mathbb{Z}_+$, generated by the DPSS starting from $\mathbf{x}(0)$ and corresponding to σ, asymptotically converges to zero.

Definition 3.2. The DPSS (1.1) is *consistently stabilizable* if there exists a switching sequence $\sigma\colon \mathbb{Z}_+ \to [1,p]$ such that, for every positive initial state $\mathbf{x}(0)$, the state trajectory $\mathbf{x}(t)$, $t \in \mathbb{Z}_+$, generated by the DPSS starting from $\mathbf{x}(0)$ and corresponding to σ, converges to zero.

It is worthwhile to underline that, while in Definition 3.1 the choice of the switching sequence σ may depend on the initial state $\mathbf{x}(0)$, in Definition 3.2 the stabilizing sequence σ is required to drive to zero the state trajectory independently of the initial state.

It is clear that consistent stabilizability implies stabilizability, while the natural question arises whether the converse is true. In that regard, [8] provides an effective response along with an additional characterization of stabilizability.

Proposition 3.1. Given a DPSS (1.1), the following statements are equivalent:

(i) The system is stabilizable;

(ii) The system is consistently stabilizable;

(iii) There exist $N > 0$ and indices $i_0, i_1, \dots, i_{N-1} \in [1, p]$, such that the matrix product $A_{i_{N-1}} A_{i_{N-2}} \dots A_{i_1} A_{i_0}$ is a positive, Schur matrix;

(iv) There exists a periodic switching sequence that leads to zero every positive initial state.

Proof. (i) $\Rightarrow$ (ii): Consider any switching sequence σ that drives to zero the initial state $\hat{\mathbf{x}}(0) = \mathbf{1}_n$, where $\mathbf{1}_n$ is the n-dimensional vector with all entries equal to 1. We prove that such switching sequence drives to zero every other positive state $\mathbf{x}(0)$. Indeed, choose a positive number M such that $0 < \mathbf{x}(0) \leq M\hat{\mathbf{x}}(0)$ and let $\mathbf{x}(t)$ and $\hat{\mathbf{x}}(t)$, $t \in \mathbb{Z}_+$, be the state evolutions originated from $\mathbf{x}(0)$ and $\hat{\mathbf{x}}(0)$, respectively, under the switching sequence σ. By the positivity assumption on the matrices A_i's, we have:

$$0 \leq \mathbf{x}(t) \leq M\hat{\mathbf{x}}(t) \quad \forall t \in \mathbb{Z}_+,$$

which ensures that $\mathbf{x}(t)$ goes to zero as $t \to +\infty$. So, the system is consistently stabilizable.

(ii) $\Rightarrow$ (iii): Let σ be the switching sequence that makes the state evolution go to zero, independently of the initial state. Set $\mathbf{x}(0) = \mathbf{1}_n$ and $\varepsilon \in (0, 1)$. Then, a positive integer N can be found such that:

$$\mathbf{x}(N) = A_{\sigma(N-1)} \dots A_{\sigma(1)} A_{\sigma(0)} \mathbf{1}_n \ll \varepsilon \mathbf{1}_n.$$

This ensures, by standard properties of positive matrices (see Proposition A.1 in the Appendix[1]), that the spectral radius of the positive matrix $A_{\sigma(N-1)} \dots A_{\sigma(1)} A_{\sigma(0)}$ is smaller than $\varepsilon < 1$, and hence the matrix is Schur. So, (iii) holds for $i_k = \sigma(k)$, $k \in [0, N-1]$.

(iii) $\Rightarrow$ (iv): If $A := A_{i_{N-1}} A_{i_{N-2}} \dots A_{i_1} A_{i_0}$ is a positive, Schur matrix, then A^k converges to zero as k goes to infinity. Consequently, the switching sequence $\sigma(t) = i_{(t \bmod N)}$ drives to zero the state evolution corresponding to every positive initial state.

(iv) $\Rightarrow$ (i): The proof is obvious. □

The above proposition states that, when dealing with DPSSs, consistent stabilizability and stabilizability are equivalent properties, and they are both

[1] If the matrices $A_1, A_2, \dots, A_p$ are not irreducible, it is always possible to choose $\delta > 0$ such that $\tilde{A}_i = A_i + \delta \mathbf{1}_n \mathbf{1}_n^T$ is irreducible for all $i = 1, \dots, p$. Then, the following inequalities hold:

$$A_{\sigma(N-1)} \dots A_{\sigma(1)} A_{\sigma(0)} \mathbf{1}_n \ll \tilde{A}_{\sigma(N-1)} \dots \tilde{A}_{\sigma(1)} \tilde{A}_{\sigma(0)} \mathbf{1}_n \ll \varepsilon \mathbf{1}_n,$$

which implies that the spectral radius of the positive matrix $A_{\sigma(N-1)} \dots A_{\sigma(1)} A_{\sigma(0)}$ is smaller than the spectral radius of the matrix $\tilde{A}_{\sigma(N-1)} \dots \tilde{A}_{\sigma(1)} \tilde{A}_{\sigma(0)}$ (whose spectral radius, in turn, is smaller than $\varepsilon < 1$ by Proposition A.1 in the Appendix).

equivalent to the possibility of stabilizing the system by means of a periodic switching sequence, independently of the positive initial state. It is worth noticing that in the general case, i.e. when there is no positivity assumption, discrete-time switched systems can be found (see [36], Chapter 3) that are stabilizable, but not consistently stabilizable (indeed, in the above proof (i) $\Rightarrow$ (ii) we made unavoidable use of the positivity assumption).

Clearly, the stabilization problem is a non-trivial one only if all matrices A_i's are not Schur (otherwise, if $i \in [1,p]$ is the index of some asymptotically stable subsystem, the switching signal $\sigma(t) \equiv i$ drives to zero every positive initial state). So, throughout this chapter, we will steadily make this assumption.

Apart from formulating stabilizability criteria that characterize the existence of a stabilizing switching law, a more important issue is to explicitly calculate such a stabilizing law. In that regard, notice that the stabilizing switching strategy is most likely not to be unique: by choosing different switching rules, one may obtain different trajectories, all of them convergent to zero, originating from the same initial state. However, in what follows, we will not be concerned with a comparison between such convergent trajectories, namely between their rates of convergence, but only with the existence and the design of a stabilizing switching strategy. In particular, the focus will be put on the search for a stabilizing switching sequence whose value at time t depends on the specific value of the state $\mathbf{x}(t)$, thus representing a state-feedback switching sequence of the form $\sigma(\mathbf{x}(t)) = u(\mathbf{x}(t))$. To this aim, it is assumed that the full state vector $\mathbf{x}(t)$ is available for feedback for all $t \in \mathbb{Z}_+$.

In the following sections, in order to determine the function $u(\cdot)\colon \mathbb{R}^n \to [1,p]$ which stabilizes the system, we will first resort to piecewise CLFs, and then we will move on to CLFs such that, for every $\mathbf{x} > 0$, $\Delta V_i(\mathbf{x}) < 0$ for some $i \in [1,p]$.

3.2 Piecewise Copositive Lyapunov Function approach

In this section we aim at designing a stabilizing switching rule by means of a piecewise Lyapunov Function, which takes the form of the minimum of p individual Lyapunov functions. In particular, we will consider Piecewise Linear CLFs of the form:

$$V(\mathbf{x}(t)) = \min_{i=1,\dots,p} \mathbf{v}_i^T \mathbf{x}(t), \tag{3.1}$$

or Piecewise Quadratic Lyapunov Functions of the form:

$$V(\mathbf{x}(t)) = \min_{i=1,\dots,p} \mathbf{x}^T(t) P_i \mathbf{x}(t). \tag{3.2}$$

This method has been first applied in the context of general switched systems and then extended to positive switched systems (both in the continuous and in the discrete-time case). For this reason, we begin our discussion with a result derived in the general case, that clearly holds true also for positive switched systems. In [9] Geromel and Colaneri resort to Piecewise Quadratic Lyapunov Functions to propose a state-feedback strategy for stabilization, designed from the solution of a set of Lyapunov-Metzler inequalities.

Before stating such result, let us recall that a *simplex* in $\mathbb{R}^p$ is the set of vectors:

$$\Lambda := \left\{ \boldsymbol{\lambda} \in \mathbb{R}^p : \sum_{i=1}^{p} \lambda_i = 1,\ \lambda_i \geq 0 \right\}.$$

We also need to introduce a class of Metzler matrices, denoted by $\mathscr{M}$, which is constituted by all Metzler matrices $\mathbb{M} \in \mathbb{R}^{p \times p}$ with elements μ_{ij} such that:

$$\mu_{ij} \geq 0,\ \sum_{i=1}^{p} \mu_{ij} = 1,\ \forall i, j.$$

Proposition 3.2. Let $A_1, \ldots, A_p \in \mathbb{R}^{n \times n}$ be the subsystems' state matrices of a discrete-time switched system. Assume that there exist $\mathbb{M} \in \mathscr{M}$ and a set of p positive definite matrices $P_1, \ldots, P_p$ satisfying the Lyapunov-Metzler inequalities:

$$A_i^T \left(\sum_{j=1}^{p} \mu_{ji} P_j \right) A_i - P_i \prec 0, \quad \forall i \in [1, p]. \tag{3.3}$$

Then, the state-feedback switching rule:

$$\sigma(\mathbf{x}(t)) := \underset{i=1,\ldots,p}{\operatorname{argmin}}\ \mathbf{x}^T(t) P_i \mathbf{x}(t) \tag{3.4}$$

stabilizes the system, i.e. it makes the state evolution goes to zero for every initial state.

Proof. Consider the Piecewise Quadratic Lyapunov Function (3.2) with matrices P_i satisfying (3.3) and assume that, at an arbitrary instant $t \in \mathbb{Z}_+$, the state switching control is given by $\sigma(\mathbf{x}(t)) = u(\mathbf{x}(t)) = i$ for some $i \in [1, p]$. Hence, we have $V(\mathbf{x}(t)) = \mathbf{x}^T(t) P_i \mathbf{x}(t)$ and, from the system dynamic equation, we get:

$$\begin{aligned} V(\mathbf{x}(t+1)) &= \min_{j=1,\ldots,p} \mathbf{x}^T(t) A_i^T P_j A_i \mathbf{x}(t) \\ &= \min_{\boldsymbol{\lambda} \in \Lambda} \mathbf{x}^T(t) A_i^T \left(\sum_{j=1}^{p} \lambda_j P_j \right) A_i \mathbf{x}(t) \\ &\leq \mathbf{x}^T(t) A_i^T \left(\sum_{j=1}^{p} \mu_{ji} P_j \right) A_i \mathbf{x}(t), \end{aligned} \tag{3.5}$$

where the inequality holds owing to the fact that each column of $\mathbb{M}$ belongs to Λ. Now, from the Lyapunov-Metzler inequalities (3.3), it follows that for every $\mathbf{x}(t) \neq 0$:

$$\mathbf{x}^T(t) \left[A_i^T \left(\sum_{j=1}^{p} \mu_{ji} P_j \right) A_i - P_i \right] \mathbf{x}(t) < 0,$$

or, equivalently:

$$\mathbf{x}^T(t) A_i^T \left(\sum_{j=1}^{p} \mu_{ji} P_j \right) A_i \mathbf{x}(t) < \mathbf{x}^T(t) P_i \mathbf{x}(t). \tag{3.6}$$

Hence, from (3.5) and (3.6), it follows:

$$V(\mathbf{x}(t+1)) < \mathbf{x}^T(t) P_i \mathbf{x}(t) = V(\mathbf{x}(t)).$$

So, taking into account that the Lyapunov function $V(\mathbf{x}(t))$ is radially unbounded, we conclude that $\mathbf{x}(t)$ converges to zero for every initial state. $\square$

It is worth noticing that the numerical determination of a solution, if any, of the Lyapunov-Metzler inequalities (3.3) with respect to the variables $\mathbb{M}$ and $P_1, \dots, P_p$ is not a simple task due to its non-linear nature, and this is why Geromel and Colaneri provide in [9] an alternative, although certainly more conservative, condition expressed by means of LMIs. This result is here omitted, as a very similar argument will be presented, in the context of DPSSs, towards the end of the section.

As mentioned above, the piecewise Lyapunov function method has been applied also to DPSSs. However, this line of research has been only partially explored and, to date, only Piecewise Linear CLFs have been considered. In this regard, a recent work by Hernandez-Vargas and coauthors should be mentioned: [11] provides a sufficient condition for the existence of a Piecewise Linear CLF, which, in turn, allows to design a stabilizing switching sequence. The result is presented in Theorem 3.1 below, but first let us introduce a new class of Metzler matrices. We shall denote by $\mathscr{N}$ the set of all Metzler matrices $\mathbb{N} \in \mathbb{R}^{p\times p}$ with elements ν_{ij}, such that:

$$\nu_{ij} \geq 0, \forall i \neq j, \ \sum_{i=1}^{p} \nu_{ij} = 0, \forall j.$$

Theorem 3.1. Let $A_1, \dots, A_p \in \mathbb{R}_+^{n\times n}$ be the subsystems' state matrices of a DPSS (1.1). Assume that there exist $\mathbb{N} \in \mathscr{N}$ and a set of p positive vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$, $\mathbf{v}_i \in \mathbb{R}_+^n$, satisfying the copositive Lyapunov inequalities:

$$(A_i - I_n)^T \mathbf{v}_i + \sum_{j=1}^{p} \nu_{ji} \mathbf{v}_j < 0, \quad \forall i \in [1, p]. \tag{3.7}$$

Then, the state-feedback switching rule:

$$\sigma(\mathbf{x}(t)) := \operatorname*{argmin}_{i=1,\dots,p} \mathbf{v}_i^T \mathbf{x}(t) \tag{3.8}$$

stabilizes the DPSS, i.e. it makes the state evolution goes to zero for every positive initial state.

Proof. Consider the Piecewise Linear CLF (3.1) with vectors $\mathbf{v}_i$ satisfying (3.7). We prove that $V(\mathbf{x}(t))$ is decreasing, namely its difference $\Delta V(\mathbf{x}(t))$ is negative, along the system trajectory corresponding to the state-switching control (3.8). Indeed, we have:

$$\begin{aligned}
\Delta V(\mathbf{x}(t)) :&= V(\mathbf{x}(t+1)) - V(\mathbf{x}(t)) \\
&= \min_{j=1,\dots,p} \{\mathbf{v}_j^T \mathbf{x}(t+1)\} - \min_{j=1,\dots,p} \{\mathbf{v}_j^T \mathbf{x}(t)\} \\
&= \min_{j=1,\dots,p} \{\mathbf{v}_j^T A_{\sigma(t)} \mathbf{x}(t)\} - \min_{j=1,\dots,p} \{\mathbf{v}_j^T \mathbf{x}(t)\} \\
&= \min_{j=1,\dots,p} \{\mathbf{v}_j^T A_{\sigma(t)} \mathbf{x}(t)\} - \mathbf{v}_{\sigma(t)}^T \mathbf{x}(t),
\end{aligned}$$

and, by definition of $\sigma(\mathbf{x}(t))$, the following inequality holds:

$$\begin{aligned}\Delta V(\mathbf{x}(t)) &\leq \mathbf{v}_{\sigma(t)}^T A_{\sigma(t)} \mathbf{x}(t) - \mathbf{v}_{\sigma(t)}^T \mathbf{x}(t) \\ &= \mathbf{v}_{\sigma(t)}^T (A_{\sigma(t)} - I_n) \mathbf{x}(t).\end{aligned} \tag{3.9}$$

Now, notice that the Lyapunov inequality (3.7) for $i = \sigma(t)$ can be rewritten as:

$$(A_{\sigma(t)} - I_n)^T \mathbf{v}_{\sigma(t)} < -\sum_{j=1}^{p} \nu_{j\sigma(t)} \mathbf{v}_j,$$

or, equivalently, as:

$$\mathbf{v}_{\sigma(t)}^T (A_{\sigma(t)} - I_n) < -\sum_{j=1}^{p} \nu_{j\sigma(t)} \mathbf{v}_j^T. \tag{3.10}$$

From (3.9) and (3.10), with $\mathbf{x}(t) \neq 0$, it follows:

$$\begin{aligned}\Delta V(\mathbf{x}(t)) &< -\sum_{j=1}^{p} \nu_{j\sigma(t)} \mathbf{v}_j^T \mathbf{x}(t) \\ &= -\nu_{\sigma(t)\sigma(t)} \mathbf{v}_{\sigma(t)}^T \mathbf{x}(t) - \sum_{\substack{j=1 \\ j\neq\sigma(t)}}^{p} \nu_{j\sigma(t)} \mathbf{v}_j^T \mathbf{x}(t) \\ &\leq -\nu_{\sigma(t)\sigma(t)} \mathbf{v}_{\sigma(t)}^T \mathbf{x}(t) - \sum_{\substack{j=1 \\ j\neq\sigma(t)}}^{p} \nu_{j\sigma(t)} \mathbf{v}_{\sigma(t)}^T \mathbf{x}(t) \\ &= -\left(\sum_{j=1}^{p} \nu_{j\sigma(t)}\right) \mathbf{v}_{\sigma(t)}^T \mathbf{x}(t) = 0.\end{aligned}$$

Thus, $V(\mathbf{x}(t))$ converges to zero, and $\mathbf{x}(t)$ converges to zero in turn. □

The previous theorem provides a sufficient condition for stabilizability along with a stabilizing switching rule designed from the solution of a set of copositive Lyapunov inequalities (notice that such inequalities are not LMIs since the unknown parameters ν_{ji} multiply the unknown vectors $\mathbf{v}_j$). An interesting application of this result will be presented in Chapter 6 in the context of HIV treatment modeling.

More recently, Tong and coauthors [37] resorted to the already introduced class of matrices denoted by $\mathscr{M}$ in order to provide an alternative stabilization strategy: again, the stabilization condition is expressed by means of a set of matrix inequalities, whose solution, if any, defines a Piecewice Linear CLF, which, in turn, allows to determine a stabilizing switching law.

Theorem 3.2. Let $A_1, \dots, A_p \in \mathbb{R}_+^{n\times n}$ be the subsystems' state matrices of a DPSS (1.1). Assume that there exist $\mathbb{M} \in \mathscr{M}$ and a set of p positive vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$, $\mathbf{v}_i \in \mathbb{R}_+^n$, satisfying the copositive Lyapunov-Metzler inequalities:

$$A_i^T \left(\sum_{j=1}^{p} \mu_{ji} \mathbf{v}_j\right) - \mathbf{v}_i < 0, \quad \forall i \in [1, p]. \tag{3.11}$$

Then, the state-feedback switching rule:

$$\sigma(\mathbf{x}(t)) := \underset{i=1,\ldots,p}{\operatorname{argmin}}\ \mathbf{v}_i^T \mathbf{x}(t) \tag{3.12}$$

stabilizes the DPSS, i.e. it makes the state evolution goes to zero for every positive initial state.

Proof. The proof follows the same line as the proof of Proposition 3.2, except that the Lyapunov function we now consider is the Piecewise Linear CLF (3.1). Suppose that at an arbitrary instant $t \in \mathbb{Z}_+$, the state switching control is given by $\sigma(\mathbf{x}(t)) = u(\mathbf{x}(t)) = i$ for some $i \in [1, p]$. Hence, $V(\mathbf{x}(t)) = \mathbf{v}_i^T \mathbf{x}(t)$ and, from the system dynamic equation, we have:

$$\begin{aligned} V(\mathbf{x}(t+1)) &= \min_{j=1,\ldots,p} \mathbf{v}_j^T A_i \mathbf{x}(t) \\ &= \min_{\boldsymbol{\lambda} \in \Lambda} \left(\sum_{j=1}^{p} \lambda_j \mathbf{v}_j^T \right) A_i \mathbf{x}(t) \\ &\leq \left(\sum_{j=1}^{p} \mu_{ji} \mathbf{v}_j^T \right) A_i \mathbf{x}(t), \end{aligned} \tag{3.13}$$

where the inequality holds owing to the fact that each column of $\mathbb{M}$ belongs to Λ. Then, from (3.11) and (3.13), it follows:

$$V(\mathbf{x}(t+1)) < \mathbf{v}_i^T \mathbf{x}(t) = V(\mathbf{x}(t)).$$

Recalling that the Lyapunov function $V(\mathbf{x}(t))$ is radially unbounded, we conclude that $\mathbf{x}(t)$ converges to zero for every positive initial state. □

Again, the Lyapunov-Metzler inequalities (3.11) are not linear because of the products of variables μ_{ji} and $\mathbf{v}_j$. In order to obtain an efficient numerical solution, Tong and coauthors focus on particular matrices belonging to $\mathscr{M}$, namely matrices A_i, $i = 1, \ldots, p$, whose diagonal entries are identical, by this meaning that for each of them $[A_i]_{jj} = [A_i]_{kk}$ for all j, k. This technique, which leads to a stabilizability condition expressed by LMIs, is analogous to the one pursued by Geromel and Colaneri in [9] when dealing with general switched systems and piecewise quadratic Lyapunov functions. Formally, the result can be summarized by the following theorem.

Theorem 3.3. Let $A_1, \ldots, A_p \in \mathbb{R}_+^{n \times n}$ be the subsystems' state matrices of a DPSS (1.1). Let $\mathbf{q}_1, \ldots, \mathbf{q}_p$, $\mathbf{q}_i \in \mathbb{R}_+^n$, be a given set of p positive vectors and assume that there exist a scalar $0 \leq \eta \leq 1$ and a set of p positive vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$, $\mathbf{v}_i \in \mathbb{R}_+^n$, satisfying the copositive Lyapunov-Metzler inequalities:

$$A_i^T \left[\eta \mathbf{v}_i + (1-\eta) \mathbf{v}_j \right] - \mathbf{v}_i + \mathbf{q}_i < 0, \quad \forall i \neq j \in [1, p]. \tag{3.14}$$

Then, the state-feedback switching rule:

$$\sigma(\mathbf{x}(t)) := \underset{i=1,\ldots,p}{\operatorname{argmin}}\ \mathbf{v}_i^T \mathbf{x}(t) \tag{3.15}$$

stabilizes the DPSS, i.e. it makes the state evolution goes to zero for every positive initial state.

Proof. Choose $\mathbb{M} \in \mathscr{M}$, such that $\mu_{ii} = \eta$ for all $i \in [1, p]$. Then the remaining elements satisfy:

$$\sum_{\substack{j=1 \\ j \neq i}}^{p} \mu_{ji} = 1 - \eta, \quad \forall i, j \in [1, p].$$

Now, by multiplying (3.14) by μ_{ji} and, thereafter, by summing up for all $j \neq i \in [1, p]$, we obtain:

$$\eta A_i^T \sum_{\substack{j=1 \\ j \neq i}}^{p} \mu_{ji} \mathbf{v}_i + (1 - \eta) A_i^T \sum_{\substack{j=1 \\ j \neq i}}^{p} \mu_{ji} \mathbf{v}_j < (1 - \eta)(\mathbf{v}_i - \mathbf{q}_i). \tag{3.16}$$

By multiplying both sides of (3.16) by $(1 - \eta)^{-1}$ and recalling that $\eta = \mu_{ii}$ for all $i \in [1, p]$, we get:

$$A_i^T \sum_{j=1}^{p} \mu_{ji} \mathbf{v}_j < \mathbf{v}_i - \mathbf{q}_i,$$

or, equivalently:

$$A_i^T \sum_{j=1}^{p} \mu_{ji} \mathbf{v}_j - \mathbf{v}_i + \mathbf{q}_i < 0,$$

which implies that also the Lyapunov-Metzler inequalities (3.11) hold true. Hence, by Theorem 3.2, the DPSS is globally asymptotically stable under the switching law (3.15). $\square$

It is clear that the Lyapunov-Metzler inequalities (3.14) expressed by means of LMIs are simpler to solve, but provide a more conservative condition, as we restrict ourselves to a subclass of all matrices belonging to $\mathscr{M}$.

3.3 An alternative Copositive Lyapunov Function approach

Another strategy to stabilize a DPSS (1.1) is proposed in [8]: Fornasini and Valcher resort to CLFs such that for every $\mathbf{x} > 0$ there exists $i \in [1, p]$ such that $\Delta V_i(\mathbf{x})$ is negative, namely:

$$\min_{i=1,\dots,p} \Delta V_i(\mathbf{x}) = \min_{i=1,\dots,p} V(A_i \mathbf{x}) - V(\mathbf{x}) < 0, \quad \forall \mathbf{x} > 0, \tag{3.17}$$

where $V(\mathbf{x})$ is a CLF either linear, quadratic or quadratic positive definite. Notice that, differently from the piecewise CLF method, we do not search for p vectors $\mathbf{v}_i$ (or p matrices P_i) defining p individual CLFs, but a unique CLF (and hence a unique vector $\mathbf{v}$ or a unique matrix P) satisfying (3.17) is required.

In the following, we first investigate, and mutually relate, conditions for the existence of a CLF satisfying (3.17) and, subsequently, we will prove that, when any such function $V(\mathbf{x})$ is available, a stabilizing switching law, based on the values taken by the various $\Delta V_i(\mathbf{x})$'s, can be found.

Theorem 3.4. Let $A_1, \dots, A_p \in \mathbb{R}_+^{n \times n}$ be positive matrices. Condition:

(i*) $\exists P = P^T \succ 0$ and $\alpha_1, \dots, \alpha_p \geq 0$, with $\sum_{i=1}^p \alpha_i = 1$, such that $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$ satisfies:

$$\sum_{i=1}^p \alpha_i \Delta V_i(\mathbf{x}) = \sum_{i=1}^p \alpha_i \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < 0 \quad \forall \mathbf{x} > 0,$$

implies any of the following equivalent facts:

(i) $\exists \mathbf{v} \gg 0$ such that $V(\mathbf{x}) = \mathbf{v}^T \mathbf{x}$ is a Linear CLF that satisfies:

$$\min_{i=1,\dots,p} \Delta V_i(\mathbf{x}) = \min_{i=1,\dots,p} \mathbf{v}^T (A_i - I_n)\mathbf{x} < 0 \quad \forall \mathbf{x} > 0;$$

(ii) $\exists P = P^T$ of rank 1 such that $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$ is a Quadratic CLF that satisfies:

$$\min_{i=1,\dots,p} \Delta V_i(\mathbf{x}) = \min_{i=1,\dots,p} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < 0 \quad \forall \mathbf{x} > 0.$$

If (i)-(ii) hold, then the following condition holds:

(iii) $\exists \tilde{P} = \tilde{P}^T \succ 0$ such that $V(\mathbf{x}) = \mathbf{x}^T \tilde{P} \mathbf{x}$ is a Quadratic Positive Definite CLF that satisfies:

$$\min_{i=1,\dots,p} \Delta V_i(\mathbf{x}) = \min_{i=1,\dots,p} \mathbf{x}^T (A_i^T \tilde{P} A_i - \tilde{P})\mathbf{x} < 0 \quad \forall \mathbf{x} > 0.$$

If (iii) holds, then:

(iv) $\exists P = P^T$ such that $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$ is a Quadratic CLF that satisfies:

$$\min_{i=1,\dots,p} \Delta V_i(\mathbf{x}) = \min_{i=1,\dots,p} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < 0 \quad \forall \mathbf{x} > 0.$$

Proof. (i*) $\Rightarrow$ (i): Let Q_i denote the matrix:

$$Q_i := \begin{bmatrix} A_i^T P A_i & A_i^T P \\ P A_i & P \end{bmatrix}.$$

It is easily seen that Q_i is positive semidefinite for every $i \in [1, p]$, indeed:

$$\mathbf{x}^T Q_i \mathbf{x} = \mathbf{x}^T \begin{bmatrix} A_i^T \\ I_n \end{bmatrix} P \begin{bmatrix} A_i & I_n \end{bmatrix} \mathbf{x} \geq 0 \quad \forall \mathbf{x} \neq 0.$$

Consequently, also the matrix $\sum_{i=1}^p \alpha_i Q_i$ is positive semidefinite:

$$\sum_{i=1}^p \alpha_i Q_i = \begin{bmatrix} \left(\sum_{i=1}^p \alpha_i A_i^T P A_i\right) & \left(\sum_{i=1}^p \alpha_i A_i^T\right) P \\ P\left(\sum_{i=1}^p \alpha_i A_i\right) & P \end{bmatrix} \succeq 0.$$

By the Schur complement's formula[2], this implies that for every $\mathbf{x}$, and hence, in particular, for every $\mathbf{x} > 0$, we have:

$$\mathbf{x}^T \left[\left(\sum_{i=1}^p \alpha_i A_i^T P A_i \right) - \left(\sum_{i=1}^p \alpha_i A_i^T \right) P \left(\sum_{i=1}^p \alpha_i A_i \right) \right] \mathbf{x} \geq 0,$$

[2]Let M be a symmetric matrix block given by:

$$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

with matrices A, B, C having suitable dimensions. If C is positive definite, then the following condition holds: M is positive semidefinite if and only if the Schur complement of the block C is positive semidefinite, namely $M \succeq 0 \Leftrightarrow A - BC^{-1}B^T \succeq 0$.

or, equivalently:

$$\mathbf{x}^T \left(\sum_{i=1}^{p} \alpha_i A_i^T P A_i \right) \mathbf{x} \geq \mathbf{x}^T \left(\sum_{i=1}^{p} \alpha_i A_i^T \right) P \left(\sum_{i=1}^{p} \alpha_i A_i \right) \mathbf{x}.$$

Now, by subtracting $\mathbf{x}^T P \mathbf{x}$ on both sides, we obtain:

$$\mathbf{x}^T \left[\sum_{i=1}^{p} \alpha_i (A_i^T P A_i - P) \right] \mathbf{x} \geq \mathbf{x}^T \left[\left(\sum_{i=1}^{p} \alpha_i A_i^T \right) P \left(\sum_{i=1}^{p} \alpha_i A_i \right) - P \right] \mathbf{x}. \quad (3.18)$$

Since, by assumption (i*), the left hand-side in (3.18) is negative for every $\mathbf{x} > 0$, so is the right hand-side:

$$\mathbf{x}^T \left[\left(\sum_{i=1}^{p} \alpha_i A_i^T \right) P \left(\sum_{i=1}^{p} \alpha_i A_i \right) - P \right] \mathbf{x} < 0, \quad \forall \mathbf{x} > 0. \quad (3.19)$$

Let A_α denote the positive matrix $A_\alpha := \sum_{i=1}^{p} \alpha_i A_i$. It follows from (3.19) that $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$ is a quadratic copositive function such that $V(A_\alpha \mathbf{x}) - V(\mathbf{x}) < 0$ for every $\mathbf{x} > 0$ and, hence, A_α is asymptotically stable, which amounts to saying that A_α is a Schur matrix. Now, notice that $\sum_{i=1}^{p} \alpha_i I_n = I_n$ and recall that A_α is a positive, Schur matrix, if and only if:

$$\bar{A}_\alpha := A_\alpha - I_n = \sum_{i=1}^{p} \alpha_i (A_i - I_n)$$

is a Metzler, Hurwitz matrix. This implies, by standard properties of Metzler matrices, that there exists a vector $\mathbf{v} \gg 0$ such that $\mathbf{v}^T \bar{A}_\alpha \ll 0$. Hence, for every positive vector $\mathbf{x}$, we have:

$$\mathbf{v}^T \bar{A}_\alpha \mathbf{x} = \sum_{i=1}^{p} \alpha_i \left[\mathbf{v}^T (A_i - I_n) \mathbf{x} \right] < 0,$$

which amounts to saying that $\min_{i=1,\dots,p} \mathbf{v}^T (A_i - I_n) \mathbf{x} < 0$.

(i) $\Leftrightarrow$ (ii): The proof follows the same lines as the proof of the analogous conditions (i) $\Leftrightarrow$ (ii) in Theorem 2.3.

(ii) $\Rightarrow$ (iii): The reasoning is very similar to the one used in the proof of (ii) $\Rightarrow$ (iii) in Theorem 2.3, except that the two continuous functions we need now are:

$$f(\mathbf{x}) := \max_{i=1,\dots,p} |\mathbf{x}^T (A_i^T A_i - I_n) \mathbf{x}|,$$
$$g(\mathbf{x}) := \max_{i=1,\dots,p} \mathbf{x}^T (A_i^T P A_i - P) \mathbf{x}.$$

(iii) $\Rightarrow$ (iv): The proof is obvious. □

Remark 3.1. A close examination of the proof (i*) $\Rightarrow$ (i) in Theorem 3.4 shows that a necessary and sufficient condition for the existence of a Linear CLF satisfying (3.17) is that there exists a Schur convex combination of the positive system matrices, namely there exist $\alpha_1, \dots, \alpha_p \geq 0$, with $\sum_{i=1}^{p} \alpha_i = 1$, such that $A_\alpha := \sum_{i=1}^{p} \alpha_i A_i$ is Schur. However, the existence of such a Schur convex combination does not imply condition (i*) and a counterexample can be found in [8].

A more complete characterization for the existence of a CLF satisfying condition (3.17) is provided in [8] for a DPSS switching between $p = 2$ subsystems. For this particular case a new set of equivalent sufficient conditions for stabilizability can be found and these conditions prove to be stronger than any of the conditions presented in Theorem 3.4. This result is presented is Proposition 3.3 below and its derivation is based on the following lemma.

Lemma 3.1 (S-PROCEDURE)**.** Let $Q_1, Q_2 \in \mathbb{R}^{n\times n}$ be two symmetric matrices, and suppose that there exist $\bar{\mathbf{x}} \neq 0$ such that $\bar{\mathbf{x}}^T Q_1 \bar{\mathbf{x}} > 0$. Then, the following statements are equivalent:

(a) $\forall \mathbf{x} \neq 0$ such that $\mathbf{x}^T Q_1 \mathbf{x} \geq 0$, one finds $\mathbf{x}^T Q_2 \mathbf{x} < 0$;

(b) $\exists \gamma \geq 0$ such that $\gamma Q_1 + Q_2 \prec 0$.

Proposition 3.3. Let $A_1, A_2 \in \mathbb{R}_+^{n\times n}$ be positive matrices. The following facts are equivalent:

(i^2) $\exists P = P^T \succ 0$ such that $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$ satisfies:

$$\min_{i=1,2} \Delta V_i(\mathbf{x}) = \min_{i=1,2} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < 0 \quad \forall \mathbf{x} \neq 0;$$

(ii^2) $\exists P = P^T \succ 0$ and $\eta > 0$ such that $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$ satisfies:

$$\min_{i=1,2} \Delta V_i(\mathbf{x}) = \min_{i=1,2} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < -\eta \mathbf{x}^T \mathbf{x} \quad \forall \mathbf{x} \neq 0;$$

(iii^2) $\exists P = P^T \succ 0$ and $\alpha \in [0, 1]$ such that:

$$\alpha(A_1^T P A_1 - P) + (1-\alpha)(A_2^T P A_2 - P) \prec 0.$$

If (i^2)-(iii^2) hold, then:

(i*) $\exists P = P^T \succ 0$ and $\alpha \in [0, 1]$ such that $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$ satisfies for every $\mathbf{x} > 0$:

$$\alpha \Delta V_1(\mathbf{x}) + (1-\alpha)\Delta V_2(\mathbf{x}) = \mathbf{x}^T \left[\alpha(A_1^T P A_1 - P) + (1-\alpha)(A_2^T P A_2 - P)\right]\mathbf{x} < 0.$$

Conditiion (i*) implies any of the following equivalent facts:

(i) $\exists \mathbf{v} \gg 0$ such that $V(\mathbf{x}) = \mathbf{v}^T \mathbf{x}$ is a Linear CLF that satisfies:

$$\min_{i=1,2} \Delta V_i(\mathbf{x}) = \min_{i=1,2} \mathbf{v}^T (A_i - I_n)\mathbf{x} < 0 \quad \forall \mathbf{x} > 0;$$

(ii) $\exists P = P^T$ of rank 1 such that $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$ is a Quadratic CLF that satisfies:

$$\min_{i=1,2} \Delta V_i(\mathbf{x}) = \min_{i=1,2} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < 0 \quad \forall \mathbf{x} > 0.$$

If (i)-(ii) hold, then the following condition holds:

(iii) $\exists \tilde{P} = \tilde{P}^T \succ 0$ such that $V(\mathbf{x}) = \mathbf{x}^T \tilde{P}\mathbf{x}$ is a Quadratic Positive Definite CLF that satisfies:

$$\min_{i=1,2} \Delta V_i(\mathbf{x}) = \min_{i=1,2} \mathbf{x}^T (A_i^T \tilde{P} A_i - \tilde{P})\mathbf{x} < 0 \quad \forall \mathbf{x} > 0.$$

If (iii) holds, then:

(iv) $\exists P = P^T$ such that $V(\mathbf{x}) = \mathbf{x}^T P\mathbf{x}$ is a Quadratic CLF that satisfies:

$$\min_{i=1,2} \Delta V_i(\mathbf{x}) = \min_{i=1,2} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < 0 \quad \forall \mathbf{x} > 0.$$

Proof. (i^2) $\Rightarrow$ (ii^2): Consider the continuous function:

$$f(\mathbf{x}) := \min_{i=1,2} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x}$$

and the compact set:

$$\mathcal{E} := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1\}.$$

By Wierstrass' theorem and assumption (i^2), it follows that $\max_{\mathbf{x}\in\mathcal{E}} f(\mathbf{x}) < -\eta$, $\eta > 0$. Now, for every $c \in \mathbb{R}_+$ consider the compact set:

$$\mathcal{E}_c := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = c\} = \{\mathbf{x} = c\mathbf{z} : \mathbf{z} \in \mathcal{E}\},$$

and notice that:

$$\begin{aligned} \max_{\mathbf{x}\in\mathcal{E}_c} f(\mathbf{x}) &= \max_{\mathbf{x}\in\mathcal{E}_c} \min_{i=1,2} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} \\ &= \max_{\mathbf{z}\in\mathcal{E}} \min_{i=1,2} \mathbf{z}^T (A_i^T P A_i - P)\mathbf{z} < -\eta. \end{aligned} \tag{3.20}$$

Condition (3.20) implies that:

$$\begin{gathered} \min_{i=1,2} \mathbf{z}^T (A_i^T P A_i - P)\mathbf{z} < -\eta \\ \min_{i=1,2} \frac{\mathbf{x}^T}{c} (A_i^T P A_i - P) \frac{\mathbf{x}^T}{c} < -\eta \\ \min_{i=1,2} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} < -\eta \cdot c^2 \end{gathered}$$

and this ensures that, for every $\mathbf{x} \neq 0$, $f(\mathbf{x}) < -\eta \mathbf{x}^T \mathbf{x}$.

(ii^2) $\Rightarrow$ (iii^2): If either A_1 or A_2 is Schur, the result is obvious. So, we assume that neither of them is. Now, set $Q_i := A_i^T P A_i - P + \eta I_n$ and notice that condition (ii^2) can be rewritten as:

$$\min_{i=1,2} \mathbf{x}^T (A_i^T P A_i - P + \eta I_n)\mathbf{x} = \min_{i=1,2} \mathbf{x}^T Q_i \mathbf{x} < 0 \quad \forall \mathbf{x} \neq 0. \tag{3.21}$$

Observe that, as A_1 is not Schur, there exists $\bar{\mathbf{x}} \neq 0$ such that $\bar{\mathbf{x}}^T (A_1^T P A_1 - P)\bar{\mathbf{x}} \geq 0$, which implies that $\bar{\mathbf{x}}^T Q_1 \bar{\mathbf{x}} > 0$. In addition, condition (3.21) implies that for every $\mathbf{x} \neq 0$ such that $\mathbf{x}^T Q_1 \mathbf{x} \geq 0$, one has $\mathbf{x}^T Q_2 \mathbf{x} < 0$. So, by Lemma 3.1, we can claim that there exists $\gamma \geq 0$ such that:

$$\gamma Q_1 + Q_2 = \gamma (A_1^T P A_1 - P + \eta I_n) + (A_2^T P A_2 - P + \eta I_n) \prec 0,$$

and hence:

$$\gamma(A_1^T P A_1 - P) + (A_2^T P A_2 - P) \prec 0,$$

which amounts to saying that (iii^2) holds for $\alpha = \frac{\gamma}{1+\gamma} \in [0, 1)$.

(iii^2) $\Rightarrow$ (i^2) and (iii^2) $\Rightarrow$ (i*) are obvious.

The remaining conditions follow from Theorem 3.4 for the particular case $p = 2$. $\square$

The above Theorem 3.4 (along with Proposition 3.3 for the case $p = 2$) mutually relates the conditions for the existence of CLFs satisfying (3.17), thus proving that, within such class of CLFs, if a Linear CLF can be found, then a Quadratic CLF can be found, too, and this latter, in turn, ensures the existence of a Quadratic Positive Definite CLF. It is worth noticing that, in a sense, this is the counterpart for stabilizability of the characterization obtained in Theorem 2.3 for stability.

The interest in the existence of a CLF satisfying (3.17) is justified by the possibility of implementing a state-feedback stabilizing switching law. This is shown in Theorem 3.5 below, which provides a stabilization strategy which is independent of the special kind of Lyapunov function we are considering. In addition, the same reasoning would apply to every copositive homogeneous function, thus making this switching rule applicable when dealing with a broader class of Lyapunov functions.

Theorem 3.5. Let a DPSS of the form (1.1) be given. Assume there exists a CLF, which is either linear, quadratic or quadratic positive definite, satisfying the following condition:

$$\min_{i=1,\dots,p} \Delta V_i(\mathbf{x}) = \min_{i=1,\dots,p} V(A_i\mathbf{x}) - V(\mathbf{x}) < 0 \quad \forall \mathbf{x} > 0.$$

Then, the state-feedback switching rule:

$$\sigma(\mathbf{x}(t)) := \min\{k \colon \Delta V_k(\mathbf{x}(t)) \leq \Delta V_i(\mathbf{x}(t)),\ \forall i \in [1, p]\} \tag{3.22}$$

stabilizes the system, i.e. it makes the state evolution goes to zero for every positive initial state.

Proof. Consider first the case when $V(\mathbf{x})$ is a Quadratic (possibly Positive Definite) CLF and hence takes the form $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$. The function:

$$\Delta V(\mathbf{x}) := \min_{i=1,\dots,p} \Delta V_i(\mathbf{x}) = \min_{i=1,\dots,p} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x}$$

is a continuous function that takes negative values in every point of the compact set:

$$\mathcal{E} := \mathbb{R}_+^n \cap \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{x}^T P \mathbf{x} = 1\}.$$

So, by Weierstrass' theorem, $\max_{\mathbf{x}\in\mathcal{E}} \Delta V(\mathbf{x}) = -\eta$, $0 < \eta \leq 1$. Now, for every $c \in \mathbb{R}_+$ consider the compact set:

$$\mathcal{E}_c := \mathbb{R}_+^n \cap \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{x}^T P \mathbf{x} = c\} = \mathbb{R}_+^n \cap \{\mathbf{x} = \sqrt{c}\mathbf{z} \colon \mathbf{z} \in \mathcal{E}\},$$

and notice that:

$$\begin{aligned}\max_{\mathbf{x}\in\mathcal{E}_c} \Delta V(\mathbf{x}) &= \max_{\mathbf{x}\in\mathcal{E}_c} \min_{i=1,\dots,p} \mathbf{x}^T (A_i^T P A_i - P)\mathbf{x} \\ &= \max_{\mathbf{z}\in\mathcal{E}} \min_{i=1,\dots,p} \mathbf{z}^T (A_i^T P A_i - P)\mathbf{z} = -\eta.\end{aligned} \tag{3.23}$$

Condition (3.23) implies that:

$$\min_{i=1,\dots,p} \mathbf{z}^T(A_i^T PA_i - P)\mathbf{z} \le -\eta$$
$$\min_{i=1,\dots,p} \frac{\mathbf{x}^T}{\sqrt{c}}(A_i^T PA_i - P)\frac{\mathbf{x}^T}{\sqrt{c}} \le -\eta$$
$$\min_{i=1,\dots,p} \mathbf{x}^T(A_i^T PA_i - P)\mathbf{x} \le -\eta \cdot c$$

and this ensures that for every $\mathbf{x} > 0$ we have:

$$\Delta V(\mathbf{x}(t)) \le -\eta \mathbf{x}^T P\mathbf{x}.$$

Finally, notice that:

$$\begin{aligned} V(\mathbf{x}(t+1)) &= V(\mathbf{x}(t)) + \big[V(A_i\mathbf{x}(t)) - V(\mathbf{x}(t))\big] \\ &= V(\mathbf{x}(t)) + \Delta V(\mathbf{x}(t)) \\ &\le (1-\eta)\mathbf{x}^T(t)P\mathbf{x}(t) \\ &\le (1-\eta)^{t+1}\mathbf{x}^T(0)P\mathbf{x}(0). \end{aligned}$$

Thus $V(\mathbf{x}(t))$ converges to zero, and $\mathbf{x}(t)$ converges to zero in turn.

The proof in case of a linear CLF $V(\mathbf{x}) = \mathbf{v}^T\mathbf{x}$ follows the same lines, upon assuming $\mathcal{E} := \mathbb{R}_+^n \cap \{\mathbf{x} \in \mathbb{R}^n : \mathbf{v}^T\mathbf{x} = 1\}$. □

Remark 3.2. Interestingly enough, under certain assumptions the stabilizing switching strategy provided by just stated Theorem 3.5 is completely equivalent to the one described by Geromel and Colaneri in [9] (see Proposition 3.2 above). This is the case when there exist $P = P^T \succ 0$ and $\alpha_1, \dots, \alpha_p \ge 0$, with $\sum_{i=1}^p \alpha_i = 1$, such that $V(\mathbf{x}) = \mathbf{x}^T P\mathbf{x}$ satisfies:

$$\sum_{i=1}^p \alpha_i \Delta V_i(\mathbf{x}) = \sum_{i=1}^p \alpha_i \mathbf{x}^T(A_i^T PA_i - P)\mathbf{x} < 0 \quad \forall \mathbf{x} \ne 0,$$

which amounts to saying that:

$$\sum_{i=1}^p \alpha_i(A_i^T PA_i - P) \prec 0. \tag{3.24}$$

Note that this is a stronger condition with respect to condition (i*) in Theorem 3.4 and coincides with any of the conditions (i^2)-(iii^2) in Proposition 3.3 when $p = 2$. We now prove that if (3.24) holds for suitable α_i's and P, then the Lyapunov-Metzler inequalities (3.3) admit a solution. Indeed, assuming that (3.24) holds, then there exists $\varepsilon > 0$ sufficiently small such that:

$$\sum_{i=1}^p \alpha_i(A_i^T PA_i - P) + \varepsilon I_n \preceq 0,$$

or, equivalently:

$$\sum_{i=1}^p \alpha_i(A_i^T PA_i + \varepsilon I_n) - P \preceq 0.$$

This, in turn, implies that for all $i \in [1, p]$:

$$A_i^T \left[\sum_{i=1}^{p} \alpha_i (A_i^T P A_i + \varepsilon I_n) \right] A_i - A_i^T P A_i \preceq 0,$$

and hence:

$$A_i^T \left[\sum_{i=1}^{p} \alpha_i (A_i^T P A_i + \varepsilon I_n) \right] A_i - A_i^T P A_i - \varepsilon I_n \prec 0,$$

which amounts to saying that each of the p matrices $\bar{P}_i = A_i^T P A_i + \varepsilon I_n$, $i \in [1, p]$, satisfies the Lyapunov-Metzler inequalities (3.3):

$$A_i^T \left(\sum_{j=1}^{p} \alpha_j \bar{P}_j \right) A_i - \bar{P}_j \prec 0.$$

Finally, it is straightforward to see that the switching strategy (3.22) is totally equivalent to the state-feedback switching rule (3.4) of Proposition 3.2:

$$\begin{aligned} \sigma(\mathbf{x}(t)) &= \operatorname*{argmin}_{i=1,\ldots,p} \mathbf{x}^T(t)(A_i^T P A_i - P)\mathbf{x}(t) \\ &= \operatorname*{argmin}_{i=1,\ldots,p} \mathbf{x}^T(t)(A_i^T P A_i)\mathbf{x}(t) \\ &= \operatorname*{argmin}_{i=1,\ldots,p} \mathbf{x}^T(t)(A_i^T P A_i + \varepsilon I_n)\mathbf{x}(t) \\ &= \operatorname*{argmin}_{i=1,\ldots,p} \mathbf{x}^T(t)\bar{P}_i\mathbf{x}(t). \end{aligned}$$

Remark 3.3. When a DPSS (1.1) satisfies condition (i*) of Theorem 3.4, different state-feedback switching strategies can be adopted and the natural question arises whether they are just the same or, if not, which of them ensures better convergence performance. In this regard, a few considerations should be done. First of all, notice that the switching strategies based on Linear CLFs and those based on Quadratic CLFs of rank 1 are just the same. Indeed, as clarified in the proof of (ii) $\Leftrightarrow$ (i) in Theorem 2.3, a matrix $P = P^T$ of rank 1 satisfies condition (ii) if and only if it can be expressed as $P = \mathbf{v}\mathbf{v}^T$, for some vector $\mathbf{v} \gg 0$. On the other hand, by the non-negativity of the quantities involved, we have:

$$\begin{aligned} \sigma(\mathbf{x}(t)) &= \min\{k\colon \mathbf{v}^T(A_k - I_n)\mathbf{x} \le \mathbf{v}^T(A_i - I_n)\mathbf{x},\ \forall i \in [1,p]\} \\ &= \min\{k\colon \mathbf{v}^T A_k\mathbf{x} \le \mathbf{v}^T A_i\mathbf{x},\ \forall i \in [1,p]\} \\ &= \min\{k\colon \mathbf{x}^T A_k^T\mathbf{v}\mathbf{v}^T A_k\mathbf{x} \le \mathbf{x}^T A_i^T\mathbf{v}\mathbf{v}^T A_i\mathbf{x},\ \forall i \in [1,p]\} \\ &= \min\{k\colon \mathbf{x}^T(A_k^T\mathbf{v}\mathbf{v}^T A_k - \mathbf{v}\mathbf{v}^T)\mathbf{x} \le \mathbf{x}^T(A_i^T\mathbf{v}\mathbf{v}^T A_i - \mathbf{v}\mathbf{v}^T)\mathbf{x}, \forall i \in [1,p]\} \end{aligned}$$

and hence the switching sequences based on $\mathbf{v}^T\mathbf{x}$ and on $\mathbf{x}^T\mathbf{v}\mathbf{v}^T\mathbf{x}$ are just the same. Secondly, since the set of Quadratic CLFs of arbitrary rank include the set of Quadratic CLFs of rank 1, the stabilizing switching laws based on the latter are a subset of those based on the former. Similarly, the class of switching laws based on Quadratic CLFs encompasses those based on Quadratic Positive Definite CLFs and hence it ensures convergence performances at least as good as the latter. So, we can conclude that, in order to optimize the convergence performances, it is always convenient to resort to switching laws based on Quadratic CLFs.

Before concluding, let us revise what we have seen throughout this chapter, where the determination of a stabilizing switching strategy has been addressed. First, the piecewise CLF method has been considered: in this context the stabilization condition is expressed by means of a set of matrix inequalities, whose solution, if any, defines a Piecewise Linear CLF, which, in turn, allows to design a stabilizing switching law. This underlying idea has led to results presented in Theorem 3.1 and in Theorem 3.2 by resorting to different classes of Metzler matrices. Unfortunately, the aformentioned matrix inequalities are non-linear and hence looking for a simpler, although more conservative, reformulation might be convenient, as seen in Theorem 3.3. Finally, an alternative approach has been presented: Theorem 3.4 investigates conditions for the existence of a special class of CLFs, while Theorem 3.5 ensures the possibility of implementing a state-feedback stabilizing switching law whenever such a Lyapunov function is available.

Part II

Application of Positive Switched Systems to HIV Treatment Modeling

Chapter 4

The Immune System and the Human Immunodeficiency Virus

In this chapter we provide some basic concepts on the immune system and the Human Immunodeficiency Virus (HIV). In particular, in the first section a brief overview over the immune system's components is presented, while in the second section the typical course of HIV disease is described, along with antiretroviral therapies commonly used for treating individuals infected. In both cases we will limit to some essential notions and we refer the reader to [2],[16],[19] for a complete and exhaustive presentation of the topics.

All figures included in the present chapter are taken from [2].

4.1 The immune system

The immune system is a remarkable defense mechanism: it provides the means to make rapid, specific and protective responses against foreign and harmful substances, microorganisms, toxins, and malignant cells. Its central role is illustrated by the tragic example of severe immunodeficiencies, as seen in both genetically determined diseases and in the Acquired Immunodeficiency Syndrome (AIDS).

The immune system can be divided into two parts, called innate and adaptive. The innate immune system acts in a nonspecific manner, which means that defense mechanisms become active independently of the invading pathogen. Some examples include physical barriers such as the skin, chemical barriers like lysozymes, the complement system, granulocytes and macrophages. The adaptive immune system is responsible for highly specific reactions which require recognition of specific "non-self" antigens. It also creates immunological memory after an initial response to a specific pathogen, which leads to an enhanced response to subsequent encounters with the same pathogen.

When a foreign, potentially pathogenic agent breaches the outer barriers of the body, mechanisms of innate immunity are first activated (also known as the primary immune defenses), while specific immune defense factors are mobilized later on to fortify and regulate these primary defenses.

White blood cells, also called leukocytes, constitute the immune system's components. Like all other components of the blood, they originate from pluripotent hematopoietic stem cells of the bone marrow and they subsequently differentiate into various types: polymorphonuclear granulocytes, lymphocytes and monocytes. The class of polymorphonuclear granulocytes includes neutrophils, eosinophils, and basophils, which are all characterized by a segmented nucleus and by the presence of differently staining granules in their cytoplasm when viewed under microscopy. Lymphocytes can be distinguished into two different types: T cells (Thymus cells) and B cells (Bone cells). Monocytes circulate in the blood and continually migrate into tissues, where they mature and become macrophages. As lymphocytes and monocytes are characterized by the apparent absence of granules in their cytoplasm, they are collectively called mononuclear agranulocytes in order to distinguish them from granulocyte cells.

The various components of the immune system play different roles in the immune response. Granulocytes and monocytes have the ability to ingest particles, microorganisms, dead cell debris and fluids by phagocytosis, which is the main mechanism of innate immunity. T and B lymphocytes are the effector cells of the adaptive immune response: the former are responsible for the so called cell-mediated immunity, while the latter produce antibodies, which constitute the humoral (namely, antibody-mediated) immune response. Both T and B cells carry on their surface specialized receptors which recognize and bind with only one specific antigen: this "lock-and-key" mechanism accounts for the high specificity of the adaptive immune response. A brief description of T cells, B cells and macrophages is presented below.

T lymphocytes

T-lymphocytes derive from precursors in hematopoietic tissue, undergo differentiation in the thymus (hence the name T lymphocytes), and are then seeded to the peripheral lymphoid tissue and to the recirculating pool of lymphocytes. They can be distinguished from other lymphocytes by the presence on the surface of a T cell Receptor (TCR), a protein that binds specifically to antigens trapped by the Major Histocompatibility Complex (MHC) molecules. The diversity of TCR is achieved by means of rearrangement of gene segments coding for its constituent chains. Differently from B cell receptors, T cell receptors are unable to recognize free antigens. Rather, they recognize cell-associated molecules, namely antigenic peptides, derived by proteolysis of the antigen, in combination with either MHC class I or II.

There are several subsets of T cells mediating a wide range of immunologic functions and differing in how they recognize antigens. T Helper cells and Cytotoxic T cells are among the most important subpopulations. T Helper cells are also called CD4+ T cells as they express the surface molecule CD4, which recognizes only MHC class II-associated antigens. The CD4 molecule also serves as the binding protein for the HIV. T Helper cells exercise an important regulatory effect on other lymphocytes: they help B cells develop into antibody-producing cells and they produce, or induce the production of, cytokines by which means they activate macrophages and enhance their microbicidal activity.

Cytotoxic T cells are also known as CD8+ T cells as they express the CD8 molecule that binds to MHC class I. They are capable of efficiently lysing target cells, like virus-infected cells, tumor cells and allogeneic cells. Such lysing mechanism involves the production of perforin, a molecule that inserts into the membrane of target cells and promotes its lysis.

B lymphocytes

B lymphocytes derive from hematopoietic stem cells by a complex set of differentiation events occurring in the bone marrow (hence the name B lymphocytes). When a mature B cell encounters the antigenic epitope specifically recognized by its surface receptor, it becomes an activated B cell. Such an activation process, which may occur directly or through the interaction with a Helper T cell, is responsible for the proliferation and differentiation of the B cell into either a plasma-cell or a memory cell. Plasma-cells secrete in the blood antibodies exhibiting the same antigen specificity as the B cell receptor. The great variety of antibodies, known as Immunoglobulins (Ig), is ensured by a process of continuous diversification of the genetically identical B-cell precursors. Memory B cells give rise to antibody-secreting cells upon re-challenge of the individual: the hallmarks of this response to re-challenge is that it is of greater magnitude, occurs more promptly and is composed of antibodies with higher affinity for the antigen.

Macrophages

Macrophages, which represent the mature form of monocytes, are large, highly phagocytic and relatively long-lived cells distributed in almost all tissues. In contrast with T and B lymphocytes, macrophages cannot recirculate or re-initiate DNA replication except in a limited way. They are important for the generation of both adaptive immune response and innate immune response: they regulate activation of T and B lymphocytes through their specialized derivatives known as Dendritic Cells, they process and present antigens, they produce proteins (chemokines and cytokines) that activate other immune system cells and they phagocytose cellular debris, pathogens, necrotic and apoptotic cells.

4.2 HIV infection

The Human Immunodeficiency Virus was first identified in 1983 and was shown to be the cause of Acquired Immunodeficiency Syndrome (AIDS) in 1984. Since the introduction of the combination of antiretroviral therapies a significant decrease in morbidity and a reduction in mortality have been reported. Nevertheless, the treatment of HIV infected patients is still of major importance in today's medicine as, to date, no curative therapy resulting in complete eradication of the virus is available.

4.2.1 The virus and its replication cycle

HIV belongs to the family of retroviruses, a class of viruses characterized by having an enzyme, called reverse transcriptase, that transcribes single-stranded genomic RNA into double-stranded DNA.

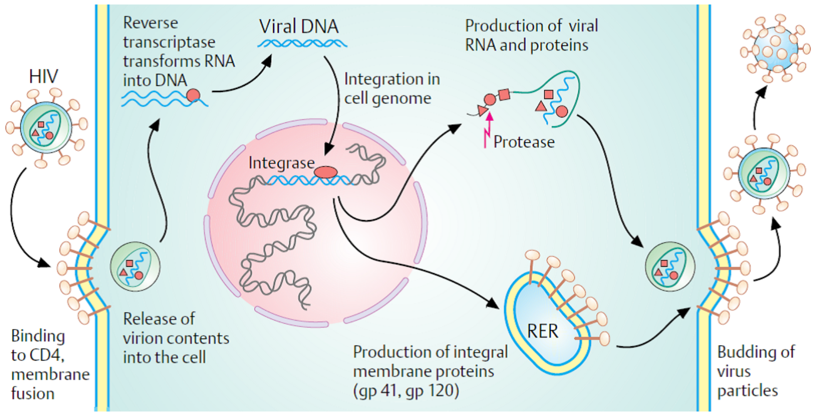

Figure 4.1: Replication cycle of HIV in the host cell.

The HIV viron is about 100nm in diameter with 72 spikes derived from glycoproteins gp120 and gp41. Its genome contains nine genes encoding the structural proteins of the core, the envelope glycoproteins, the enzymes involved in viral replication and integration and other proteins essential for viral production. The attachment of the viral particle to the surface of the host cell requires binding of gp120 with the CD4 molecule. Consequently, HIV can infect CD4+ T lymphocytes and other cells bearing the CD4 marker on their surface, such as monocytes, tissue macrophages and Langherans cells.

Figure 4.1 illustrates HIV's replication cycle in the host cell. Once the contents of the virion have been released into the cytoplasm of the cell, the reverse transcriptase starts to transcribe the RNA in double-stranded DNA, which is then incorporated into the host-cell genome. Immediately afterward, the production of viral proteins, supplied by cell's protein synthesis mechanism, can be initiated: immature viral particles are first formed, which undergo a maturation process facilitated by the enzyme protease, and eventually become new viruses.

4.2.2 HIV disease progression

HIV infection is predominantly an infection of the immune system: it is characterized by the eventual depletion of CD4+ T cells, leading to severe immune deficiency which in turn provokes devastating effects on the patient's health.

As illustrated in Figure 4.2, the typical course HIV infection goes through three stages: the primary infection, the asymptomatic or latency period and the final progression to AIDS. The median time between primary HIV infection and the development of AIDS is approximately 10 years.

The primary infection can either remain inapparent or manifest through symptoms such as fever, chills and lymphadenopathy. This initial phase of infection is associated with a burst of circulating free viruses and a decline in the level of CD4+ T cells, CD8+ T cells and B cells. After resolution of the acute syndrome, both CD4+ and CD8+ T cell levels usually rebound to near normal levels, while the overall virus population decreases sharply owing to the development of potent cellular and humoral responses: the cellular cytotoxic im-

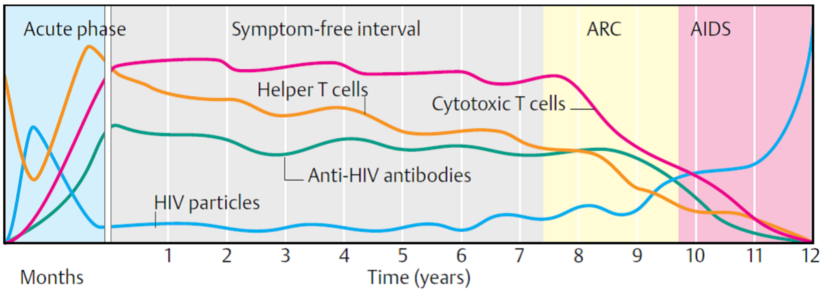

Figure 4.2: Typical course of HIV infection.

mune response is activated, antibodies produced by activated B cells bind free viral particles and make them digestible for macrophages. Nevertheless, HIV is not completely cleared from the body. Indeed, during the course of the second stage (also called chronic or persistent infection), ongoing virus replication can be detected consistently in the blood and in the lymphoid tissue, even if the patient does not exhibit any evidence of disease. During this latency period CD4+ T cells counts usually decrease gradually.[1] When the CD4+ T cell count has dropped lower than 250 cells/mm^3, the individual is said to have AIDS. At this stage, the immune system collapses and the risk for the development of opportunistic diseases is high: the loss of integrity of the immune function allows ubiquitous environmental organisms with limited virulence to become life-threatening pathogens. AIDS is also characterized by the occurrence of certain malignant neoplasms (like Kaposi's sarcoma) and central nervous systems manifestations in its late stages.

4.2.3 Antiretroviral drugs for HIV infection

Since the identification of HIV as the causative agent for AIDS, a number of antiretroviral drugs for treating individuals infected with HIV have been developed. There are five classes of substances (Figure 4.3) available for HIV therapy, acting on different stages of the HIV life-cycle:

- Fusion Inhibitors (FIs) interfere with the binding, fusion and entry of an HIV virion to a human cell;
- Nucleosidic Reverse Transcriptase Inhibitors (NRTIs) bind to the active center of reverse transcriptase and are integrated in the DNA strands, resulting in "chain termination";
- Nonnucleosidic Reverse Transcriptase Inhibitors (NNRTIs) bind to the catalytic center of reverse transcriptase and inactivate the enzyme, thereby inhibiting the production of viral cDNA;
- Integrase Nuclear Strand Transfer Inhibitors (INSTIs) inhibit the viral enzyme integrase, which is responsible for integration of viral DNA into the DNA of the infected cell;

[1]As we will see in Chapter 5, such graduality is very important when modeling HIV dynamics, as it allows us to assume that CD4+ T cell counts are approximately constant.

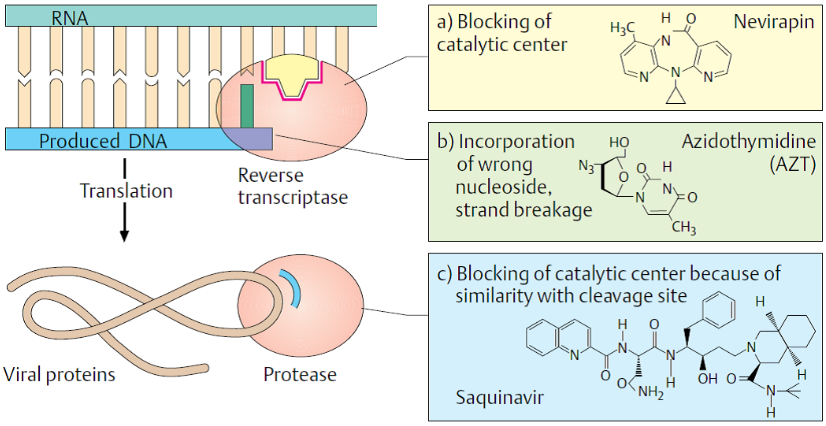

Figure 4.3: Antiretroviral drugs for HIV infection.

- Protease Inhibitors (PI) inhibit viral protease and thus viral maturation.

A combination of at least three drugs from at least two substance classes is usually administered. Such combination of antiretrovirals, known as Highly Active Antiretroviral Therapy (HAART), often results in suppression of plasma HIV RNA and a significant increase in CD4+ T cell counts. However, HAART is unable to completely eradicate HIV infection owing to the persistent HIV replication in reservoirs sites such as lymphoid tissue and latently infected CD4+ T cells, as well as the continual evolution of HIV envelop and protease genes.

4.2.4 Viral mutation and switching regimen

One striking feature of the HIV is the high level of genetic variability within a single infected patient. The leading cause of such a diversity is the high error rate of reverse transcriptase: this enzyme does not have any proof-correction activity, and appears to introduce errors every 10^4 nucleotides. In addition, fast HIV replication cycle and prolonged duration of HIV infection are major contributors to viral diversity. Such level of genetic variability, much higher than most human RNA viruses, is a prominent issue in HIV infection. First of all, it contributes to the ability of HIV to evade the humoral immune response: when, towards the end of the latency period, around 10 million HIV variants per day are produced, it becomes impossible for the immune system to control the infection effectively. Secondly, the generation of many variants of HIV in the presence of antiretroviral drugs may cause drug resistance, and hence the reduction in effectiveness of a drug.

A relatively recent strategy to manage viral mutation and limit, or delay, the emergence of drug-resistant variants is the proactive switching and alternation of HAART regimens. The design of drug sequencing, however, is still a point of discussion as there is no general consensus among clinicians on the optimal time to change therapy. There is, indeed, a sort of tradeoff: on the one hand, switching drugs too early risks poor adherence to a new drug regimen and limits future treatments options; on the other hand, switching drugs too late allows the accumulation of mutations and the selection of drug-resistant variants. In this context, developing HIV mutation models and performing simulation studies may help to design switching strategies able to delay the emergence of highly resistant mutant viruses.

Chapter 5

Viral Mutation Treatment Models

Since 1990s several mathematical models of HIV infection have been proposed [22],[38],[34],[11],[10], each of them describing one or more different aspects of HIV dynamics: the interaction with the immune system's components, the primary infection and the asymptomatic stage, the long-term behavior of the infection and the progression to AIDS, and, more recently, the problem of viral mutation leading to the generation of many variants of HIV. These models have the ability to reflect clinical results they have been thought to describe, but, obviously, none of them can fully explain all events observed to occur in practice.

In what follows we focus on two different models of HIV mutation treatment, both of them involving positive switched systems introduced in Part I of the present work.

5.1 A 4 variant, 2 drug combination model

The first model we present is taken from [11] and deals with 4 viral variants and 2 drug combinations. As we will see in Chapter 6, Hernandez-Vargas and coauthors in the same work [11] used this model to perform interesting simulation studies testing the effectiveness of switching among two treatments.

5.1.1 Model's assumptions

In order to allow control analysis and optimization of treatment switching, the viral mutation model should not be too complicated and, to this aim, some simplifying assumptions are needed. First of all, scalar dynamics for each mutant are considered, by this meaning that we only focus on the viral concentration $V_i(t)$ of mutant i and we do not consider its possible different states, such as infected T cells or infected macrophages. Secondly, we assume that the viral clearance rate is constant, even if it actually depends on the treatment which is being administered and on the viral genetics. In addition, another important assumption deals with the mutation rate between species: this rate is taken as a constant, but in practice it depends on the replication rate of the variants and on the treatment regime. Finally, we consider a deterministic model, which does not include random variations. Even if this is a significant limitation, it should be noted that, owing to the linearity of the system, the deterministic model describes the expected behavior of a stochastic model.

5.1.2 Mutation model

As mentioned above, the model has 4 different viral variants (also called "genotypes" or "strains") and 2 possible drug therapies that can be administered: this corresponds to a 4-th order positive switched system switching among 2 different subsystems. The state variables are briefly described as follow:

- $x_1(t)$ represents the Wild Type (WT) variant, a genotype susceptible to both therapies, which would be the most prolific strain in the absence of any drugs;
- $x_2(t)$ represents Genotype 1 (G1), a genotype that is resistant to therapy 1, but is susceptible to therapy 2;
- $x_3(t)$ represents Genotype 2 (G2), a genotype that is resistant to therapy 2, but is susceptible to therapy 1;
- $x_4(t)$ represents the Highly Resistant Genotype (HRG), a variant with low proliferation rate, but that is resistant to all drug therapies.

Describing the behavior of any viral variant means taking into account three different terms: virus proliferation, natural death and virus mutation. Recalling the assumptions presented in the previous section, the dynamics of the i-th viral population under the j-th drug therapy can be described by the following ordinary differential equation:

$$\dot{x}_i(t) = \rho_{i,j} x_i(t) - \delta x_i(t) + \sum_{\substack{l=1 \\ l \neq i}}^{4} \mu m_{i,l} x_l(t), \qquad (5.1)$$

where:

- $\rho_{i,j}$ is the replication rate of genotype i under therapy j;
- δ is the death (or decay) rate;
- μ is the mutation rate;
- $m_{i,l} \in \{0, 1\}$ represents the genetic connections between genotypes, that is $m_{i,l} = 1$ if and only if it is possible for genotype l to mutate into genotype i and $m_{i,l} = 0$ otherwise.

Taking into account all different viral variants, equation (5.1) can be rewritten in vector form as:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \rho_{1,j} & 0 & 0 & 0 \\ 0 & \rho_{2,j} & 0 & 0 \\ 0 & 0 & \rho_{3,j} & 0 \\ 0 & 0 & 0 & \rho_{4,j} \end{bmatrix} \mathbf{x}(t) - \delta \mathbf{x}(t) + \mu \begin{bmatrix} 0 & m_{1,2} & m_{1,3} & m_{1,4} \\ m_{2,1} & 0 & m_{2,3} & m_{2,4} \\ m_{3,1} & m_{3,2} & 0 & m_{3,4} \\ m_{4,1} & m_{4,2} & m_{4,3} & 0 \end{bmatrix} \mathbf{x}(t)$$

Table 5.1: Replication rates for viral variants and therapy combinations.

Variant	Therapy 1	Therapy 2
Wild Type (x_1)	$\rho_{1,1} = 0.05$	$\rho_{1,2} = 0.05$
Genotype 1 (x_2)	$\rho_{2,1} = 0.40$	$\rho_{2,2} = 0.05$
Genotype 2 (x_3)	$\rho_{3,1} = 0.05$	$\rho_{3,2} = 0.40$
HR Genotype (x_4)	$\rho_{4,1} = 0.30$	$\rho_{3,2} = 0.30$

Finally, if the two different drug therapies are administered according to the switching rule $\sigma(t)$, $\sigma(t) \in [1,2]$, we can define the matrices $R_{\sigma(t)} := diag\{\rho_{i,\sigma(t)}\}$ and $M := [m_{i,l}]$ and obtain the following CPSS:

$$\dot{\mathbf{x}}(t) = (R_{\sigma(t)} - \delta I_4 + \mu M)\mathbf{x}(t), \tag{5.2}$$

where the matrix $R_{\sigma(t)} - \delta I_4 + \mu M$ is certainly a Metzler matrix.

5.1.3 Model's parameters

The various replication rates for viral variants and therapy combinations are shown in Table 5.1: they describe a symmetric scenario, in the sense that therapy 1 inhibits G2 with the same intensity as therapy 2 inhibits G1 (in practice, this scenario is rather idealized since there will be some small differences in relative proliferation ability). Two further considerations underlie numerical values reported in Table 5.1. First of all it is assumed that genetic distance from WT makes the reproduction rate decrease: this is the reason why $\rho_{4,1} = \rho_{4,2}$ is smaller than $\rho_{2,1}$ and $\rho_{3,2}$. Secondly, we take the replication rate in the absence of drugs as $0.5day^{-1}$ (as suggested by clinical data) and we consider therapy's effectiveness equal to 90%: this justifies the fact that $\rho_{1,1}$, $\rho_{1,2}$, $\rho_{2,2}$ and $\rho_{3,1}$ are all equal to $0.05day^{-1}$ (however, the replication rate of G1 is lower under drug therapy 1 and the same holds for reproduction rate of G2 under drug therapy 2).

As regards genetic connections between genotypes, we consider the symmetric and circular mutation graph shown in Figure 5.1: only connections WT $\leftrightarrow$ G1, G1 $\leftrightarrow$ HRG, HRG $\leftrightarrow$ G2 and G2 $\leftrightarrow$ WT are allowed. Hence, the mutation matrix results as follow:

$$M = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

For the sake of simplicity, we assume that other connections requiring double mutations are of negligible probability.

As far as the remaining parameters are concerned, we take the viral mutation rate as $\mu = 10^{-4}$ and the viral clearance rate as $\delta = 0.24day^{-1}$, corresponding to a half life of slightly less than 3 days.

5.1.4 Model's discretization and final rearrangement

The viral mutation treatment model just introduced is described in the continuous time, however, since measurements can only reasonably be made infre-

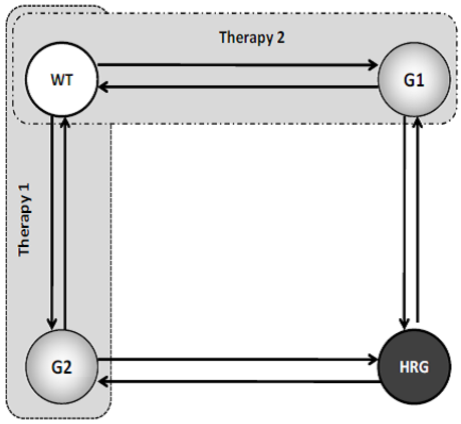

Figure 5.1: Mutation tree for the 4 variant, 2 drug combination model.

quently, we are interested in its discrete-time counterpart. Hence, we consider a regular treatment interval τ during which treatment is fixed and, denoting by $k \in \mathbb{N}$ the number of sampling intervals since $t = 0$, we obtain the following DPSS:

$$\mathbf{x}(k+1) = A_{\sigma(k)}\mathbf{x}(k), \tag{5.3}$$

where $\mathbf{x}(k) = \mathbf{x}(k\tau)$ is the sampled state and the positive matrix $A_{\sigma(k)}$ is given by:

$$A_{\sigma(k)} := e^{(R_{\sigma(k)} - \delta I + \mu M)\tau}.$$

Finally, an important remark concerns system's stabilizability: even if necessary and sufficient conditions for the stabilizability of a DPSS do not exist, biological reasons suggest that the system (5.3) is not stabilizable. Indeed, clinical experience shows that highly resistant genotypes emerging from viral mutation process escape the effects of treatments and immune system (namely, the state evolution asymptotically diverges). For this reason, in order to ensure at least the subsystems' stability, it is convenient to introduce exponential weighting on the new coordinates:

$$\hat{\mathbf{x}}(k) = e^{-\beta k\tau}\mathbf{x}(k\tau).$$

The above variable transformation leads to:

$$\begin{aligned}\hat{\mathbf{x}}((k+1)\tau) &= e^{-\beta(k+1)\tau}\mathbf{x}((k+1)\tau)\\ &= e^{-\beta\tau}A_{\sigma(k\tau)}e^{-\beta k\tau}\mathbf{x}(k\tau)\\ &= e^{-\beta\tau}A_{\sigma(k\tau)}\hat{x}(k\tau)\end{aligned}$$

Hence, we have obtained the following DPSS:

$$\hat{\mathbf{x}}(k+1) = \hat{A}_{\sigma(k)}\hat{\mathbf{x}}(k), \tag{5.4}$$

where $\hat{A}_{\sigma(k)} := e^{-\beta\tau}A_{\sigma(k)}$, $\hat{\mathbf{x}}(k) = \hat{\mathbf{x}}(k\tau)$ and $\sigma(k) = \sigma(k\tau)$ is constant in the interval $[k\tau, (k+1)\tau]$.

5.2 A 64 variant, 3 drug combination model

A far more complex model with respect to the one introduced in the previous section is presented in [10]: this model takes into account 64 viral variants and switching among 3 different drug combinations. In the same work [10], a Model Predictive Control approach is proposed in order to analyze drug regimens and maximize the delay till viral escape.

5.2.1 Model's assumptions

In order to provide a more extensive description for HIV dynamics, one at least of the assumptions presented in section 5.1.1 should be released. In particular, differently from the previous case, a set of states for each possible genotype is considered. This is consistent with the ability of HIV to infect a number of different cells, such as CD4+ T cells, macrophages and dendritic cells. For the sake of simplicity, three states for each variant are included: free viral particles, infected CD4+ T cells and latently infected cells, namely those cells which can be activated and start reproducing virus later. It is worth noticing that such state variables' selection reflects the fact that monitoring viral levels in the plasma, together with CD4+ T cell counts, plays an important role in deciding which therapy should be administered.

As far as the remaining assumptions of section 5.1.1 are concerned, they are still valid: independence of death rates and mutation rates from treatment and mutant, and the deterministic (namely non-stochastic) nature of the model.

In addition, a further assumption is needed, namely the time-invariance of uninfected CD4+ T cell counts: this is consistent with clinical data of the initial infection stage, until full progression to a dominant highly resistant mutant (as seen in Chapter 4), and has the great advantage of making the dynamics essentially linear.

5.2.2 Mutation model

The model deals with 64 viral variants and 3 different drug combinations. Recalling that three different states for each possible genotype are taken into account, this corresponds to a positive switched system switching among 3 different subsystems where each subsystem has $64 \cdot 3 = 192$ state variables. In particular, for each i-th mutant, with $i = 1, \ldots, 64$, we have the following state variables:

- $T_i^*(t)$ represents the (active) infected CD4+ T cell population;
- $L_i(t)$ represents the latently infected cells;
- $V_i(t)$ represents the viral population.

It is worth recalling that, among infected CD4+ T lymphocytes, a proportion of cells passes into the (active) infected cell population, whereas the remaining part passes into the latently infected cell population: we consider this proportion equal to ψ.

Furthermore, in all equations presented below, T represents the uninfected CD4+ T cell population, which is assumed being approximately constant as mentioned above.

Describing the behavior of the infected cell population requires taking into account four different terms: natural death, activation of latently infected cells, infection of healthy CD4+ T cells and virus mutation. The resulting equation is given by:

$$\dot{T}_i^*(t) = -\delta_1 T_i^*(t) + \alpha L_i(t) + \psi b_{i,j} T V_i(t) + \sum_{\substack{l=1 \\ l \neq i}}^{64} \mu m_{i,l} V_l(t) T, \tag{5.5}$$

where:

- δ_1 is CD4+ T cells' death rate;
- α is the activation rate;
- $b_{i,j}$ represents the infectivity of genotype i under treatment j. It depends on the genotype and the therapy that is being used according to the following relation:

  $$b_{i,j} = \bar{b} \beta_{i,j} f_i,$$

  where $\bar{b}$ is the infectivity rate, $\beta_{i,j}$ represents the infection efficiency for genotype i under treatment j and f_i represents the fitness of genotype i;
- μ is the mutation rate;
- $m_{i,l} \in \{0, 1\}$ represents the genetic connections between genotypes.

The dynamic evolution of the latently infected cell population is controlled by the same mechanisms which govern the behavior of the infected cell population, except for the absence of the viral mutation term:

$$\dot{L}_i(t) = -\delta_2 L_i(t) - \alpha L_i(t) + (1 - \psi) b_{i,j} T V_i(t), \tag{5.6}$$

where δ_2 is the death rate of the latently infected cells.

In order to describe the behavior of the viral population $V_i(t)$, the following terms should be considered: virus proliferation, infection of healthy CD4+ T cells and natural death. The dynamic equation is hence given by:

$$\dot{V}_i(t) = e_{i,j} T_i^*(t) - \varphi T V_i(t) - \delta_3 V_i(t), \tag{5.7}$$

where:

- $e_{i,j}$ represents the viral proliferation rate of genotype i under treatment j. It depends on the fitness of the genotype and the therapy that is being administered according to the following relation:

  $$e_{i,j} = \bar{e} \varepsilon_{i,j} f_i,$$

  where $\bar{e}$ is the proliferation rate, $\varepsilon_{i,j}$ represents the production efficiency for genotype i under treatment j and f_i represents the fitness of genotype i;
- φ represents the rate at which a viral particle meets (and infects) an healthy T Helper cell;

- δ_3 is the virus' death rate.

Note that if we remove the assumption of constant healthy CD4+ T cell counts, all the equations (5.5), (5.6) and (5.7) become non-linear owing to the product $TV_i(t)$.

If we consider the state vector of genotype i, $\mathbf{x}_i(t) = \begin{bmatrix} T_i^*(t) & L_i(t) & V_i(t) \end{bmatrix}^T$, equations (5.5), (5.6), (5.7) can be rewritten in vector form as:

$$\dot{\mathbf{x}}_i(t) = \begin{bmatrix} \delta_1 & \alpha & \psi b_{i,j} T \\ 0 & -(\alpha + \delta_2) & (1-\psi) b_{i,j} T \\ e_{i,j} & 0 & -(\varphi T + \delta_3) \end{bmatrix} \mathbf{x}_i(t) + \mu T \begin{bmatrix} 0 & m_{1,2} & m_{1,3} \\ m_{2,1} & 0 & m_{2,3} \\ m_{3,1} & m_{3,2} & 0 \end{bmatrix} \mathbf{x}_i(t)$$

Now, it is convenient to introduce the matrix:

$$\Lambda_{i,j} = \begin{bmatrix} \delta_1 & \alpha & \psi b_{i,j} T \\ 0 & -(\alpha + \delta_2) & (1-\psi) b_{i,j} T \\ e_{i,j} & 0 & -(\varphi T + \delta_3) \end{bmatrix}.$$

Finally, if the three drug regimens are alternated according to the switching rule $\sigma(t)$, $\sigma(t) \in [1, 3]$, and the matrix M is defined as $M := [m_{i,l}]$, we can build up the whole state vector $\mathbf{x}(t) = \begin{bmatrix} \mathbf{x}_1(t) & \mathbf{x}_2(t) & \dots & \mathbf{x}_{64}(t) \end{bmatrix}^T$ and obtain the following CPSS:

$$\dot{\mathbf{x}}(t) = \left(\begin{bmatrix} \Lambda_{1,\sigma} & 0 & \dots & 0 \\ 0 & \Lambda_{2,\sigma} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \Lambda_{64,\sigma} \end{bmatrix} + \mu T M \right) \mathbf{x}(t), \tag{5.8}$$

where the state matrix is certainly a Metzler matrix.

5.2.3 Model's parameters

Differently from the previous model case, we are not interested in all numerical values of the numerous system's parameters and the reason is that in Chapter 6 we will only hint at results obtained by Hernandez-Vargas and coauthors using the 64 variant, 3 drug combination model and the MPC approach. We provide however a few interesting considerations concerning the mutation tree and the fitness of genotype.

As far as the mutation tree (and hence the matrix M) is concerned, the viral variants are organized in a three-dimensional lattice as shown in Figure 5.2: the wild type genotype g_1 is susceptible to all therapies, while a number of independent mutations are required to achieve resistance to all therapies, represented by genotype g_{64}. Clearly, viral variant g_1 would be the most prolific variant in the absence of any drugs, while genotype g_{64} has low proliferation rate. Arrows in Figure 5.2 indicates the efficiency of the drugs: for instance, the genotypes $g_1, g_5, \dots, g_{61}$ are all on the same face of the lattice and are fully susceptible to therapy 1, while the opposite face, $g_4, g_8, \dots, g_{64}$ describes all genotypes highly resistant to therapy 1.

Finally, the fitness of genotype i, f_i, is a decreasing factor owing to the fact that, in the absence of treatment, mutation reduces the fitness of the genotype.

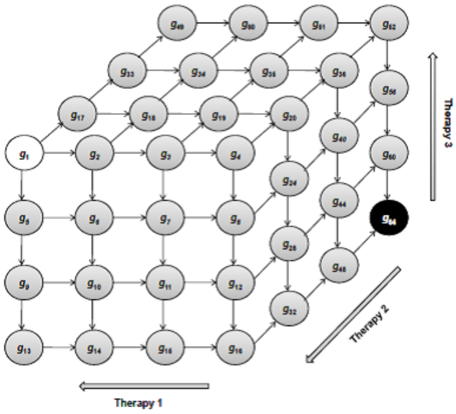

Figure 5.2: Mutation tree for the 64 variant, 3 drug combination model.

5.2.4 Model's discretization

Analogously to what done for the 4 variant, 2 drug combination model, and for the same reason, it is convenient to introduce the discrete-time counterpart of model (5.8). Denoting by τ the sampling interval, the DPSS is given by:

$$\mathbf{x}(k+1) = A_{\sigma(k)}\mathbf{x}(k), \tag{5.9}$$

where $\mathbf{x}(k) = \mathbf{x}(k\tau)$ is the sampled state and the positive matrix $A_{\sigma(k)}$ is given by:

$$A_{\sigma(k)} := e^{(diag\{\Lambda_{i,\sigma(t)}\}+\mu TM)\tau}.$$

Chapter 6

Switching Strategies to Mitigate HIV Escape

In the present chapter we address the problem of Highly Active Antiretroviral Therapy (HAART) scheduling using a control theoretic approach. First of all, the optimal control problem is introduced along with its great criticality: computing the optimal solution is a hard task to be accomplished even numerically. For this reason, looking for suboptimal solutions may be convenient and, in this regard, two different approaches are considered. Firstly, switching strategies designed by means of linear CLFs are tested using the 4 variant, 2 drug combination model. Secondly, the Model Predictive Control approach is briefly described along with simulation results performed on the 64 variant, 3 drug combination model. As we will see, the remarkable fact is that both approaches guarantee performances that are not far away from the solution of the optimal control problem.

6.1 Problem formulation

As already seen in the previous chapters, current HAART regimens are only able to partially and temporarily halt the HIV replication: not only they reduce the growth of certain viral populations while leaving that of others unchanged, but there also exist viral variants resistant to all antiretroviral drugs currently in use.

Furthermore, biological reasons ensure that, if the total viral load is small enough during a finite time of treatment, then there is a significant probability that the total virus load becomes zero and stays at zero. Hence, finding the best time to change HAART therapy (namely, finding a suitable switching rule) in order to maintain the total viral load at low levels is of great importance for patient treatment.

Formulation as an optimal control problem is straightforward when considering the total viral load as the cost function to be minimized. In particular, if we consider the 4 variant, 2 drug combination model (5.4), recalling that each state variable represents a different genotype, it is logical to define a cost function over an infinite, or finite, time horizon respectively as:

$$J_\infty := \sum_{k=0}^{\infty} \mathbf{1}_4^T \hat{\mathbf{x}}(k) \tag{6.1}$$

$$J_T := \mathbf{1}_4 \hat{\mathbf{x}}(T), \tag{6.2}$$

where T is an appropriate final time.

Analogously, if we consider the 64 variant, 3 drug combination model (5.9) and we recall that for each mutant i the state vector is defined as:

$$\mathbf{x}_i(t) = \begin{bmatrix} T_i^*(t) & L_i(t) & V_i(t) \end{bmatrix}^T,$$

the cost function over finite horizon takes the form:

$$J_T := \mathbf{c}\mathbf{x}(T), \tag{6.3}$$

where $\mathbf{c} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 & \dots & 0 & 0 & 1 \end{bmatrix}^T$.

Hence, finding the optimal alternation of HAART regimens means finding the optimal switching law that minimizes one of the cost functions just introduced.

6.2 Optimal control problem

Before considering the 4 variant, 2 drug combination model, let us formulate the problem in more general terms. To this aim, consider a DPSS of the form (1.1) switching among p subsystems and the cost function to be minimized:

$$J_T = \mathbf{c}^T \mathbf{x}(T) + \sum_{k=0}^{T-1} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k).$$

Notice that each i-th subsystem is associated with a (possibly) different weighting vector $\mathbf{q}_i$, which means that when the i-th subsystem is active, then the penalty on the vector state is equal to $\mathbf{q}_i$. The vector $\mathbf{c}$, instead, provides weighing to the final state reached by the system.

The first approach we might think of to compute the optimal solution is a "brute force" algorithm which analyzes all possible combinations for the switching sequence. Assuming the initial state $\mathbf{x}(0)$ to be given and considering a decision time equal to τ, during which treatment is fixed, we can proceed as follows:

- we compute the set of all states that can be reached by the system at any time instant;
- we evaluate $p^{\frac{T}{\tau}}$ possible switching sequences and we determine which one minimizes the performance criterion.

Clearly, this exhaustive search approach becomes computationally prohibitive as soon as the final time T grows up.

An alternative, less computationally demanding, way to calculate the optimal switching signal $\sigma^o(\cdot)$ and the corresponding trajectory $\mathbf{x}^o(\cdot)$ is the so called "dynamic programming"[1] approach. This approach is based on the principle of optimality formulated by Bellman and stated in Proposition 6.1 below.

Proposition 6.1 (BELLMAN'S PRINCIPLE OF OPTIMALITY)**.** Let $\sigma^o(\cdot)$ and $\mathbf{x}^o(\cdot)$ be respectively the optimal strategy and the optimal trajectory corresponding to the initial state $\mathbf{x}(0) = \mathbf{x}_0$ and the time interval $[0, T]$. Then, for any arbitrary instant $\bar{t}$ in $[0, T]$ the switching strategy $\sigma^o(\bar{t}), \sigma^o(\bar{t}+1), \ldots, \sigma^o(T-1)$ represents the optimal strategy corresponding to the initial state $\mathbf{x}^o(\bar{t})$ and the time interval $[\bar{t}, T]$.

The proof of this principle is rather straightforward: suppose that the optimal strategy, when starting from $\mathbf{x}^o(\bar{t})$, is $\bar{\sigma}(\bar{t}), \bar{\sigma}(\bar{t}+1), \ldots, \bar{\sigma}(T-1)$ and apply the switching sequence:

$$\sigma^o(0), \sigma^o(1), \ldots, \sigma^o(\bar{t}-1), \bar{\sigma}(\bar{t}), \bar{\sigma}(\bar{t}+1), \ldots, \bar{\sigma}(T-1);$$

then, the cost function J_T would assume a smaller value than J_T^o, thus contradicting the hypothesis of optimality on $\sigma^o(\cdot)$.

The characteristic feature of the dynamic programming approach is the fact that we proceed backward. Intuitively, the underlying idea is the following:

- if we suppose we have somehow determined $\mathbf{x}^o(T-1)$, the remaining decision $\sigma(T-1)$ which minimizes the cost function over the interval $[T-1, T]$ coincides with $\sigma^o(T-1)$;
- then, we can go one step backward: assuming the state $\mathbf{x}^o(T-2)$ to be given, the remaining decision $\sigma(T-2)$ minimizing the cost function over the interval $[T-2, T]$ coincides with the value of the optimal law $\sigma^o(T-2)$;
- we can go on moving backward until we reach the initial condition $\mathbf{x}(0)$ and hence the optimal value $\sigma^o(0)$.

With no pretense of rigorousness, the above reasoning can be put into mathematical terms as follows: the optimal trajectory $\mathbf{x}^o(\cdot)$ and the optimal switching signal $\sigma^o(\cdot)$ satisfy the following difference equation, known as the Hamilton-Jacobi-Bellman equation[2]:

$$H(\mathbf{x}^o(T), T) = \mathbf{1}_n^T \mathbf{x}^o(T) \tag{6.4}$$

$$H(\mathbf{x}^o(k), k) = \min_{i=1,\ldots,p} \{H(\mathbf{x}^o(k+1), k+1) + \mathbf{q}_i \mathbf{x}(k)\}, \; k = 0, \ldots, T-1, \tag{6.5}$$

where $\mathbf{q}_i = \mathbf{q}_{\sigma^o(k)}$.

Obviously, the optimal trajectory $\mathbf{x}^o(\cdot)$ must satisfy the further condition:

$$\mathbf{x}(0) = \mathbf{x}_0, \tag{6.6}$$

$$\mathbf{x}^o(k+1) = A_{\sigma^o(k)} \mathbf{x}^o(k), \; k = 0, \ldots, T-1. \tag{6.7}$$

[1] Dynamic programming algorithms solve problems by combining the solutions to subproblems that are not disjoint but rather overlapping (namely, they share "subsubproblems"). A typical dynamic programming algorithm recursively defines the value of an optimal solution and then constructs such an optimal solution in a bottom-up fashion.

[2] In this context our aim is to provide only a brief introduction, at a very intuitive level, to the Hamilton-Jacobi-Bellman equation. Deeper insights into its solution, along with a more rigorous formulation of the problem, can be found for instance in [21],[3].

Hence, using equations (6.4), (6.5), (6.6) and (6.7), we obtain the following system:

$$\begin{cases} H(\mathbf{x}^o(k), k) = \min\limits_{i=1,\dots,p} \{H(\mathbf{x}^o(k+1), k+1) + \mathbf{q}_i\mathbf{x}(k)\} \\ \mathbf{x}^o(k+1) = A_{\sigma^o(k)}\mathbf{x}^o(k) \end{cases} \tag{6.8}$$

with boundary conditions:

$$\begin{cases} H(\mathbf{x}^o(T), T) = \mathbf{1}_n^T\mathbf{x}^o(T) \\ \mathbf{x}(0) = \mathbf{x}_0. \end{cases} \tag{6.9}$$

Notice that the state equation (6.7) with initial condition (6.6) must be integrated forward, whereas equation (6.5), initialized by condition (6.4), must be integrated backward, both according to the coupling condition provided by the switching rule $\sigma^o(\cdot)$. In this sense the optimal control problem (6.8)-(6.9) is a two point boundary value problem, and cannot be solved using regular integration techniques.

Even if a number of algorithms have been proposed to solve the system (6.8)-(6.9), and hence compute the optimal solution, we will present only the simplest and more intuitive procedure. Consider the case where $\mathbf{q}_i = 0$ for all $i = 1, \dots, p$, namely there is a terminal cost only. Note that this does not entail loss of generality as we can always introduce a new variable $y(k)$, having equation $y(k+1) = y(k) + \mathbf{q}_{\sigma(k)}\mathbf{x}(k)$ and initial condition $y(0) = 0$, so that:

$$J_T = \mathbf{c}^T\mathbf{x}(T) + y(T).$$

Now, define recursively the sequence of matrices:

$$\begin{aligned} \Omega_0 &= \mathbf{c} \\ \Omega_1 &= \begin{bmatrix} A_1^T\Omega_0 & A_2^T\Omega_0 & \dots & A_p^T\Omega_0 \end{bmatrix} = \begin{bmatrix} A_1^T\mathbf{c} & A_2^T\mathbf{c} & \dots & A_p^T\mathbf{c} \end{bmatrix} \\ &\vdots \\ \Omega_{k+1} &= \begin{bmatrix} A_1^T\Omega_k & A_2^T\Omega_k & \dots & A_p^T\Omega_k \end{bmatrix} \end{aligned}$$

Then, we can compute the optimal switching strategy and the optimal trajectory as:

$$\begin{aligned} \sigma^o(0) &= \operatorname*{argmin}_j \, (\Omega_T^T)^{(j)}\mathbf{x}(0) \\ \mathbf{x}^o(1) &= A_{\sigma^o(0)}\mathbf{x}(0) \\ \sigma^o(1) &= \operatorname*{argmin}_j \, (\Omega_{T-1}^T)^{(j)}\mathbf{x}^o(1) \\ &\vdots \\ \sigma^o(k) &= \operatorname*{argmin}_j \, (\Omega_{T-k}^T)^{(j)}\mathbf{x}^o(k) \\ \mathbf{x}^o(k+1) &= A_{\sigma^o(k)}\mathbf{x}(k) \\ &\vdots \\ \sigma^o(T-1) &= \operatorname*{argmin}_j \, (\Omega_1^T)^{(j)}\mathbf{x}^o(T-1) \\ \mathbf{x}^o(T) &= A_{\sigma^o(T-1)}\mathbf{x}(T-1) \end{aligned}$$

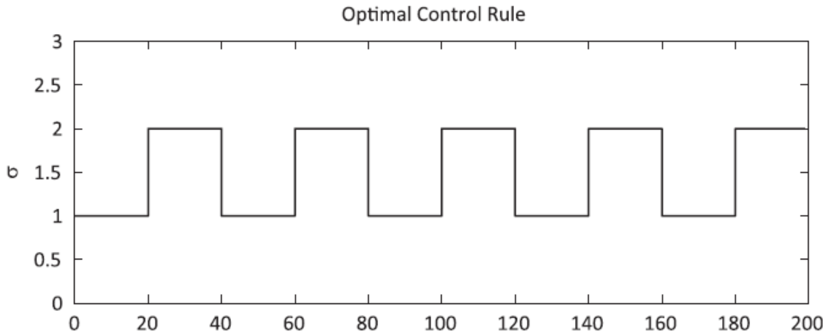

Figure 6.1: Optimal control rule.

Notice that, even if we build up the optimal sequence onwards, we are still resorting to Bellman's principle of optimality, as at each step $\bar{t} \in [0, T]$, the optimal value $\sigma^o(\bar{t})$ is computed as the optimal strategy corresponding to the initial state $\mathbf{x}^o(\bar{t})$ and the time interval $[\bar{t}, T]$.

The implementation of the previous strategy requires storing the columns of Ω^T_{T-k} whose number would be $1 + p + p^2 + p^3 + \cdots + p^T$. Clearly, this exponential growth could be too computationally demanding. In general, many of the columns of the matrices Ω^T_{T-k} may be redundant and can be removed. to accomplish this goal, Hernandez-Vargas and coauthors in [14] provide a number of algorithms based on linear programming techniques, whose presentation certainly goes beyond our objectives.

Finally, we go back to the problem of treatment scheduling and present results obtained by Hernandez-Vargas and coauthors in [11]. They consider the 4 variant, 2 drug combination model given by equation (5.3) and the cost function:

$$J_T := \mathbf{1}_4^T \mathbf{x}(T).$$

Furthermore, they consider a period of $T = 200$ days and take the decision time τ equal to 20 days (this is consistent with the fact that clinical visits during HIV treatment typically have a frequency of once a month or less).

Simulation results [11] can be summarized as follows:

- The optimal control rule is shown in Figure 6.1: we can notice that the switching among the two therapies occurs quite regularly, approximately every 20 days;

- The total viral load at the end of treatment is equal to 664.99.

For comparison purposes, authors also perform a simulation using a single therapy. As we might expect, in this case the total viral load at the end of the treatment is much greater than 664.99, and, indeed, it is equal to 3.05×10^3. This dramatic difference reflects clinical data suggesting the proactive alternation of HAART regimens.

6.3 Bounds to the optimal cost

We have seen in the previous section that computing the solution to the optimal control problem might be really computationally demanding, or even prohibitive, depending on the length of the treatment period T. For this reason it is convenient to look for alternative strategies, namely suboptimal strategies.

The aim of the present section is to present results derived in [11] providing upper bounds on the performance of the optimal feedback strategy, both in the finite- and in the infinite-time horizon case. It is worthwhile to underline that these results are not only interesting by themselves, but they also provide justification for testing certain switching strategies.

6.3.1 Infinite time horizon case

Consider a DPSS of the form (1.1) and the following cost function over an infinite time horizon:

$$J_\infty := \sum_{k=0}^{\infty} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k).$$

Recall that, as seen in Chapter 3, section 3.2, $\mathscr{N}$ denotes the set of all Metzler matrices $\mathbb{N} \in \mathbb{R}^{p\times p}$ with elements ν_{ij}, such that:

$$\nu_{ij} \geq 0, \forall i \neq j,\ \sum_{i=1}^{p} \nu_{ij} = 0, \forall j.$$

Then, an upper bound on the optimal value J_∞^o of J_∞ can be derived, as shown in Proposition 6.2 below.

Proposition 6.2. Let $A_1, \dots, A_p \in \mathbb{R}_+^{n\times n}$ be the subsystems' state matrices of a DPSS (1.1). Let $\mathbf{q}_1, \dots, \mathbf{q}_p$, $\mathbf{q}_i \in \mathbb{R}_+^n$, be a given set of p positive vectors and assume that there exist $\mathbb{N} \in \mathscr{N}$ and a set of p positive vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$, $\mathbf{v}_i \in \mathbb{R}_+^n$, satisfying the copositive Lyapunov inequalities:

$$(A_i - I_n)^T \mathbf{v}_i + \sum_{j=1}^{p} \nu_{ji} \mathbf{v}_j + \mathbf{q}_i < 0, \quad \forall i \in [1, p]. \tag{6.10}$$

Apply the state-feedback switching rule:

$$\sigma(\mathbf{x}(t)) := \operatorname*{argmin}_{i=1,\dots,p} \mathbf{v}_i^T \mathbf{x}(t) \tag{6.11}$$

and consider the cost function:

$$J_\infty = \sum_{k=0}^{\infty} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k).$$

Then, the following bound on the cost function holds:

$$J_\infty \leq \min_{i=1,\dots,p} \mathbf{v}_i^T \mathbf{x}_0.$$

Proof. If inequalities (6.10) hold, then, by Theorem 3.1, the switching rule (6.11) stabilizes the DPSS. In addition, if we consider the Piecewise Linear CLF:

$$V(\mathbf{x}(k)) = \min_{i=1,\dots,p} \mathbf{v}_i^T \mathbf{x}(k),$$

then, as seen in the proof of Theorem 3.1, we have:

$$\Delta V(\mathbf{x}(k)) \leq \mathbf{v}_{\sigma(k)}^T (A_{\sigma(k)} - I_n)\mathbf{x}(k). \tag{6.12}$$

Now, notice that the Lyapunov inequality (6.10) for $i = \sigma(k)$ can be rewritten as:

$$(A_{\sigma(k)} - I_n)^T \mathbf{v}_{\sigma(k)} < -\sum_{j=1}^{p} \nu_{j\sigma(k)} \mathbf{v}_j - \mathbf{q}_{\sigma(k)},$$

or, equivalently, as:

$$\mathbf{v}_{\sigma(k)}^T (A_{\sigma(k)} - I_n) < -\sum_{j=1}^{p} \nu_{j\sigma(k)} \mathbf{v}_j^T - \mathbf{q}_{\sigma(k)}^T. \tag{6.13}$$

From (6.12) and (6.13), with $\mathbf{x}(k) \neq 0$, it follows:

$$\begin{aligned}
\Delta V(\mathbf{x}(k)) &< -\sum_{j=1}^{p} \nu_{j\sigma(k)} \mathbf{v}_j^T \mathbf{x}(k) - \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k) \\
&\leq -\left(\sum_{j=1}^{p} \nu_{j\sigma(k)}\right) \mathbf{v}_{\sigma(k)}^T \mathbf{x}(k) - \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k) \\
&= -\mathbf{q}_{\sigma(k)}^T \mathbf{x}(k).
\end{aligned}$$

By summing up on both sides, for each k from zero to infinity, we have:

$$\sum_{k=0}^{\infty} \Delta V(\mathbf{x}(k)) \leq -\sum_{k=0}^{\infty} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k),$$

or, equivalently:

$$\sum_{k=0}^{\infty} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k) \leq V(\mathbf{x}(0)) - V(\mathbf{x}(\infty)). \tag{6.14}$$

Now, recall that the state evolution asymptotically converges to zero, and hence $V(\mathbf{x}(\infty)) = 0$. From (6.14), this implies that:

$$J_\infty = \sum_{k=0}^{\infty} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k) \leq \min_{i=1,\dots,p} \mathbf{v}_i^T \mathbf{x}_0.$$

□

6.3.2 Finite time horizon case

Consider a DPSS of the form (1.1) and the following cost function already introduced in section 6.2 when dealing with the optimal control problem:

$$J_T := \mathbf{c}^T \mathbf{x}(T) + \sum_{k=0}^{T-1} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k).$$

In order to compute an upper bound on the optimal value J_T^o of J_T we need to slightly modify the relevant inequalities considered in the previous subsection. In this case, however, instead of a set of inequalities, we need to deal with a set of difference equations and a set of vectors $\mathbf{v}_i(k)$ which depend on the time instant k.

Proposition 6.3. Let $A_1, \dots, A_p \in \mathbb{R}_+^{n\times n}$ be the subsystems' state matrices of a DPSS (1.1). Let $\mathbf{q}_1, \dots, \mathbf{q}_p$, $\mathbf{q}_i \in \mathbb{R}_+^n$, be a given set of p positive vectors and assume that there exist $\mathbb{N} \in \mathscr{N}$ and a set of $p \cdot T$ positive vectors $\{\mathbf{v}_1(k), \dots, \mathbf{v}_p(k)\}$, $\mathbf{v}_i(k) \in \mathbb{R}_+^n$, $k = 0, \dots, T-1$, satisfying the difference equations:

$$\mathbf{v}_i(k) = A_i^T \mathbf{v}_i(k+1) + \sum_{j=1}^{p} \nu_{ji} \mathbf{v}_j(k) + \mathbf{q}_i, \ \mathbf{v}_i(T) = \mathbf{c}, \tag{6.15}$$

for $i = 1, \dots, p$. Apply the state-feedback switching rule:

$$\sigma(\mathbf{x}(k)) := \underset{i=1,\dots,p}{\operatorname{argmin}} \ \mathbf{v}_i^T(k)\mathbf{x}(k)$$

and consider the cost function:

$$J_T = \mathbf{c}^T \mathbf{x}(T) + \sum_{k=0}^{T-1} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k).$$

Then, the following bound on the cost function holds:

$$J_T \le \min_{i=1,\dots,p} \ \mathbf{v}_i^T(0)\mathbf{x}_0.$$

Proof. Consider the function:

$$V(\mathbf{x}(k), k) = \min_{i=1,\dots,p} \ \mathbf{v}_i^T(k)\mathbf{x}(k)$$

Then, we have:

$$\begin{aligned} V(\mathbf{x}(k+1), k+1) &= \min_{i=1,\dots,p} \ \mathbf{v}_i^T(k+1)\mathbf{x}(k+1) \\ &= \min_{i=1,\dots,p} \ \mathbf{v}_i^T(k+1) A_{\sigma(k)} \mathbf{x}(k) \\ &\le \mathbf{v}_{\sigma(k)}^T(k+1) A_{\sigma(k)} \mathbf{x}(k) \end{aligned} \tag{6.16}$$

Now, from the difference equation (6.15) for $i = \sigma(k)$ we have:

$$A_{\sigma(k)}^T \mathbf{v}_{\sigma(k)}(k+1) = \mathbf{v}_{\sigma(k)}(k) - \sum_{j=1}^{p} \nu_{j\sigma(k)} \mathbf{v}_j(k) - \mathbf{q}_{\sigma(k)},$$

or, equivalently:

$$\mathbf{v}_{\sigma(k)}^T(k+1) A_{\sigma(k)} = \mathbf{v}_{\sigma(k)}^T(k) - \sum_{j=1}^{p} \nu_{j\sigma(k)} \mathbf{v}_j^T(k) - \mathbf{q}_{\sigma(k)}^T. \tag{6.17}$$

Hence, from (6.16) and (6.17), it follows:

$$\begin{aligned}
V(\mathbf{x}(k+1), k+1) &\leq V(\mathbf{x}(k), k) - \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k) - \sum_{j=1}^{p} \nu_{j\sigma(k)} \mathbf{v}_j^T(k)\mathbf{x}(k) \\
&\leq V(\mathbf{x}(k), k) - \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k) - \mathbf{v}_{\sigma(k)}^T(k)\mathbf{x}(k) \sum_{j=1}^{p} \nu_{j\sigma(k)} \\
&= V(\mathbf{x}(k), k) - \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k),
\end{aligned}$$

which amounts to:

$$V(\mathbf{x}(k+1), k+1) - V(\mathbf{x}(k), k) \leq -\mathbf{q}_{\sigma(k)}^T \mathbf{x}(k).$$

By summing up on both sides, for each k from zero to $T-1$, we have:

$$\sum_{k=0}^{T-1} V(\mathbf{x}(k+1), k+1) - V(\mathbf{x}(k), k) \leq -\sum_{k=0}^{T-1} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k),$$

so that:

$$\begin{aligned}
J_T &= \mathbf{c}^T \mathbf{x}(T) + \sum_{k=0}^{T-1} \mathbf{q}_{\sigma(k)}^T \mathbf{x}(k) \\
&\leq \mathbf{c}^T \mathbf{x}(T) - \sum_{k=0}^{T-1} V(\mathbf{x}(k+1), k+1) - V(\mathbf{x}(k), k) \\
&= \mathbf{c}^T \mathbf{x}(T) - V(\mathbf{x}(T), T) + V(\mathbf{x}_0, 0) \\
&= \min_{i=1,\ldots,p} \mathbf{v}_i^T(0)\mathbf{x}_0
\end{aligned}$$

□

It is worth noticing that, differently from the infinite horizon case where conditions (6.10) may be infeasible, in the finite horizon case equations (6.15) are always feasible. Indeed, once we choose the matrix $\mathbb{N} \in \mathscr{N}$ as the zero-matrix (and hence $\nu_{ij} = 0$ for every i, j), we only need to solve backward the difference equation (6.15).

6.4 Guaranteed cost control over infinite horizon

We consider the cost function over infinite horizon (6.1):

$$J_\infty := \sum_{k=0}^{\infty} \mathbf{1}_4^T \hat{\mathbf{x}}(k)$$

and the 4 variant, 2 drug combination model given by (5.4), where we choose β so as to ensure that all matrices $\hat{A}_\sigma$ are Schur.

Now, consider the copositive Lyapunov inequalities (6.10) with vectors $\mathbf{q}_i = \mathbf{1}_4$ for all i. It is straightforward to see that the subsystems' stability assumption ensures that such inequalities are always feasible, as we only need to choose

$\mathbb{N} \in \mathscr{N}$ equal to the zero-matrix (namely $\nu_{ij} = 0$ for every i, j) and vectors $\mathbf{v}_i$ given by:

$$\mathbf{v}_i = -(\hat{A}_i^T - I_n)^{-1}\mathbf{q}. \tag{6.18}$$

Notice that, since $\hat{A}_i$ is a Schur matrix, $\hat{A}_i^T - I_n$ is a Metzler Hurwitz matrix, and hence it is non-singular and its inverse is a negative matrix, so that vectors $\mathbf{v}_i$ in (6.18) are positive.

Finally, we consider the switching rule:

$$\sigma(\hat{\mathbf{x}}(k)) := \underset{i=1,\ldots,p}{\operatorname{argmin}} \ \mathbf{v}_i^T \hat{\mathbf{x}}(k). \tag{6.19}$$

By applying Proposition 6.2, we can claim that such switching rule guarantees an upper bound on the cost function:

$$J_\infty \leq \min_{i=1,\ldots,p} \mathbf{v}_i^T \hat{\mathbf{x}}_0.$$

As for the optimal control case, we take a decision time equal to 20 days and a period of treatment equal to 200 days. Simulation results [11] can be summarized as follows:

- The dynamic evolution of the state variables (namely, of the 4 viral variants) is illustrated in Figure 6.2: we can notice that, for an initial period of time, the switching rule (6.19) is able to maintain a low Wild Type concentration and to suppress the concentrations of Genotype 1 and 2. However, as we might expect, the Highly Resistant Genotype eventually grows since none of the therapies affect this genotype;
- The switching rule given by (6.19) is illustrated in Figure 6.3: as for the optimal control case, it shows a certain periodicity (and again the period is approximately equal to 20 days);
- The total viral load at the end of treatment is equal to 664.99: quite surprisingly, this value coincides with the optimal performance previously derived.

6.5 Guaranteed cost control over finite horizon

We consider the cost function over finite horizon (6.2):

$$J_T := \mathbf{1}_4^T \hat{\mathbf{x}}(T).$$

Differently from the previous case, the condition that all matrices $\hat{A}_i$ are Schur matrices can be removed. Therefore, we consider the 4 variant, 2 drug combination model (5.4) with $\beta = 0$ (which implies $\hat{\mathbf{x}}(k) = \mathbf{x}(k)$ for all k).

Now, we can solve backward in time the difference equations (6.15) with vectors $\mathbf{q}_i = \mathbf{1}_4$ for all i: to this aim we choose $\mathbb{N} \in \mathscr{N}$ equal to the zero-matrix and start from the final condition $\mathbf{v}_i(T) = \mathbf{1}_4$ for all i.

Finally, we consider the switching rule:

$$\sigma(\hat{\mathbf{x}}(k)) := \underset{i=1,\ldots,p}{\operatorname{argmin}} \ \mathbf{v}_i^T(k) \hat{\mathbf{x}}(k). \tag{6.20}$$

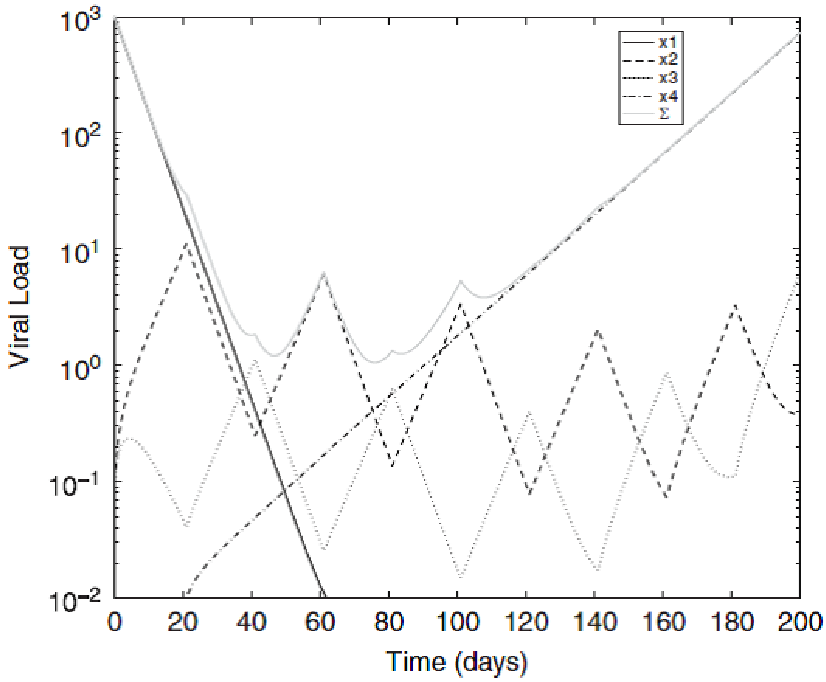

Figure 6.2: Dynamic evolution of the state variables using the guaranteed cost control over infinite horizon.

Table 6.1: Total viral load at the end of treatment using different control rules.

Control rule	Performance
Single therapy	3.05×10^{13}
Optimal control	664.99
Guaranteed cost control over infinite horizon	664.99
Guaranteed cost control over finite horizon	664.99

By applying Proposition 6.3, we can claim that this switching rule guarantees an upper bound on the cost function:

$$J_T \leq \min_{i=1,\ldots,p} \mathbf{v}_i^T(0)\mathbf{x}_0.$$

Again, we take a decision time equal to 20 days and a period of treatment equal to 200 days. Simulation results shows that the control rule (6.20) has the same performance obtained in the infinite time horizon case, namely the total viral load concentration at the end of the treatment is equal to 664.99.

6.6 Comparison

All results presented so far are summarized in Table 6.1.

Fist of all, an important comment about suboptimal strategies should be make: even if in this particular case both guaranteed cost controls have the

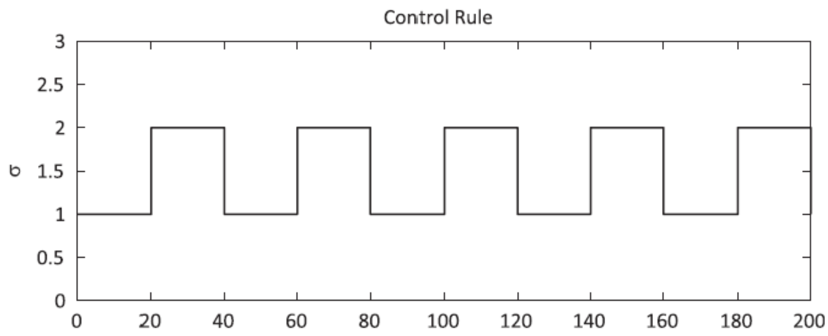

Figure 6.3: Guaranteed cost control over infinite horizon.

same performance, this is not a general result, but rather a consequence of the symmetry of the replication rates values (see Table 5.1 in Chapter 5).

Comparing the performance of the optimal switching law with other suboptimal strategies, we see that they all give the same total virus load at the end of the treatment. What is more, Hernandez-Vargas and coauthors claim that guaranteed cost controls have the same performances as the optimal rule even if different initial conditions are considered. This is certainly a very interesting, as well as surprising, result, since computing suboptimal strategies is far less computationally demanding that computing the optimal rule. On the other hand, there is no proof that any of the guaranteed cost control is the same as the optimal control.

Finally, it is worth saying that periodicity of the switching law previously noticed might be a consequence of the symmetry in the replication rates.

6.7 Model Predictive Control Approach

A completely different approach for determining near optimal switching drug schedules is Model Predictive Control (MPC). Generally speaking, MPC appears to be suitable for applications to the biomedical area, owing to its robustness to disturbances, model uncertainties and the possibility of handling constraints. For these reasons, also in the context of HIV treatment scheduling the MPC approach has been addressed [10],[39],[32],[13].

In what follows, we will first present the basic idea of MPC, and later on, without being concerned with its rigorous mathematical formulation nor with the computation of a numerical solution, we will present results derived by Hernandez-Vargas and coauthors [10] using the 64 variant, 3 drug combination model introduced in Chapter 5.

The original feature of MPC approach is the fact that the optimization procedure resorts to predictions based on a model (hence the name Model Predictive Control). To be more precise, the basic idea is the following (see Figure 6.4):

- At time instant t the controller, using measurements just collected, predicts the future dynamic behavior of the system over a prediction horizon T_p and computes the optimal control rule over a control horizon T_c;

- Only the first input of the optimal control sequence is applied (the re-

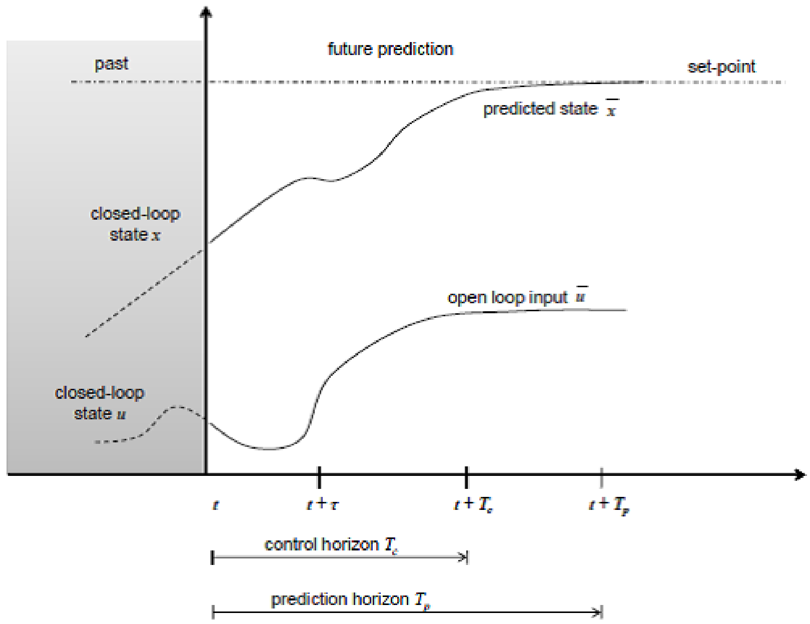

Figure 6.4: Model Predictive Control approach.

maining inputs are disregarded and will not be taken into account any more);

- At time instant $t+\tau$ new measurements are available and the whole procedure, namely prediction and optimization, can be reiterated to find a new input function.

We can think of the prediction and the control horizons as two sliding windows moving forward: the former defines how far into the future the controller predicts the state evolution, while the latter defines how far into the future it plans the control action.

It is worth noticing that, due to disturbances, measurement noise and model-plant mismatch, the actual system behavior is most likely to be different from the predicted one. However, implementing only the first step of the optimal sequence (and then performing again the prediction and optimization procedures), means incorporating a useful feedback mechanism.

As far as the prediction horizon T_p is concerned, an interesting consideration should be done. This parameter, indeed, plays a very important role in the performance of the MPC schemes: on the one hand, it is desirable to use short prediction horizons for computational reasons (the shortest T_p, the less costly the solution of the optimization problem); on the other hand, choosing too short prediction horizons leads to optimal control rule which differ significantly from the optimal solution over infinite horizon.

As already mentioned above, we can now present results derived in [10] using the 64 variant, 3 drug combination model. In order to be consistent with application in a clinical setting, authors fixed the decision time τ to 90 days and decided not to take measurements at faster intervals.

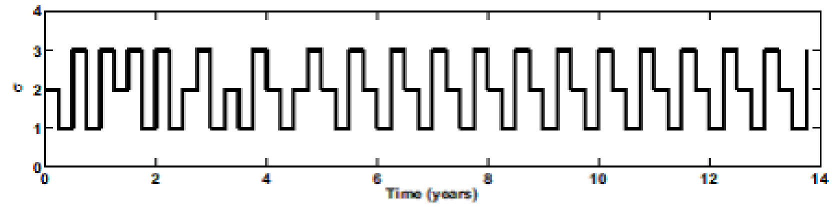

Figure 6.5: Switching control rule given by the MPC approach.

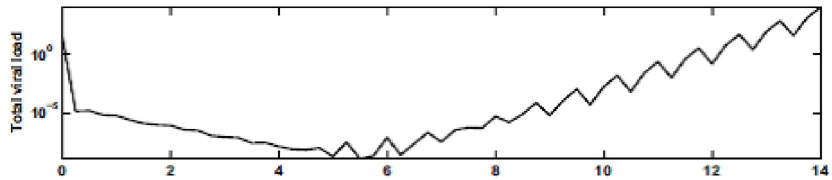

Figure 6.6: Total viral load given by the MPC approach.

Simulation results over a period of 14 years and using a 10 months long prediction horizon, can be summarized as follows:

- The switching rule is illustrated in Figure 6.5: we can notice that it is irregular for the first 4 years, but it shows a quite regular behavior for the remaining years;

- Time trend of the total viral load is illustrated in Figure 6.6: it can be seen that the critical value of 1000 copies/ml (defining the so called "virologic failure") is reached after 12 years.

For comparison purposes, in the same work [10] a simulation study over a period of 4 years is performed applying the strategy commonly used in clinical practice, known as "Virologic Failure Treatment Strategy". This switching rule and the corresponding trend of the total viral load are reported in Figures 6.7 and 6.8 respectively: notice that the viral load starts to escape after 4 years only. This means that the switching rule obtained by means of MPC algorithm can extend the virologic failure for 8 years more compare to the clinical assessment.

Finally, Hernandez-Vargas and coauthors resorted also to a "brute force" approach to compute the optimal control law over a period of 5 years (recall that the evaluation of $3^{\frac{T}{\tau}}$ possible combinations is required and hence high computa-

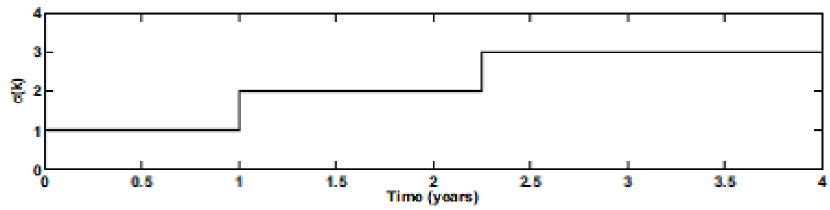

Figure 6.7: Switching control rule given by the Virologic Failure Treatment Strategy.

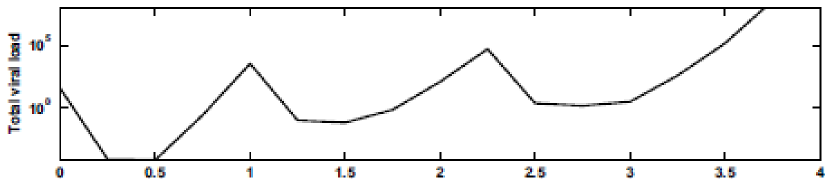

Figure 6.8: Total viral load given by the Virologic Failure Treatment Strategy.

Table 6.2: Total viral load at the end of treatment using different control strategies.

Control strategy	Performance
Model Predictive Control	2.31×10^{-9}
Virologic Failure Treatment Strategy	5.14×10^{28}

tional resources are needed). Simulation results for a 5 years period treatment are summarized in Table 6.2: interestingly, the difference on the viral load obtained with the optimal control and the MPC approach is very small. Hence, MPC appears to be an appropriate framework for this problem, as it provides close results to the optimal control problem, while reducing dramatically computational resources.

6.8 Conclusions

Simulation studies presented throughout this chapter establish the importance of alternating HAART regimens: using different drugs at the right moment is of great importance for patient treatment, as it allows to suppress HIV RNA levels maximally and to prevent future selection of resistant mutations. However, finding the optimal switching rule can require high computational resources as the optimal control problem results in a two boundary value problem. An alternative, suboptimal, solution is represented by guaranteed cost controls, which achieve good results compared to the optimal one. A completely different approach, but as promising as the previous one, is MPC, whose performance is similar to the optimal control and has the advantage of being quite robust to model uncertainties.

Of course, all these simulation studies may help to optimally schedule HIV treatment, a problem which is still a point of discussion among clinicians. Nevertheless, we are far away from a full characterization of switching strategies able to forestall drug failure and a number of interesting open problems can be found: on the one hand, the optimal control problem for positive switched system has been only partially explored to date; on the other hand, in order to provide a more extensive description of HIV dynamics, more complicated mutation trees might be taken into consideration, further states for each mutant (such us infected macrophages) might be considered and random variations might be introduced. All these considerations make the problem of HIV treatment scheduling an open and active field of research.

Appendix A

Positive systems

The purpose of this appendix is to introduce some essential notions on positive systems along with some basic properties of positive matrices and Metzler matrices. Not aiming for a complete characterization of the subject, we will limit us to those remarkable facts which are of relevance to the comprehension of results provided in the present work. For the sake of conciseness, we have chosen to omit (almost) all proofs and to refer the reader to specialized literature on the topic.

By a positive system we mean a state-space model having the peculiar property that its state trajectories and outputs always remain non-negative for any non-negative initial condition and any non-negative input. It is worth noticing that the positivity constrain is commonly encountered in many applications: a variety of models having positive linear behavior can be found in Engineering, Management, Science, Economics, Social Sciences, Biology and Medicine.

As it will be clearer later, the theory of positive systems is firmly based on positive matrices for the discrete-time case and on Metzler matrices for the continuous-time case. For this reason, in the following, we will first provide insights into these classes of matrices and then we will move on to positive systems.

A.1 Positive matrices

In this section square positive matrices, namely square matrices whose elements are all non-negative, are considered. In particular, we discuss their classification into irreducible matrices and reducible matrices and we provide insights into their spectral properties.

Definition A.1. A real positive matrix $A \in \mathbb{R}_+^{n\times n}$ is called:

- *primitive* if there exists an integer $h > 0$ such that $A^h \gg 0$, namely:
$$\left[A^h\right]_{r,s} > 0,\ \forall r, s \in [1, n];$$
the smallest h such that $A^h \gg 0$ is called "exponent of primitivity";
- *irreducible* if for any pair $(r, s), 1 \leq r, s \leq n$, there exists an integer $h > 0$ (in general depending on r and s) such that:
$$\left[A^h\right]_{r,s} > 0;$$

- *reducible* if there exist r and s such that:

$$\left[A^h\right]_{r,s} = 0, \ \forall h \geq 0.$$

Notice that strictly positive square matrices are primitive matrices with exponent of primitivity equal to 1, while primitive matrices are a subset of irreducible matrices. In addition, the set of reducible matrices complements those of irreducible matrices within the whole set of positive matrices.

As far as spectral properties are concerned, positive matrix theory takes the famous Perron-Frobenius theorem as its foundation: Perron proved it in 1907 for strictly positive matrices, while Frobenius provided its extension to irreducible matrices in 1912.

Theorem A.1 (Perron-Frobenius theorem for irreducible matrices)**.** Let $A \in \mathbb{R}_+^{n \times n}$ be an irreducible matrix. Then, the following properties hold:

(i) Existence of a strictly positive eigenvalue and a strictly positive eigenvector: there exist a real number $\lambda_0 > 0$ and a vector $\mathbf{v}_0 \gg 0$ such that:

$$A\mathbf{v}_0 = \lambda_0 \mathbf{v}_0;$$

(ii) Maximality of λ_0: the eigenvalue λ_0 is maximal in modulus among all the eigenvalues of A, namely $|\lambda| \leq \lambda_0$ for any eigenvalue $\lambda \in \Lambda(A)$ (and hence λ_0 is the spectral radius of A);

(iii) Structure of the peripheral spectrum: every eigenvalue λ with $|\lambda| = \lambda_0$ is a simple root of the characteristic polynomial of A; in addition, there exists a positive integer η, called "index of imprimitivity", such that:

  - all the eigenvalues of A with $|\lambda| = \lambda_0$ are given by:

$$\lambda = \lambda_0 e^{j\frac{2\pi}{\eta}k}, \ k = 0, 1, \ldots, \eta - 1;$$

  - the whole spectrum of A is invariant under a rotation of the complex plane by $\frac{2\pi}{\eta}$ (namely, the whole spectrum of A is invariant with respect to multiplication by $e^{j\frac{2\pi}{\eta}}$);

(iv) Uniqueness of $\mathbf{v}_0$: any positive eigenvector of A corresponding to λ_0 is a multiple of $\mathbf{v}_0$;

(v) Components on the Jordan basis: with respect to the Jordan basis, any positive vector $\mathbf{x} > 0$ has a positive component with respect to $\mathbf{v}_0$;

(vi) Monotonicity of λ_0: if $\bar{A}$ is such that $\bar{A} > A$, its maximal eigenvalue satisfies $\bar{\lambda}_0 > \lambda_0$.

The positive eigenvalue λ_0 and the strictly positive vector $\mathbf{v}_0$ are called "Perron-Frobenius eigenvalue" and "Perron-Frobenius eigenvector".

As already mentioned above, proving just stated properties of irreducible matrices goes beyond our objectives. However, various proofs of the Perron-Frobenius theorem can be found in the literature, see for instance [6],[31].

A further useful characterization of the Perron-Frobenius eigenvalue is provided by the following proposition.

Proposition A.1. Let $A \in \mathbb{R}_+^{n\times n}$ be an irreducible matrix and let λ_0 be its Perron-Frobenius eigenvalue. Then, the following properties hold:

(i) If $\lambda > \lambda_0$, the matrix $(\lambda I_n - A)^{-1}$ is strictly positive;
if $0 \le \lambda < \lambda_0$, the matrix $(\lambda I_n - A)^{-1}$, if it exists, has at least one negative entry;

(ii) For every vector $\mathbf{x} > 0$ and every $\lambda \in \mathbb{R}_+$, we have:

$$\begin{aligned} A\mathbf{x} > \lambda\mathbf{x} &\Rightarrow \lambda_0 > \lambda, \\ A\mathbf{x} < \lambda\mathbf{x} &\Rightarrow \lambda_0 < \lambda; \end{aligned}$$

(iii) λ_0 can be characterized as:

$$\begin{aligned} \lambda_0 &= \sup\{\lambda \in \mathbb{R}_+ \colon A\mathbf{x} \ge \lambda\mathbf{x},\ \text{for some } \mathbf{x} > 0\} \\ &= \inf\{\lambda \in \mathbb{R}_+ \colon A\mathbf{x} \le \lambda\mathbf{x},\ \text{for some } \mathbf{x} > 0\}. \end{aligned}$$

Interestingly enough, some properties stated for irreducible matrices in the previous Perron-Frobenius theorem hold true in a weaker version for positive matrices that are not necessarily irreducible. These results are presented in Theorem A.2 below, but first the normal form of a reducible matrix should be introduced.

Proposition A.2. Let $A \in \mathbb{R}_+^{n\times n}$ be a reducible matrix of dimension $n > 1$. Then, there exists a permutation matrix Π such that $\hat{A} = \Pi^T A \Pi$ is a lower triangular matrix block having the following *normal form*:

$$\hat{A} = \Pi^T A \Pi = \begin{bmatrix} \hat{A}_{1,1} & & & & & & \\ 0 & \hat{A}_{2,2} & & & & & \\ 0 & 0 & \ddots & & & & \\ 0 & 0 & \dots & \hat{A}_{h,h} & & & \\ \star & \star & \dots & \star & \hat{A}_{h+1,h+1} & & \\ \star & \star & \dots & \star & \star & \ddots & \\ \star & \star & \dots & \star & \star & \star & \hat{A}_{k,k} \end{bmatrix} \tag{A.1}$$

where each diagonal block $\hat{A}_{i,i}$, $i = 1, 2, \dots, k > 1$, is an irreducible matrix (possibly the zero matrix of dimension 1). In addition, if $\hat{A}$ is not a diagonal block matrix, namely $h < k$, then for every index $i > h$ there exists $j \ne i$ such that $\hat{A}_{i,j} \ne 0$. The blocks $\hat{A}_{i,i}$, $i = 1, 2, \dots, h$, are called "isolated blocks".

Theorem A.2 (Perron-Frobenius theorem for reducible matrices)**.** Let $A \in \mathbb{R}_+^{n\times n}$ be a reducible matrix. Then, the following properties hold:

(i) Existence of a maximal non-negative eigenvalue with a positive eigenvector: there exist a real number $\lambda_0 \ge 0$ and a vector $\mathbf{v}_0 > 0$ such that:

$$A\mathbf{v}_0 = \lambda_0 \mathbf{v}_0;$$

(ii) Maximality of λ_0: the eigenvalue λ_0 is maximal in modulus among all eigenvalues of A, namely $|\lambda| \le \lambda_0$ for any eigenvalue $\lambda \in \Lambda(A)$;

(iii) STRUCTURE OF THE PERIPHERAL SPECTRUM: there exist positive integers $\eta_1, \dots, \eta_g$, $g \leq n$, such that all eigenvalues of A with $|\lambda| = \lambda_0$ are given by:

$$\lambda = \lambda_0 e^{j\frac{2\pi}{\eta_h}k_h}, \ h = 1, 2, \dots, g, \ k_h = 1, 2, \dots, \eta_h;$$

(iv) EXISTENCE OF A STRICTLY POSITIVE EIGENVECTOR IN U_{λ_0}: the eigenspace U_{λ_0} corresponding to the maximal eigenvalue λ_0 has a strictly positive eigenvector $\mathbf{v}_0$ if and only if, in the normal form (A.1), λ_0 is an eigenvalue of all isolated blocks $\hat{A}_{i,i}, i = 1, 2, \dots, h$, but this is not true for any of the remaining diagonal blocks $\hat{A}_{i,i}, i > h$;

(v) MONOTONICITY OF λ_0: if $\bar{A}$ is such that $\bar{A} > A$, its maximal eigenvalue satisfies $\bar{\lambda}_0 \geq \lambda_0$.

Again, we omit the proof of the previous theorem, which can be found for instance in [6].

A.2 Metzler matrices

The focus of the present section will be on Metzler matrices, an important class of matrices that has attracted a great deal of attention over the past decades and still represents the object of an intense study.

Definition A.2. A real matrix $A \in \mathbb{R}^{n \times n}$ is called a *Metzler matrix* if all the off-diagonal entries are non-negative, namely $a_{ij} \geq 0, \forall i \neq j$.

Definition A.3. A real matrix $A \in \mathbb{R}^{n \times n}$ is a *Metzler Hurwitz matrix* if it is a Metzler matrix and the real part of each eigenvalue of A is negative.

It is worth noticing that any Metzler matrix A can be represented as:

$$A = A_+ - \lambda I_n, \quad \lambda \in \mathbb{R}, \ A_+ \in \mathbb{R}_+^{n \times n}. \tag{A.2}$$

Consequently, as the spectrum of A and the spectrum of A_+ are linked by the relation $\Lambda(A) = \Lambda(A_+) - \lambda$, (A.2) defines a Metzler Hurwitz matrix if and only if $\lambda > \lambda_0$, where λ_0 is the Perron eigenvalue of A_+.

Several equivalent definitions for Metzler Hurwitz matrices involving concepts from many areas of linear algebra are available in the literature (just think that Berman and Plemmons in [1] provide fifty equivalent definitions for M-matrices, which are the opposite of Metzler Hurwitz matrices). Far away from providing a systematic treatment of such a class of matrices, in the following proposition we list a selection of their most useful properties.

Proposition A.3. Let $A \in \mathbb{R}^{n \times n}$ be a Metzler matrix. Then, the following statements are equivalent:

(i) A is a Metzler Hurwitz matrix;

(ii) The coefficients of the characteristic polynomial of A are all positive, namely if $\Delta_A(z) = \alpha_0 + \alpha_1 z + \dots + \alpha_{n-1} z^{n-1} + z^n$, then $\alpha_i > 0$ for every $i = 0, \dots, n-1$;

(iii) A is non-singular and A^{-1} is a negative matrix, namely $A^{-1} < 0$;

(iv) For every vector $\mathbf{d} \in \mathbb{R}^n$ the condition $A\mathbf{d} \leq 0$ implies $\mathbf{d} \geq 0$;

(v) There exists a vector $\mathbf{d} \gg 0$ such that $A\mathbf{d} \ll 0$;

(vi) A has all negative diagonal elements and there exists a positive diagonal matrix D such that $-AD$ is strictly diagonally dominant, namely:

$$-a_{ii}d_i > \sum_{j \neq i} |-a_{ij}| d_j, \; i = 1, \ldots, n;$$

(vii) A has all negative diagonal elements and there exists a positive diagonal matrix D such that $-D^{-1}AD$ is strictly diagonally dominant;

(viii) The matrix $-A$ is a P-matrix, namely all its principal minors are positive;

(ix) There exist lower and upper triangular matrices L and U, respectively, with positive diagonals such that:

$$A = -LU;$$

(x) There exists a positive diagonal Lyapunov solution, namely there exists a positive diagonal matrix D such that $DA + A^T D$ is negative definite.

A more exhaustive characterization of Metzler Hurwitz matrices (actually, of their opposites, the M-matrices) can be found for instance in [1],[15].

Both positive matrices and Metzler matrices are central to the study of positive systems, which constitute the object of the following sections. In particular, we present necessary and sufficient conditions for positivity and, subsequently, we briefly discuss asymptotic stability. As far as other classical properties of dynamical systems are concerned, such as reachability, observability and input-output maps, we refer the reader to [6],[5],[18].

A.3 Discrete-time positive systems

Consider the discrete-time system described by the equations:

$$\begin{cases} \mathbf{x}(t+1) = F\mathbf{x}(t) + G\mathbf{u}(t), \\ \mathbf{y}(t) = H\mathbf{x}(t) + D\mathbf{u}(t). \end{cases} \tag{A.3}$$

System (A.3) is (internally) positive if and only if for any $\mathbf{x}(0) \in \mathbb{R}^n_+$ and every $\mathbf{u}(t) \in \mathbb{R}^m_+$ we have $\mathbf{x}(t) \in \mathbb{R}^n_+$ and $\mathbf{y}(t) \in \mathbb{R}^p_+$ for all $t \geq 0$.

Proposition A.4. The discrete-time system (A.3) is positive if and only if $F \in \mathbb{R}^{n\times n}_+$, $G \in \mathbb{R}^{n\times m}_+$, $H \in \mathbb{R}^{p\times n}_+$, $D \in \mathbb{R}^{p\times m}_+$.

Proof. SUFFICIENCY: It is obvious that if F, G, H, D have all non-negative entries, any non-negative initial state and any non-negative input will generate a non-negative trajectory and output at all times.

NECESSITY: If we choose $\mathbf{u}(0) = 0$ and $\mathbf{x}(0) = \mathbf{e}_i$, the state $\mathbf{x}(1)$ and the output $\mathbf{y}(0)$ are respectively the i-th column of F and the i-th column of H and this implies that both F and H must be positive matrices. Analogously, if we choose $\mathbf{u}(0) = \mathbf{e}_i$ and $\mathbf{x}(0) = 0$, the state $\mathbf{x}(1)$ and the output $\mathbf{y}(0)$ are respectively the i-th column of G and the i-th column of D and hence both G and D must be positive matrices. □

As far as asymptotic stability is concerned, it is worth noticing that asymptotic stability with respect to initial conditions in the positive orthant is equivalent to asymptotic stability with respect to arbitrary initial conditions in $\mathbb{R}^n$. Indeed, any initial state $\mathbf{x}(0) \in \mathbb{R}^n$ can be written as $\mathbf{x}(0) = \mathbf{x}_1 - \mathbf{x}_2$ where $\mathbf{x}_1, \mathbf{x}_2 > 0$. Hence, all the conditions for asymptotic stability of linear systems, hold true for positive systems. Nevertheless, positive systems have some peculiar properties that are quite useful for the stability analysis, as shown in the following proposition.

Proposition A.5. Let $F \in \mathbb{R}_+^{n\times n}$. Then, the following statements are equivalent:

(i) All eigenvalues of F have moduli less than 1, namely $|\Lambda(F)| < 1$;

(ii) The matrix $F - I_n$ is a Metzler Hurwitz matrix;

(iii) The coefficients of the characteristic polynomial of $F - I_n$, $\Delta_{F-I_n}(z)$, are all positive;

(iv) There exists a linear CLF $V(\mathbf{x}) := \mathbf{v}^T\mathbf{x}$ such that:

$$\Delta V(\mathbf{x}) := \mathbf{v}^T(F - I_n)\mathbf{x} \ll 0, \ \forall \mathbf{x} > 0;$$

(v) The Lyapunov equation:

$$F^T X F - X = -Q,$$

with $-Q$ a suitable definite negative matrix, admits a positive diagonal solution, namely there exists a positive diagonal matrix D such that $F^T D F - D$ is negative definite.

A.4 Continuous-time positive systems

Consider the continuous-time system described by the equations:

$$\begin{cases} \dot{\mathbf{x}}(t) = F\mathbf{x}(t) + G\mathbf{u}(t), \\ \mathbf{y}(t) = H\mathbf{x}(t) + D\mathbf{u}(t). \end{cases} \tag{A.4}$$

Analogously to its discrete-time counterpart, the system (A.4) is (internally) positive if and only if for any $\mathbf{x}(0) \in \mathbb{R}_+^n$ and every $\mathbf{u}(t) \in \mathbb{R}_+^m$ we have $\mathbf{x}(t) \in \mathbb{R}_+^n$ and $\mathbf{y}(t) \in \mathbb{R}_+^p$ for all $t \geq 0$.

Proposition A.6. The continuous-time system (A.4) is positive if and only if F is a Metzler matrix and $G \in \mathbb{R}_+^{n\times m}$, $H \in \mathbb{R}_+^{p\times n}$, $D \in \mathbb{R}_+^{p\times m}$.

Proof. SUFFICIENCY: Recall that the state at any time instant $t \in \mathbb{Z}_+$, starting from the initial condition $\mathbf{x}(0)$ and under the effect of the input $\mathbf{u}(\cdot)$, can be expressed as:

$$\mathbf{x}(t) = e^{Ft}\mathbf{x}(0) + \int_0^t e^{F(t-\tau)} G\mathbf{u}(\tau)\, d\tau,$$

and hence we only need to prove that e^{Ft} is a positive matrix for every $t \geq 0$. This is rather straightforward if we recall that any Metzler matrix can be represented as $F = F_+ - \lambda I_n$ for some $\lambda \in \mathbb{R}$ and some positive matrix $F_+ \in \mathbb{R}_+^{n\times n}$. Indeed:

$$e^{Ft} = e^{(F_+ - \lambda I_n)t} = e^{-\lambda t} e^{F_+ t} = e^{-\lambda t} \sum_{i=0}^{+\infty} F_+^i \frac{t^i}{i!},$$

which implies that e^{Ft} is a positive matrix for every $t \geq 0$.

NECESSITY: If we choose $\mathbf{x}(0) = 0$, the positivity assumption ensures $\dot{\mathbf{x}}(0) = G\mathbf{u}(0) \geq 0$ for every $\mathbf{u}(0) \geq 0$, which in turn implies $G \in \mathbb{R}_+^{n\times m}$. Analogously, $\mathbf{y}(0) = D\mathbf{u}(0)$ and, by the positivity assumption, we have $\mathbf{y}(0) \geq 0$ for every $\mathbf{u}(0) \geq 0$, which amounts to $D \in \mathbb{R}_+^{p\times m}$. Finally, if we choose $\mathbf{x}(0) = \mathbf{e}_i$ and $\mathbf{u}(0) = 0$, $\dot{\mathbf{x}}(0) = F\mathbf{x}(0)$ is the i-th column of F. Since the trajectory of a positive system can not leave the positive orthant $\mathbb{R}_+^n$, each element of the vector $\dot{\mathbf{x}}(0)$, except for the i-th, must be positive or zero. Therefore, the off-diagonal elements of F must be positive or zero, which amounts to saying that F is a Metzler matrix. □

To conclude, the following proposition, which is the continuous-time counterpart of Proposition A.5, characterizes asymptotic stability for a continuous-time positive system.

Proposition A.7. Let $F \in \mathbb{R}^{n\times n}$ be a Metzler matrix. Then, the following statements are equivalent:

(i) All the eigenvalues of F have negative real parts, namely $\Re\{\Lambda(F)\} < 0$;

(ii) F is a Metzler Hurwitz matrix;

(iii) The coefficients of the characteristic polynomial of F, $\Delta_F(z)$, are all positive;

(iv) There exists a linear CLF $V(\mathbf{x}) := \mathbf{v}^T\mathbf{x}$ such that:

$$\dot{V}(\mathbf{x}) := \mathbf{v}^T F\mathbf{x} \ll 0, \ \forall \mathbf{x} > 0;$$

(v) The Lyapunov equation:

$$F^T X + XF = -Q,$$

with $-Q$ a suitable definite negative matrix, admits a positive diagonal solution, namely there exists a positive diagonal matrix D such that $F^T D + DF$ is negative definite.

Bibliography

[1] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia: SIAM, 1994.

[2] G. R. Burmester and A. Pezzutto. *Color Atlas of Immunology*. Stuttgart: Thieme, 2003.

[3] M. M. Denn. *Optimization by variational methods*. New York: McGraw-Hill, 1969.

[4] T. S. Doan, A. Kalauch, and S. Siegmund. “A constructive approach to linear Lyapunov functions for positive switched systems using Collatz-Wielandt Sets”. In: *IEEE Trans. Automat. Control* 58.3 (2013), pp. 748–751.

[5] L. Farina and S. Rinaldi. *Positive Linear Systems: Theory and Applications*. New York: Wiley, 2000.

[6] E. Fornasini. *Appunti di teoria dei sistemi*. Padova: Edizioni Libreria Progetto, 2013.

[7] E. Fornasini and M. E. Valcher. “Linear copositive Lyapunov functions for continuous-time positive switched systems”. In: *IEEE Trans. Automat. Control* 55.8 (2010), pp. 1933–1937.

[8] E. Fornasini and M. E. Valcher. “Stability and stabilizability criteria for discrete-time positive switched systems”. In: *IEEE Trans. Automat. Control* 57.5 (2012), pp. 1208–1221.

[9] J. C. Geromel and P. Colaneri. “Stability and stabilization of discrete time switched systems”. In: *International Journal of Control* 79.7 (2006), pp. 719–728.

[10] E. Hernandez-Vargas, R. Middleton, and P. Colaneri. “Optimal and MPC switching strategies for mitigating viral mutation and escape”. In: *Proceedings of the 2011 IFAC World Congress*. Milano, Italy, 2011.

[11] E. Hernandez-Vargas et al. “Discrete-time control for switched positive systems with application to mitigating viral escape”. In: *International Journal of Robust and Nonlinear Control* (2010). DOI: 10.1002/rnc.1628.

[12] E. A. Hernandez-Vargas. “A control theoretic approach to mitigate viral escape in HIV”. PhD Thesis. Maynooth: Hamilton Institute, National University of Ireland, 2011.

[13] E. A. Hernandez-Vargas, P. Colaneri, and R. H. Middleton. “Switching strategies to mitigate HIV mutation”. In: *IEEE Transactions on Control Systems Technology* 22.4 (2014), pp. 1623–1628.

[14] E. A. Hernandez-Vargas et al. “Dynamic Optimization Algorithms to mitigate HIV escape”. In: *49th IEEE Conference on Decision and Control*. Atlanta, USA, 2010, pp. 827–832.

[15] R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge: Cambridge University Press, 1991.

[16] Fundamental Immunology. *W. E. Paul*. Lippincott Williams and Wilkins, 2003.

[17] A. Jadbabaie, J. Lin, and A. S. Morse. “Coordination of groups of mobile autonomous agents using nearest neighbor rules”. In: *IEEE Trans. Aut. Contr.* 48.6 (2003), pp. 988–1001.

[18] T. Kaczorek. *Positive 1D and 2D Systems*. London: Springer-Verlag, 2002.

[19] F. H. Kayser et al. *Color Atlas of Medical Microbiology*. Thieme, 2005.

[20] F. Knorn, O. Mason, and R. Shorten. “On linear co-positive Lyapunov functions for sets of linear positive systems”. In: *Automatica* 45.8 (2009), pp. 1943–1947.

[21] H. Kwakernaak and R. Sivan. *Linear Optimal Control systems*. New York: Wiley-interscience, 1972.

[22] A. Landi et al. “Modeling and control of HIV dynamics”. In: *Computer Methods and Programs in Biomedicine* (2007).

[23] D. Liberzon. *Switching in systems and Control*. Boston: Birkhauser, 2003.

[24] H. Lin and P. J. Antsaklis. “Stability and stabilizability of switched linear systems: a short survey of recent results”. In: *Proceedings of the 2005 IEEE International Symposium on Intelligent Control*. Limassol, Cyprus, 2005, pp. 24–29.

[25] H. Lin and P. J. Antsaklis. “Stability and stabilizability of switched linear systems: a survey of recent results”. In: *IEEE Trans. Automat. Control* 54.2 (2009), pp. 308–322.

[26] Xingwen Liu. “Stability analysis of switched positive systems: a switched linear copositive Lyapunov function method”. In: *IEEE Trans. Circuits Syst. II, Exp. Briefs* 56.5 (2009), pp. 414–418.

[27] O. Mason and R. Shorten. “A result on the existence of common quadratic Lyapunov functions for pairs of stable discrete-time LTI systems”. In: *Proceedings of the American Control Conference*. Denver,Colorado, 2003, pp. 4471–4475.

[28] O. Mason and R. Shorten. “On linear copositive Lyapunov functions and the stability of switched positive linear systems”. In: *IEEE Trans. Aut. Contr.* 52.7 (2007), pp. 1346–1349.

[29] O. Mason and R. Shorten. “On the simultaneous diagonal stability of a pair of positive linear systems”. In: *Linear Algebra Appl.* 413 (2006), pp. 13–23.

[30] O. Mason and R. Shorten. “Quadratic and copositive Lyapunov functions and the stability of positive switched linear systems”. In: *Proocedings of the 2007 American Control Conference*. New York City, USA, 2007, pp. 657–662.

[31] H. Minc. *Nonnegative Matrices*. New York: Wiley, 1988.

[32] G. Pannocchia, M. Laurino, and A. Landi. “A Model Predictive Control strategy toward optimal structured interruptions in anti-HIV therapy”. In: *IEEE Transactions on Biomedical Engineering* 57.5 (2010), pp. 1040–1050.

[33] A. Paul et al. “A switched system model for stability analysis of distributed power control algorithms for cellular communications”. In: *Proceedings of the 2004 American Control Conference*. Boston, Massachusetts, 2004.

[34] A. Perelson and P. Nelson. “Mathematical analysis of HIV-1 dynamics in vivo”. In: *SIAM Review* 41.1 (1999), pp. 3–44.

[35] R. Shorten, F. Wirth, and D. Leith. “A positive system model of TCP-like congestion control: Asymptotic results”. In: *IEEE/ACM Trans. Networking* 14.3 (2006), pp. 616–629.

[36] Z. Sun and S. S. Ge. *Switched Linear Systems: Control and Design*. London: Springer-Verlag, 2005.

[37] Y. Tong, C. Wang, and L. Zhang. “Stabilization of discrete-time switched positive linear systems via time- and state-dependent switching laws”. In: *IET Control Theory Appl.* 6.11 (2012), pp. 1603–1609.

[38] D. Wodarz and M. A. Nowak. “Specific therapy regimes could lead to long-term immunological control of HIV”. In: *Proc. National Acad. Sci.* USA, 1999, pp. 14464–14469.

[39] R. Zurakowski and A. R. Teel. “A Model Predictive Control based scheduling methos for HIV therapy”. In: *Journal of Theoretical Biology* 238 (2006), pp. 368–382.

Ringraziamenti

Un primo, sentito, enorme ringraziamento alla professoressa Maria Elena Valcher. Avrei molti motivi per ringraziarla, ma sarò breve: mi ha proposto un argomento di tesi che va incontro perfettamente a ciò che stavo cercando; ha sempre creduto, dall'inizio alla fine, che la laurea nella sessione di luglio fosse un obiettivo raggiungibile; è sempre stata presente quando ho avuto bisogno del suo aiuto, offrendomi una disponibilità senza limiti come pochi altri avrebbero fatto.

Ringrazio sinceramente il professor Ettore Fornasini per aver fatto crescere grandemente il mio interesse per la Teoria dei Sistemi e dei Controlli: mi ha sempre stupito il modo in cui riesce a far apparire le cose meno complicate di quanto non lo siano in realtà.

Ringrazio infinitamente la mia famiglia perchè in questi cinque anni non ha mai smesso di credere in me dandomi molta più fiducia di quanta io ne avessi in me stessa, ha sempre cercato di farmi ragionare evitando che prendessi decisioni azzardate, mi ha sempre sostenuta (e, lo riconosco, sopportata) nei momenti difficili.

Infine un grazie di cuore a Chiara e Eleonora: quando ho scelto di seguire Teoria dei Sistemi sapevo di poter contare sul loro aiuto. Grazie perchè in questi ultimi due anni è stato bello aiutarsi, sostenersi la sera prima degli esami, condividere le ansie, le paure, ma anche le soddisfazioni.

Irene Zorzan