UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Dynamic output-feedback controllers for discrete-time linear systems with markovian jumping parameters, imperfect mode observation and additive noise perturbations

Pablo Ivan Zamora Mercado

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)

ICMC USP
SÃO CARLOS

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

Pablo Ivan Zamora Mercado

Dynamic output-feedback controllers for discrete-time linear systems with markovian jumping parameters, imperfect mode observation and additive noise perturbations

Master dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Program in Computer Science and Computational Mathematics. *EXAMINATION BOARD PRESENTATION COPY*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. Eduardo F. Costa

USP – São Carlos
February 2020

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

Z25d

Zamora Mercado, Pablo Iván
Dynamic output-feedback controllers for discrete-time linear systems with markovian jumping parameters, imperfect mode observation and additive noise perturbations / Pablo Iván Zamora Mercado; orientador Eduardo Fontoura Costa. -- São Carlos, 2020.
63 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2020.

1. Markov jump linear systems. 2. Hybrid systems. 3. Optimal control. 4. Optimization. 5. Stochastic Control. I. Fontoura Costa, Eduardo, orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Pablo Ivan Zamora Mercado

Controladores dinâmicos de feedback de saída para sistemas lineares de tempo discreto com parâmetros de salto markoviano, observação de modo imperfeito e perturbações aditivas de ruído

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional. *EXEMPLAR DE DEFESA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. Eduardo F. Costa

USP – São Carlos
Fevereiro de 2020

To my father Gerardo who gave me the greatest gift anyone could give another person: He believed in me.

To my mother Ana Maria who taught me about gratitude, patience, faith and to be strong no matter what.

To my Grandfather Hugo, for his wisdom and having his arms always open for me.

ACKNOWLEDGEMENTS

I want to thank CAPES Scholarship for the funding and also for supporting my research activities.

I wish to express my deepest gratitude to Professor Dr. Eduardo F. Costa, who was like a Father for during my time in São Carlos, for his guidance and overall insights in this field that have made this an inspiring experience for me. For his support, encouragement and patience I'm in debt.

I would like to thank to Luiz and Junior who are part of the Complex Systems, Particles and Control Research group, his help was determinant during research.

To my family who was always there no matter what, Gerardo, Ana Maria, Vivian and Fernando thank you.

To Omar my deepest gratitude, for helping in my darkest times.

To my friends in São Carlos Daniel, Paulo, Lito, Elvis, Mario, Juan Luis, and the guys who were always there when I needed them the most.

To Ramalakwmi Justiniano, my girlfriend, best friend and true love, for putting up with me being far away. Thanks for all your support

“ ...I leave Sisyphus at the foot of the mountain. One always finds one’s burden again. But Sisyphus teaches the higher fidelity that negates the gods and raises rocks. He too concludes that all is well. This universe henceforth without a master seems to him neither sterile nor futile. Each atom of that stone, each mineral flake of that night-filled mountain, in itself, forms a world. The struggle itself toward the heights is enough to fill a man’s heart.

One must imagine Sisyphus happy...”

(The Myth of Sisyphus, Albert Camus)

RESUMO

ZAMORA, P.I. **Controladores dinâmicos de feedback de saída para sistemas lineares de tempo discreto com parâmetros de salto markoviano, observação de modo imperfeito e perturbações aditivas de ruído**. 2020. 60 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

Uma classe de controladores de retorno de saída dinâmicos estacionários para sistemas lineares de salto markoviano em tempo discreto (MJLS), considerando a minimização de o custo médio a longo prazo é estudado.

Uma classe de controladores de feedback de saída dinâmicos estacionários para sistemas lineares de salto markoviano em tempo discreto (MJLS), considerando a minimização de o custo médio a longo prazo é estudado. A cadeia de Markov que governa os parâmetros não precisa ser ergódica e é permitido que seja periódica e contenha estados transitórios / classes não comunicantes, o que aumenta a classe do sistema, compreendendo sistemas periódicos como uma subclasse. Uma formulação compacta de otimização é obtida para o independente de modo/ baseado em detector controlador de ordem parcial / total, permitindo explorar a complexidade e consequentemente obtenha o melhor desempenho implementável para um aplicativo. Desenvolvemos um algoritmo de viabilidade - otimização em dois estágios, usando a abordagem baseada no operador. Apresentamos um conjunto de exemplos numéricos no contexto aleatório não trivial sistemas sujeitos a saltos para representar nossos resultados e comparar o desempenho com um algoritmo genético clássico, resultando em uma clara vantagem para o algoritmo proposto

).

Palavras-chave: Sistemas Lineares,parametros com saltos Markovianos,Controle Otimo, Controle Estocástico, Otimização.

ABSTRACT

ZAMORA, P.I. **Dynamic output-feedback controllers for discrete-time linear systems with markovian jumping parameters, imperfect mode observation and additive noise perturbations**. 2020. 60 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

A class of stationary dynamic output-feedback controllers for discrete-time Markovian Jumping Linear Systems(MJLS) considering the minimization of a long run average cost is studied. The Markov chain that governs the parameters is not required to be ergodic, and it is allowed to be periodic and contain transient states / non-communicating classes, which enlarges the class of system, e.g. now comprising periodic systems as a sub-class. A compact optmization formulation is obtained for the mode-independent/detector-based controller of partial/full order, allowing one to explore the complexity and consequently obtain the best implementable performance for an application. We develop a two stage feasibility - optimization algorithm using the operator-based approach. We present a set of numerical examples in the context of random non-trivial systems that are subject to jumps in order to represent our results and compare the performance with a classical genetic algorithm, resulting in a clear advantage for our proposed algorithm.

Keywords: Linear systems, Markov jump parameters, optimal control, Stochastic Control, Optimization.

CONTENTS

CHAPTER

1

INTRODUCTION

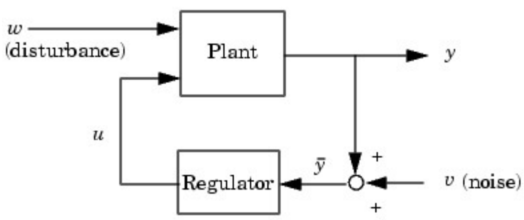
The first chapter of this thesis is devoted to explain the main characteristics of Markov Jump Linear Systems, the principal topic of this work. We will justify his study and the control scheme selected and his main characteristics. A description on how the chapters are organized is given at the end.

1.1 Markov Jump Linear Systems

Influenced by mechanics and engineering in 1940, control theory emerged. These sciences induced the need to model physical phenomena as state space ordinary differential equations. This interest led to the rise of three theoretical paradigms in control (POLDERMAN; WILLEMS, 2010). Pontryagin developing the maximum principle, Bellman dynamical programming and Kalman with his Linear Quadratic Gaussian(LQG) con- troller(BERTSEKAS, 1987). A good background material in to the foundations are(SONTAG, 1998; POLDERMAN; WILLEMS, 2010). We make special emphasis in the LQG problem whose creation was a breakthrough, in the sense that it achieves an optimal disturbance attenuation. Figure 1

Figure 1 – LQG Problem

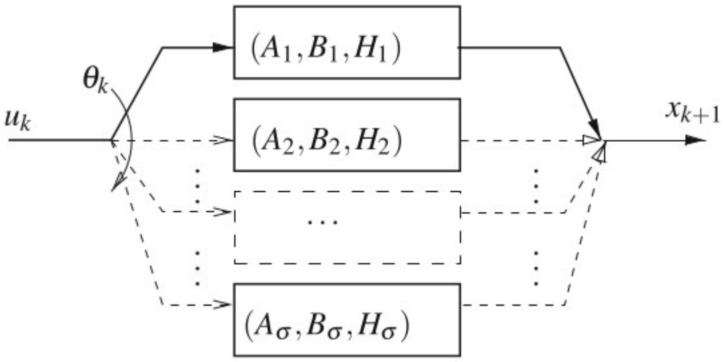
Source: Polderman and Willems (2010).



This result, considered for the first time the stochasticity in control and solved problems that arose in the design of gunfire control systems, during the second world war as stated

Figure 2 – A scheme of MJLS

Source: Vargas, Costa and Val (2013).



on (ÅSTRÖM, 1970). The study of a subgroup of this theory, resulted in an original paper by Krasovskii and Lidskii in 1961 (KRASOVSKII; LIDSKII, 1961), introducing a class of systems, that has abrupt parameter changes according to a stochastic process. Markov jump linear systems are a subclass of stochastic systems that can be characterized for having a continuous-valued (discrete-valued) dynamics that is linear in time and a discrete-valued dynamic that is modeled as a Markov chain that is independent of the other set of continuous-valued (discrete-valued) dynamics. For the Markov jump linear system, a more didactic explanation can be done, considering that a dynamical system G, is described by a non autonomous ODE in matrix form (A_1, B_1, H_1). In a certain moment, suppose an abrupt change caused the model to change and be described by (A_2, B_2, H_2), more generally, we can imagine that the system is subject to a series of changes that make it switch, over time, among a countable set of models θ_k.

1.2 Control of Markov Jump Linear Systems

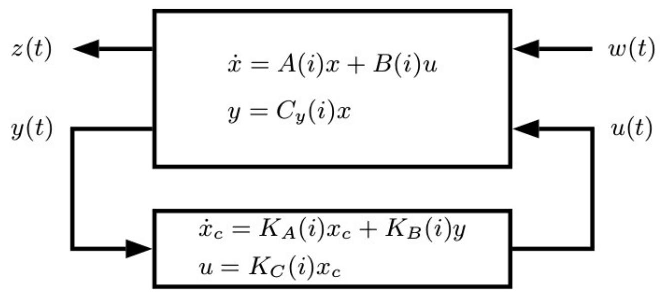
In this section, we will giver some brief explanations of main concepts on MJLS we will reference forward, we also denote the main articles where the theoretical framework of this work lies.

The Markov Jump Linear Systems(MJLS) framework allows a more complex approach in order to model processes that have abrupt changes or disturbances, a new way to deal with the control of those systems. Much work has been done extending the theory of control of linear systems to MJLS (COSTA; FRAGOSO; MARQUES, 2005). And much investigation has to be done to extend results to other systems and controllers applications of MJLS include. Considering the broad applications the development of this theory has brought,for example for macroeconomic systems(VAL; BASAR, 1999), networked control (HESPANHA; NAGHSHTABRIZI; XU, 2007), Fault isolation in Multi-Agent systems(MESKIN; KHORASANI, 2009), Finance and Portfolio Optimization(COSTA; ARAUJO, 2008),(ELLIOTT T. K. SIU; LAU, 2007) , Quadcopter failure Control(DRAGAN; COSTA, 2016a), A DC motor device(VARGAS; COSTA; VAL, 2013) , Robotics and fault tolerant manipulators (SIQUEIRA; TERRA, 2009).

1.2.1 Output feedback controllers

In the state feedback scheme we assume complete access to the state vector x, which is not always possible in real world problems due to physical limitations and/or costs involved. This assumption makes it impossible to apply a state feedback approach,so new schemes are introduced. One possibility is to use *static* output feedback in the form $u(t) = Fy(t)$; this is an easy to implement controller, however the task of finding a suitable F is frequently very complex. Another possibility is to use the so called *dynamic* controller, which we denote by G_c. we will use the system output vector,

Figure 3 – State feedback stabilization block diagram



For a discrete-dynamical system one can construct the closed loop system coupling the output of the controller to the control variable of the system and the output of the system to the input of the controller.

1.2.2 Class of controllers with partial or non-observation of the mode

The theoretical framework for optimal control of MJLS with mode observation is quite mature and it has many strong results(COSTA; VARGAS, 2011; COSTA; FRAGOSO; MARQUES, 2005; COSTA; FRAGOSO; TODOROV, 2013; CHIZECK; WILLSKY, 1986; DRAGAN; COSTA, 2016a). The situation is different in control problems with partial or no information of the Markov state, there are few existing results(VARGAS; COSTA; VAL, 2013; DOLGOV; CHLEBEK; HANEBECK, 2016). This type of problem is much harder because the controller can, in principle, be used to both optimize the objective function and to help raise information about the Markov state and these tasks are sometimes conflicting. Now, although the problem is far more complex, there is a great practical appeal for studying it, because the Markov state is not available in many applications. It is particularly interesting for applications that the controller does not take information on the chain into account at all. Which makes it easier to implement in a embedded controller for example. This approach has narrow results; previous works have been focusing in a particular class of controllers, the state feedback controllers, the results are presented in (VARGAS; COSTA; VAL, 2016; BORTOLIN, 2012; OLIVEIRA, 2014; SILVA, 2012) where the state variable, x is assumed to be available.

In this work we assume x is not available and it is worth to mention that, (DRAGAN; COSTA, 2016a) is a natural starting point because it uses a general approach the one we are looking for; it computes the gains for a dynamic stabilizing output feedback controller sadly the approach cannot be used totally in the sense that it is based on filtering and control coupled Riccati equations that can be solved by well-known,sadly this efficient methods are not possible to use in the case when we cannot observe the governing Markov chain.That is why we also study results of the work of the KIT Intelligent Sensor-Actuator-Systems (ISAS) Research Group (DOLGOV; HANEBECK, 2015; DOLGOV; CHLEBEK; HANEBECK, 2016; DOLGOV; HANEBECK, 2017) as the ones from Vargas et al.(VARGAS; COSTA; VAL, 2013) that explores the state feedback approach(also without the observation of the mode) we will also use the results from (TODOROV; FRAGOSO; COSTA, 2018; COSTA; FRAGOSO; TODOROV, 2015; de Oliveira; Costa; Daafouz, 2018) were the detector based solution is explored, by detector, we mean that an process estimates the situation of the markov chain governing the switching.

1.3 Motivation

The interest in this type of problem is due to a combination between the potential for applications and, at same time, availability of results and algorithms that can be used for finding solutions. For instance, the controller given in (DRAGAN; COSTA, 2016a) is computed based on filtering and control coupled Riccati equations that can be solved by well-known, efficient methods. Moreover, the formulation in (DRAGAN; COSTA, 2016a) allows one to explore the trade-off between the complexity and performance of the controller; more specifically, the dimension of the dynamic controller can be selected by the user, ranging from $n_c = 0$ to a full order controller with $n_c = n$, where n is the dimension of the state component x of the plant. Another relevant aspect for application of an MJLS is that the jump variable θ is not necessarily observed in a perfect and immediate way (VARGAS; COSTA; VAL, 2016; DOLGOV; CHLEBEK; HANEBECK, 2016; DOLGOV; HANEBECK, 2017), which motivated the development of mode-independent and detector-based controllers, see e.g. (VARGAS; COSTA; VAL, 2016; DOLGOV; HANEBECK, 2017) and (TODOROV; FRAGOSO; COSTA, 2018; COSTA; FRAGOSO; TODOROV, 2015; de Oliveira; Costa; Daafouz, 2018) respectively. The controller studied in this note gives one more step towards flexibility and applicability. It is detector-based featuring variable dimension n_c, and the case of static output feedback is also included. In this way, at one extreme one can design a controller of dimension $n_c = 0$ and only one mode (as in (VARGAS *et al.*, 2016)), and at the other extreme a full order mode dependent controller (as in (DRAGAN; COSTA, 2016a)). The intermediary cases with $1 \leq n_c \leq n$ and mode-independent/detector-based controllers for the considered class of problem are studied here for the first time in literature.

CHAPTER

2

NOTATION AND PRELIMINARY RESULTS

In the second Chapter an extended presentation is made for the notation that is used along the Thesis, we also show some fundamental results that will be used in the next chapters and support its development.

2.1 Notations

Let $\mathbb{R}^r$ denote the usual r-th dimensional Euclidean space, $\mathcal{M}_{r\times d}$ ($\mathcal{M}_d$) be the linear space formed by all real matrices $\mathbb{R}^{r\times d}$ ($\mathbb{R}^{d\times d}$), let $\mathcal{S}_d \subset \mathcal{M}_d$ be the linear subspace of symmetric matrices of size $d\times d$ and $\mathcal{S}_d^+$ ($\mathcal{S}_d^0$) the closed (open) convex cone of positive semi-definite (definite) matrices.

$\mathcal{M}_{r\times d}^{\mathcal{N}}$ is the $\mathcal{N}$-th cartesian product $\mathcal{M}_{r\times d}\times\mathcal{M}_{r\times d}\times\cdots\times\mathcal{M}_{r\times d}$ indexed by the set $\Theta := \{1,\ldots,\mathcal{N}\}$ forming a linear space of indexed matrices, that is, for an arbitrary $\mathbf{V} := (V(1),...,V(\mathcal{N})) \in \mathcal{M}_{r\times d}^{\mathcal{N}}$, we have that $V(i)\in\mathcal{M}_{r\times d}$; the sets $\mathcal{S}_d^{\mathcal{N}}$, $\mathcal{S}_d^{\mathcal{N}(+)}$ and $\mathcal{S}_d^{\mathcal{N}(0)}$ are defined following the same principle.

We also employ the ordering $\mathbf{V} > \mathbf{U}$ ($\mathbf{V}\geq\mathbf{U}$) for elements of $\mathcal{S}_d$, meaning that $V(i)-U(i)$ is positive definite (semi-definite) for all $i\in\Theta$. Relations involving elements of $\mathcal{M}_{r\times d}^{\mathcal{N}}$ are defined in an element-wise way, e.g. $\mathbf{S}=\mathbf{U}\mathbf{V}$ is such that $S(i)=U(i)V(i)$, in addition, with $\mathbf{V}\in\mathcal{M}_{r\times d}^{\mathcal{N}}$ and $T\in\mathbb{R}^{d\times l}$ the product $\mathbf{V}T$ represents the indexed set $(V(1)T,...,V(N)T)$, following the same logic for $T\mathbf{V}$. We define $\mathbb{I}_n$ as the indexed identity matrix $(I_n(1),\ldots,I_n(N))$.
We denote $tr\{\cdot\}$ as the trace operator. For this space an inner product is defined,

$$\langle\mathbf{V},\mathbf{U}\rangle = \sum_{i=1}^{\mathcal{N}} tr[V(i)^\intercal\, U(i)]. \tag{2.1}$$

for an arbitrary $\mathbf{V},\mathbf{U}\in\mathcal{M}_{r\times d}$ and a corresponding norm $||V||_2^2 = \langle\mathbf{V},\mathbf{V}\rangle$, inducing a Hilbert space structure on $\mathcal{M}_d^{\mathcal{N}}$.

If $f : \mathcal{M}_{r\times d} \to \mathbb{R}$ is a differentiable function on the domain $\mathcal{M}_{r\times d}$, we denote the partial derivative $\partial f(G)/\partial G$ as $\partial_G f(\cdot)$ whenever $G \in \mathcal{M}_{r\times d}$.

We define the independent stochastic processes $\{\theta_k\}_{k\geq 0}$, $\{w_k\}_{k\geq 0}$ and $\{v_k\}_{k\geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$. It is assumed that w_k and v_k are independent, and Gaussian random sources of noise, with zero-mean covariances:
$E[w_k] = 0$, $E[v_k] = 0$, $E[w_k w_k^\intercal] = \mathcal{W}$, $E[v_k v_k^\intercal] = \mathcal{V}$ and $E[w_k v_k^\intercal] = 0$. Also, $E[w_t w_k^\intercal] = 0$, $E[v_t v_k^\intercal] = 0, t \neq k$.
$\{\theta_k\}_{k\geq 0}$ is an homogeneous Markov chain whose state θ_k takes values in the set $\Theta = \{1, 2, ..., \mathcal{N}\}$, with a transition probability matrix $P = (p_{i,j})$, $(i, j) \in \Theta \times \Theta$, that is, $p_{i,j} = \mathcal{P}(\theta_{t+1} = j | \theta_t = i)$. $\pi_k = (\pi_k(1), \ldots, \pi_k(\mathcal{N})$ denotes the distribution of θ_k.

The Cesàro limit distribution of the Markov chain can be computed as

$$\rho^{\pi_0}(j) = \sum_{k=1}^{\mathcal{N}} \pi_0(k)\rho_{k,j}, \quad 1 \leq j \leq N \tag{2.2}$$

where the vector $\pi_0 = (\pi_0(1), \ldots, \pi_0(N))$ is the initial distribution of the Markov chain. $\rho_{k,j}$ are elements of the following matrix Π, whose existence is demonstrated in (DOOB, 1953)

$$\Pi = \lim_{\tau\to\infty} \frac{1}{\tau} \sum_{k=0}^{\tau} P^k. \tag{2.3}$$

For the same space, $\mathbb{1}_A$ be the indicator function over a set A for a set, is defined for any $\omega \in \Omega$ (Dirac measurement).

Finally for sake of generality, we define a stochastic process $\{\gamma_k\}_{k\geq 0}$, with $\gamma_k \in \Gamma = \{1, \ldots, \mathcal{L}\}$, satisfying $\mathcal{P}(\gamma_k = \ell | \theta_k = i) = q_{i,\ell}$, we call this process a "detection process", with a detection probability matrix $P_\alpha = (q_{l,j})$, $(l, j) \in \Gamma \times \Gamma$, that is, $q_{l,j} = \mathcal{P}(\gamma_{t+1} = l | \gamma_t = j)$. for this new index we have $\mathcal{M}_d^{\mathcal{L}} \subset \mathcal{M}_d^{\mathcal{N}}$ and respectively for the other spaces defined above.

2.2 Lyapunov type operators

We introduce on the space $\mathcal{S}_d^{\mathcal{N}}$ the linear operator $\mathcal{L}_{cl}$, considering matrices $\mathcal{A} \in \mathcal{M}_{r\times d}^{\mathcal{N}}$ in the space defined above and the elements of the stochastic matrix P defined by:

$$(\mathcal{L}_{cl}\mathbf{V})(i) = \sum_{j\in\Theta} p_{ji}\mathcal{A}(j)V(j)\mathcal{A}(j)^\intercal, \quad 1 \leq i \leq \mathcal{N}, \tag{2.4}$$

for all $\mathbf{V} = (V(1), \ldots, V(\mathcal{N})) \in \mathcal{S}_d^{\mathcal{N}}$. By direct calculation one obtains that the adjoint operator of $\mathcal{L}_{cl}$ with respect to the inner product (2.1) is described by:

$$(\mathcal{L}_{cl}^*\mathbf{V})(i) = \sum_{j\in\Theta} p_{ij}\mathcal{A}(i)^\intercal V(j)\mathcal{A}(i), \quad 1 \leq i \leq \mathcal{N}. \tag{2.5}$$

CHAPTER

3

THE DETECTOR BASED DYNAMIC OUTPUT-FEEDBACK CONTROLLER PROBLEM

In this chapter, the problem, focus of this thesis is presented. We present a class of discrete-time dynamical system, that needs to be controlled by a detector based output feedback controller; we present operators and functionals that will help us to develop the theoretical framework of the solution proposal for them. With these tools we develop an deterministic cost function that later on will be formulated as an optimization problem finishing with a compact cost function that needs to be optimized.

3.1 Problem formulation

Let a class of discrete-time dynamical system **G** be of the form,

$$\begin{cases} x_{k+1} = A_{(\theta_k)} x_k + B_{(\theta_k)} u_k + H_{(\theta_k)} w_k \\ y_k = C_{(\theta_k)} x_k + J_{(\theta_k)} v_k \end{cases} \tag{3.1}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^p$ is the control variable and $y_k \in \mathbb{R}^m$ is the system output available to feed the controller. The system matrices belong to given sets of matrices with dimensions $\mathbf{A} \in \mathcal{M}_n^{\mathcal{N}}$, $\mathbf{B} \in \mathcal{M}_{n\times p}^{\mathcal{N}}$, $\mathbf{H} \in \mathcal{M}_{n\times l}^{\mathcal{N}}$, $\mathbf{C} \in M_{m\times n}^{\mathcal{N}}$ and $\mathbf{J} \in M_{m\times r}^{\mathcal{N}}$.

Consider also the following stationary linear dynamic controller G_c,

$$\begin{cases} \hat{x}_{k+1} = F_{(\gamma_k)} \hat{x}_k + K_{(\gamma_k)} y_k \\ u_k = L_{(\gamma_k)} \hat{x}_k + M_{(\gamma_k)} y_k, \end{cases} \tag{3.2}$$

where $\hat{x}_k \in \mathbb{R}^{n_c}$ denotes the internal state of the controller[1], $y_k \in \mathbb{R}^m$ the input of the controller, $u_k \in \mathbb{R}^p$ the output.

[1] The dimension n_c can be defined as any value in the interval $0 \leq n_c \leq n$.

The goal of this work is to design a dynamic/static output feedback controller that depends only on the detector, that is, γ_k is the only measure available to the controller. This is general enough as to include the case of no observation of θ_k (by making $\mathscr{L}=1$), cluster observation of θ_k, and perfect observation of θ_k (with $\mathscr{L}=\mathcal{N}$ and $q_{i,\ell}=1$ when $i=\ell$) (de Oliveira; Costa; Daafouz, 2018).

The controller matrices belong to sets of matrices $\mathbf{F}\in\mathcal{M}_{n_c}^{\mathcal{N}};\mathbf{K}\in\mathcal{M}_{n_c\times m}^{\mathscr{L}};\mathbf{L}\in\mathcal{M}_{p\times n_c}^{\mathscr{L}}$ and $\mathbf{M}\in\mathcal{M}_{p\times m}^{\mathscr{L}}$ of appropriate dimensions; these matrices are to be determined by an optimization problem whose objective function (or cost function) is given by:

$$\mathbf{J}=\lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}E\left[\sum_{k=0}^{\mathcal{K}}\left[x_k^\intercal Q_{(\theta_k)}x+u_k^\intercal R_{(\theta_k)}u_k\right]\right], \tag{3.3}$$

with weighting matrices that belong to the sets $\mathbf{Q}\in\mathcal{S}_n^{\mathcal{N}}$ and $\mathbf{R}\in\mathcal{S}_p^{\mathcal{N}(+)}$.

The controller (3.2) is coupled to the system, yielding a closed loop system that can be rewritten as

$$\mathbf{x}_{k+1}=\mathrm{A}_{\mathrm{cl}(\theta_k,\gamma_k)}\mathbf{x}_k+\mathrm{H}_{\mathrm{cl}(\theta_k,\gamma_k)}\mathbf{w}_k \tag{3.4}$$

where $\mathbf{x}_k\in\mathbb{R}^{(n+n_c)}$, $\mathbf{w}_k\in\mathbb{R}^{(l+r)}$,

$$\mathbf{x}_k=[x_k^\intercal\ \hat{x}_k^\intercal]^\intercal\quad \mathbf{w}_k=[w_k^\intercal\ v_k^\intercal]^\intercal, \tag{3.5}$$

$$\mathrm{A}_{\mathrm{cl}}(i,\ell)=\begin{bmatrix}A(i)+B(i)M(\ell)C(i) & B(i)L(\ell)\\ K(\ell)C(i) & F(\ell)\end{bmatrix},$$
$$\mathrm{H}_{\mathrm{cl}}(i,\ell)=\begin{bmatrix}H(i) & B(i)M(\ell)J(i)\\ 0 & K(\ell)J(i)\end{bmatrix}.$$

Remark 1. Assumptions and notations established describe a minimal set of properties involved in the system. As in (DRAGAN; COSTA, 2016a) the ergodicity hypothesis is not taken into account,in particular, results apply for MJLS with periodic Markov chains unlike for example (DOLGOV; CHLEBEK; HANEBECK, 2016).

3.2 Associated functionals and operators

3.2.1 Second moment and related results

We start defining the second moment matrix for our system using (3.5),

$$\mathscr{X}_k(i)=E[\mathbf{x}_k\mathbf{x}_k^\intercal\mathbb{1}_{\{\theta_k=i\}}],\quad \forall i\in\Theta,\ \forall k\geq 0. \tag{3.6}$$

Note that $\mathscr{X} \in \mathcal{S}_{r_x}^{\mathcal{N}}$, where $r_x = (n+n_c)\times(n+n_c)$. Based on (3.5) we consider the partitioned form

$$\mathscr{X}_k(i) = \begin{bmatrix} \mathscr{X}_k(i)_1 & \mathscr{X}_k(i)_{12} \\ (\mathscr{X}_k(i)_{12})^\intercal & \mathscr{X}_k(i)_2 \end{bmatrix} \tag{3.7}$$

$$= \begin{bmatrix} E[x_k x_k^\intercal \mathbb{1}_{\{\theta_k=i\}}] & E[x_k \hat{x}_k^\intercal \mathbb{1}_{\{\theta_k=i\}}] \\ E[\hat{x}_k x_k^\intercal \mathbb{1}_{\{\theta_k=i\}}] & E[\hat{x}_k \hat{x}_k^\intercal \mathbb{1}_{\{\theta_k=i\}}] \end{bmatrix}. \tag{3.8}$$

We will also present an adaptation of (COSTA; FRAGOSO; MARQUES, 2005, Proposition 3.35) for our specific case, using (3.5),(3.4) and using previous definition of second moment (3.6). We state the following,

Lemma 1 (Adapted from (COSTA; FRAGOSO; MARQUES, 2005))**.**

$$\mathscr{X}_{k+1}(i) = \sum_{j\in\Theta,\ell\in\Gamma} p_{ji}q_{\ell,j}\left[\mathrm{A}_{\mathrm{cl}}(j,\ell)\mathscr{X}_k(j)\mathrm{A}_{\mathrm{cl}}(j,\ell)^\intercal + \mathrm{H}_{\mathrm{cl}}(j,\ell)\mathrm{S}\mathrm{H}_{\mathrm{cl}}(j,\ell)^\intercal \pi_k(j)\right] \tag{3.9}$$

Proof.

$$E[(\mathbf{x}_{k+1})(\mathbf{x}_{k+1})^\intercal \mathbb{1}_{\{\theta_{k+1}=i\}}] =$$

$$\sum_{j\in\Theta,\ell\in\Gamma} E\left[(\mathrm{A}_{\mathrm{cl}(\theta_k,\gamma_k)}\mathbf{x}_k + \mathrm{H}_{\mathrm{cl}(\theta_k,\gamma_k)}\mathbf{w}_k)(\mathrm{A}_{\mathrm{cl}(\theta_k,\gamma_k)}\mathbf{x}_k + \mathrm{H}_{\mathrm{cl}(\theta_k,\gamma_k)}\mathbf{w}_k)^\intercal \mathbb{1}_{\{\theta_{k+1}=i\},\{\theta_k=j\}}\mathbb{1}_{\{\gamma_{k+1}=\ell\}}\right] =$$

Recalling that θ_{k+1},γ_{k+1} and x_{k+1} are conditional independent events, $\mathcal{P}(\theta_{k+1}=i,\theta_k=j,\gamma_k=\ell) = \mathcal{P}(\theta_{k+1}=i|\theta_k=j,\gamma_k=\ell)\times\mathcal{P}(\gamma_k=\ell|\theta_k=i)\times\mathcal{P}(\theta_k=i) = \mathcal{P}(\theta_k=j)p_{ji}\,q_{\ell,j} = E[\mathbb{1}_{\{\theta_k=j\}}]p_{ji}\,q_{\ell,j}$

$$= \sum_{j\in\Theta,\ell\in\Gamma} E\left[(\mathrm{A}_{\mathrm{cl}}(j,\ell)\mathbf{x}_k + \mathrm{H}_{\mathrm{cl}}(j,\ell)\mathbf{w}_k)(\mathrm{A}_{\mathrm{cl}}(j,\ell)\mathbf{x}_k + \mathrm{H}_{\mathrm{cl}}(j,\ell)\mathbf{w}_k)^\intercal \mathbb{1}_{\{\theta_k=j\}}\right]p_{ji}q_{\ell,j}$$
$$= \sum_{j\in\Theta,\ell\in\Gamma} \left[(\mathrm{A}_{\mathrm{cl}}(j,\ell)E[\mathbf{x}_k\mathbf{x}_k^\intercal\mathbb{1}_{\{\theta_k=j\}}]\mathrm{A}_{\mathrm{cl}}(j,\ell)^\intercal + \mathrm{H}_{\mathrm{cl}}(j,\ell)E[\mathbf{w}_k\mathbf{w}_k^\intercal\mathbb{1}_{\{\theta_k=j\}}]\mathrm{H}_{\mathrm{cl}}(j,\ell)^\intercal\right]p_{ji}$$

using previous definition of second moment (3.6) and the fact that $E[\mathbf{w}_k\mathbf{w}_k^\intercal\mathbb{1}_{\{\theta_k=j\}}] = E[\mathbf{w}_k\mathbf{w}_k^\intercal]\mathcal{P}(\theta_k = j) = \mathrm{S}\pi_k(j)$ where,

$$\mathbf{w}_k\mathbf{w}_k^\intercal = [w_k^\intercal\ v_k^\intercal]^\intercal[w_k^\intercal\ v_k^\intercal] = \begin{bmatrix} w_k w_k^\intercal & w_k v_k^\intercal \\ v_k w_k^\intercal & v_k v_k^\intercal \end{bmatrix} \tag{3.10}$$

$$\mathrm{S} = \begin{bmatrix} \mathcal{W} & 0 \\ 0 & \mathcal{V} \end{bmatrix} \tag{3.11}$$

That gives us,

$$= \sum_{j\in\Theta,\ell\in\Gamma} p_{ji}q_{\ell,j}\{\mathrm{A}_{\mathrm{cl}}(j,\ell)\mathscr{X}_k(i)\mathrm{A}_{\mathrm{cl}}(j,\ell)^\intercal + \mathrm{H}_{\mathrm{cl}}(j,\ell)\mathrm{S}\mathrm{H}_{\mathrm{cl}}(j,\ell)^\intercal\pi_k(j)\} \tag{3.12}$$

which completes the proof $\square$

3.2.2 Modified Lyapunov type operator

We present a modified lyapunov operator using (2.4)(2.5) and based on the coefficients of the linear system (3.4) and the elements of the stochastic matrix P and P_α the detector process transition probability, we introduce on the space $\mathcal{S}_d^{\mathcal{N}}$ the linear operator $\mathcal{L}_{cl}$ defined by:

$$(\mathcal{L}_{cl}\mathbf{V})(i) = \sum_{j\in\Theta,\ell\in\Gamma} p_{ji}q_{\ell,j}\mathrm{A}_{\mathrm{cl}}(j,\ell)V(j)\mathrm{A}_{\mathrm{cl}}(j,\ell)^{\intercal}, \quad 1\le i\le\mathcal{N}, \tag{3.13}$$

for all $\mathbf{V}=(V(1),\ldots,V(N))\in\mathcal{S}_d^{\mathcal{N}}$. By direct calculation one obtains that the adjoint operator of $\mathcal{L}_{cl}$ with respect to the inner product (2.1) is described by:

$$(\mathcal{L}_{cl}^{*}\mathbf{V})(i) = \sum_{j\in\Theta,\ell\in\Gamma} p_{ij}q_{\ell,i}\mathrm{A}_{\mathrm{cl}}(i,\ell)^{\intercal}V(j)\mathrm{A}_{\mathrm{cl}}(i,\ell), \quad 1\le i\le\mathcal{N}. \tag{3.14}$$

3.3 The long run average cost $\mathcal{J}$

Given the new closed-loop system, a new cost functional (3.3) can be written,

$$\mathbf{J} = \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}E\left[\sum_{k=0}^{\mathcal{K}}\left[x_k^{\intercal}Q_{(\theta_k)}x_k+u_k^{\intercal}R_{(\theta_k)}u_k\right]\right]$$

Using the linearity of the expectation operator and replacing output function fo the controller (3.2),

$$\begin{aligned}
&E\left[\, x_k^{\intercal}Q_{(\theta_k)}x_k+u_k^{\intercal}R_{(\theta_k)}u_k \,\right] = \\
&E\left[\, x_k^{\intercal}Q_{(\theta_k)}x+(L_{(\gamma_k)}\,\hat{x}_k+M_{(\gamma_k)}(C_{(\theta_k)}x_k+J_{(\theta_k)}v_k))^{\intercal}\, R_{(\theta_k)}\,(L_{(\gamma_k)}\,\hat{x}_k+M_{(\gamma_k)}(C_{(\theta_k)}x_k+J_{(\theta_k)}v_k))\,\right] = \\
&= E\left[x_k^{\intercal}Q_{(\theta_k)}x_k+(\hat{x}_k^{\intercal}L_{(\gamma_k)}^{\intercal}+x_k^{\intercal}C_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}+v_k^{\intercal}J_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal})\,R_{(\theta_k)}\,(L_{(\gamma_k)}\,\hat{x}_k+M_{(\gamma_k)}C_{(\theta_k)}x_k+M_{(\gamma_k)}J_{(\theta_k)}v_k)\,\right] \\
&= E\left[x_k^{\intercal}Q_{(\theta_k)}x_k+\hat{x}_k^{\intercal}L_{(\gamma_k)}^{\intercal}R_{(\theta_k)}L_{(\gamma_k)}\hat{x}_k+\hat{x}_k^{\intercal}L_{(\gamma_k)}^{\intercal}RM_{(\gamma_k)}C_{(\theta_k)}x_k+\hat{x}_k^{\intercal}L_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}J_{(\theta_k)}v_k+\right. \\
&\quad +x_k^{\intercal}C_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}L_{(\gamma_k)}\,\hat{x}_k+x_k^{\intercal}C_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}C_{(\theta_k)}x_k+x_k^{\intercal}C_{(\theta_k)}^{\intercal}M^{\intercal}RMJ_{(\theta_k)}v_k+ \\
&\quad \left.+v_k^{\intercal}J_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}L_{(\gamma_k)}\,\hat{x}_k+v_k^{\intercal}J_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}C_{(\theta_k)}x_k+v_k^{\intercal}J_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}MJ_{(\theta_k)}v_k\right]
\end{aligned}$$

Based on the assumptions of the problem statement, it may be inductively verified that for each $k\ge 0$ the random vectors x_k,w_k and v_k are independent among them, therefore:

$$\begin{aligned}
&E\left[\hat{x}_k^{\intercal}L_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}J_{(\theta_k)}v_k\right] && E\left[x_k^{\intercal}C_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}J_{(\theta_k)}v_k\right] \\
&= E\left[\hat{x}_k^{\intercal}L_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}J_{(\theta_k)}\right]E\left[v_k\right]=0 && = E\left[x_k^{\intercal}C_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}J_{(\theta_k)}\right]E\left[v_k\right]=0
\end{aligned}$$

$$\begin{aligned}
&E\left[v_k^{\intercal}J_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}L_{(\gamma_k)}\,\hat{x}_k\right] && E\left[v_k^{\intercal}J_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}C_{(\theta_k)}x_k\right] \\
&= E\left[v_k^{\intercal}J_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}L_{(\gamma_k)}\right]E\left[\hat{x}_k\right]=0 && = E\left[v_k^{\intercal}J_{(\theta_k)}^{\intercal}M_{(\gamma_k)}^{\intercal}R_{(\theta_k)}M_{(\gamma_k)}C_{(\theta_k)}\right]E\left[[\hat{x}_k\right]=0
\end{aligned}$$

This allows us to obtain,

$$
\begin{aligned}
&E\left[\ x_k^\intercal Q_{(\theta_k)} x_k + u_k^\intercal R_{(\theta_k)} u_k\ \right] = \\
&E\left[\ (x_k^\intercal Q_{(\theta_k)} x + \hat{x}_k^\intercal L_{(\gamma_k)}^\intercal R_{(\theta_k)} L_{(\gamma_k)} \hat{x}_k + \hat{x}_k^\intercal L_{(\gamma_k)}^\intercal R_{(\theta_k)} M_{(\gamma_k)} C_{(\theta_k)} x_k + x_k^\intercal C_{(\theta_k)}^\intercal M_{(\gamma_k)}^\intercal R_{(\theta_k)} L_{(\gamma_k)}\ \hat{x}_k + \right. \\
&\left. + x_k^\intercal C_{(\theta_k)}^\intercal M_{(\gamma_k)}^\intercal R_{(\theta_k)} M_{(\gamma_k)} C_{(\theta_k)} x_k) + v_k^\intercal J_{(\theta_k)}^\intercal M_{(\gamma_k)}^\intercal R_{(\theta_k)} M_{(\gamma_k)} J_{(\theta_k)} v_k\ \right]
\end{aligned}
\tag{3.15}
$$

and use (3.5) to rewrite,

$$
\begin{aligned}
&E\left[\ x_k^\intercal Q_{(\theta_k)} x_k + u_k^\intercal R_{(\theta_k)} u_k\ \right] = \\
&= E\left[\begin{bmatrix} x_k^\intercal \hat{x}_k^\intercal \end{bmatrix} \begin{bmatrix} Q_{(\theta_k)} + C_{(\theta_k)}^\intercal M_{(\gamma_k)}^\intercal R_{(\theta_k)} M_{(\gamma_k)} C_{(\theta_k)} & C_{(\theta_k)}^\intercal M_{(\gamma_k)}^\intercal R_{(\theta_k)} L_{(\gamma_k)} \\ L_{(\gamma_k)}^\intercal R_{(\theta_k)} M_{(\gamma_k)} C_{(\theta_k)} & L_{(\gamma_k)}^\intercal R_{(\theta_k)} L_{(\gamma_k)} \end{bmatrix} \begin{bmatrix} x_k \\ \hat{x}_k \end{bmatrix} + v_k^\intercal \left[J^\intercal R_{(\theta_k)} M_{(\gamma_k)} J \right] v_k \right] \\
&= E\left[\mathbf{x}_k^\intercal \tilde{Q}_{(\theta_k)} \mathbf{x}_k + v_k^\intercal [J^\intercal R_{(\theta_k)} M_{(\gamma_k)} J] v_k \right]
\end{aligned}
$$

If we take the second term and use properties of the trace operator,his linearity,and commute it with the expectation operator:

$$
\begin{aligned}
E\left[v_k^\intercal J_{(\theta_k)}^\intercal M_{(\gamma_k)}^\intercal R_{(\theta_k)} M_{(\gamma_k)} J_{(\theta_k)} v_k \right] &= \sum_{i\in\Theta,\ell\in\Gamma} tr\left[\ E[J(i)^\intercal M(\ell)^\intercal R(i) M(\ell) J(i) v_k v_k^\intercal] \mathbb{1}_{\{\theta_k=i\}} \mathbb{1}_{\{\ \gamma_k=\ell\}}\ \right] \\
&= \sum_{i\in\Theta,\ell\in\Gamma} tr\left[J(i)^\intercal M(\ell)^\intercal R(i) M(\ell) J(i)\ E[v_k v_k^\intercal \mathbb{1}_{\{\theta_k=i\}}] \mathbb{1}_{\{\ \gamma_k=\ell\}} \right]
\end{aligned}
$$

If we take into account that,

$$E[v_k v_k^\intercal \mathbb{1}_{\{\theta_k=i\}}] = E[v_k v_k^\intercal] \mathcal{P}(\theta(t)=i) \times \mathcal{P}(\theta(t)=i)$$

using problem assumptions we have, $E[v_k v_k^\intercal] = V$ so,

$$E[v_k v_k^\intercal \mathbb{1}_{\{\theta_k=i\}}] = V \pi_k(i)$$

Resulting in:

$$E\left[v_k^\intercal J_{(\theta_k)}^\intercal M(\ell)^\intercal R(i) M(\ell) J_{(\theta_k)} v_k \right] = \sum_{i\in\Theta,\ell\in\Gamma} \pi_k(i)\ q_{\ell,i}\ tr\left[M(\ell) J(i) V J(i)^\intercal M(\ell)^\intercal R(i) \right] \tag{3.16}$$

In intern product notation,

$$E\left[v_k^\intercal J_{(\theta_k)}^\intercal M(\ell)^\intercal R(i) M(\ell) J_{(\theta_k)} v_k \right] = \langle \mathbf{R}^\intercal \mathbf{M} \mathbf{J}, \mathbf{M} \mathbf{J} \mathbf{V} \rangle \tag{3.17}$$

rewritting the first term

$$\begin{aligned}
E\left[\mathbf{x}_k^\mathsf{T}\mathrm{Q}_{\mathrm{cl}(\theta_k,\gamma_k)}\mathbf{x}_k\right] &= \sum_{i\in\Theta,\ell\in\Gamma}\left[\,E[\mathbf{x}_k^\mathsf{T}\mathrm{Q}_{\mathrm{cl}}(i,\ell)\mathbf{x}_k\,\mathbb{1}_{\{\theta_k=i\}}]\right]\\
&= \sum_{i\in\Theta} E\left[tr[\mathrm{Q}_{\mathrm{cl}}(i,\ell)\;(\mathbf{x}_k\mathbf{x}_k^\mathsf{T})]\mathbb{1}_{\{\theta_k=i\}}\right]\\
&= \sum_{i\in\Theta} tr\left[\mathrm{Q}_{\mathrm{cl}}(i,\ell)\;E[\mathbf{x}_k\mathbf{x}_k^\mathsf{T}\mathbb{1}_{\{\theta_k=i\}}]\right]
\end{aligned}$$

using the definition of Second Moment of $\mathbf{x}_k$ at mode i, we have

$$= \sum_{i\in\Theta,\ell\in\Gamma} tr\left[\mathrm{Q}_{\mathrm{cl}}(i,\ell)\;\mathscr{X}_k(i)\right]$$

That is,

$$E\left[\mathbf{x}_k^\mathsf{T}\mathrm{Q}_{\mathrm{cl}}(i,\ell)\mathbf{x}_k\right] = \sum_{i\in\Theta,\ell\in\Gamma} tr\left[\mathrm{Q}_{\mathrm{cl}}(i,\ell)\;\mathscr{X}_k(i)\right] \tag{3.18}$$

In intern product notation,

$$E\left[\mathbf{x}_k^\mathsf{T}\mathrm{Q}_{\mathrm{cl}(\theta_k,\gamma_k)}\mathbf{x}_k\right] = \langle\mathbf{Q},\mathscr{X}\rangle \tag{3.19}$$

Combining (3.18) and (3.16) to obtain the value performance (3.3) the result states as follows,

Proposition 2. The performance index (3.3), for the system (3.4) can be expressed as

$$\begin{aligned}
\mathcal{J} = \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\sum_{k=0}^{\mathcal{K}}&\Bigg[\sum_{i\in\Theta,\ell\in\Gamma} tr[\mathrm{Q}_{\mathrm{cl}}(i,\ell)\mathscr{X}_k(i)]\\
&+\sum_{i\in\Theta,\ell\in\Gamma}\pi_k(i)q_{\ell,i}\,tr\left[M(\ell)J(i)VJ(i)^\mathsf{T}M(\ell)^\mathsf{T}R(i)\right]\Bigg].
\end{aligned} \tag{3.20}$$

where,

$$\mathrm{Q}_{\mathrm{cl}}(i,\ell) = \begin{bmatrix} Q(i)+C^\mathsf{T}(i)M(\ell)^\mathsf{T}R(i)M(\ell)C(i) & C^\mathsf{T}(i)M(\ell)^\mathsf{T}R(i)L(\ell)\\ L(\ell)^\mathsf{T}R(i)M(\ell)C(i) & L(\ell)^\mathsf{T}R(i)L(\ell)\end{bmatrix}$$

If a limiting $\mathscr{X} = \lim_{k\to\infty}\mathscr{X}_k$ exists, then

$$\mathcal{J} = \langle\mathcal{Q},\mathscr{X}\rangle + \langle\mathbf{R}^\mathsf{T}\mathbf{MJ},\mathbf{MJ}V\rangle \tag{3.21}$$

where $\mathcal{Q}\in\mathcal{M}_p^{\mathcal{N}}$, $\mathbf{R}\in\mathcal{M}_p^{\mathcal{N}}$, $\mathbf{J}\in\mathcal{M}_{m\times r}^{\mathcal{N}}$, $\mathbf{M}\in\mathcal{M}_{p\times m}^{\mathscr{L}}$, and $V\in\mathcal{M}_{p\times m}$ satisfy

$$\begin{aligned}
&\mathbf{R} = (R(1),\ldots,R(\mathcal{N})),\;\mathbf{J} = (J(1),\ldots,J(\mathcal{N})),\\
&\mathcal{Q} = (\mathrm{Q}_{\mathrm{cl}}(1),\ldots,\mathrm{Q}_{\mathrm{cl}}(\mathcal{N}))\;\mathbf{M} = (M(1),\ldots,M(\mathscr{L}))
\end{aligned}$$

3.4 A compact formulation for the cost

Although the cost formulation in (3.20) is quite compact, it is valid under the condition that the limiting $\mathscr{X}_k$ exists, which is not always the case – e.g. when the Markov chain is periodic. In this section we get around this inconvenience by working with a "time-average" of $\mathscr{X}_k$. We also develop an even more compact notation. Let us write

$$U(\ell) = \begin{bmatrix} F(\ell) & K(\ell) \\ L(\ell) & M(\ell) \end{bmatrix}$$

and $\mathrm{U} = (U(1), \ldots, U(\mathcal{N})) \in \mathcal{M}_{r_{\mathrm{U}} \times r_{\mathrm{U}}}$, where $r_{\mathrm{U}} = (n_c + p) \times (n_c + m)$. Let the operator $\mathbf{F} : \mathcal{S}_d^{\mathcal{N}} \times \mathcal{M}_{r_{\mathrm{U}}}^{\mathscr{L}} \to \mathcal{S}_{d\mathcal{N}}^{\mathcal{N}}$ be such that, for $V \in \mathcal{M}_{r\mathcal{N}}$, we have

$$\mathbf{F}(\mathbf{V}, \mathrm{U}) = \mathrm{diag}\Big((\mathcal{L}_{cl}\mathbf{V})(1) \ldots, (\mathcal{L}_{cl}\mathbf{V})(\mathcal{N}) \Big),$$

We also define its adjoint operator $\mathbf{G} : \mathcal{S}_d^{\mathcal{N}} \times \mathcal{M}_{r_{\mathrm{U}}}^{\mathscr{L}} \to \mathcal{S}_{d\mathcal{N}}^{\mathcal{N}}$

$$\mathbf{G}(\mathbf{V}, \mathrm{U}) = \mathrm{diag}\Big((\mathcal{L}_{cl}^{\star}\mathbf{V})(1), \ldots, (\mathcal{L}_{cl}^{\star}\mathbf{V})(\mathcal{N}) \Big),$$

Note that the controller matrix U is implicit in the closed-loop operators in the above expressions. Along the lines of the above definitions, we write

$$\begin{aligned} \Sigma = \mathrm{diag}\Bigg(&\sum_{j\in\Theta,\ell\in\Gamma} p_{j1} q_{\ell,j} \rho_j T(j,\ell), \ldots \\ &\ldots, \sum_{j\in\Theta,\ell\in\Gamma} p_{j\mathcal{N}} q_{\ell,j} \rho_j T(j,\ell) \Bigg) \end{aligned} \tag{3.22}$$

and

$$\begin{aligned} \Sigma_k = \mathrm{diag}\Bigg(&\sum_{j\in\Theta,\ell\in\Gamma} p_{j1} \pi_j(k) q_{\ell,j} T(j,\ell), \ldots \\ &\cdots, \sum_{j\in\Theta,\ell\in\Gamma} p_{j\mathcal{N}} \pi_j(k) q_{\ell,j} T(j,\ell) \Bigg). \end{aligned} \tag{3.23}$$

where we write $T(j,\ell) = \mathrm{H}_{\mathrm{cl}}(j,\ell)\ \mathrm{S}\ \mathrm{H}_{\mathrm{cl}}(j,\ell)^{\mathsf{T}}$ for convenience. Also,

$$\begin{aligned} \mathrm{Q} &= \mathrm{diag}(\mathrm{Q}_{\mathrm{cl}}(1), \ldots, \mathrm{Q}_{\mathrm{cl}}(\mathcal{N})) \\ \mathrm{Z}_k &= \mathrm{diag}(\mathscr{X}_k(1), \ldots, \mathscr{X}_k(\mathcal{N})) \end{aligned}$$

In terms of the operators and notation above, the result of Lemma 1 becomes simply

$$\mathrm{Z}_{k+1} = \mathbf{F}(\mathrm{Z}_k, \mathrm{U}) + \Sigma_k \tag{3.24}$$

In order to handle periodic states of the Markov chain,we define δ as the least common multiple of $\delta_1, \ldots, \delta_{\mathcal{N}}$, where δ_i, $i \in \Theta$, is the period of the Markov state i (or $\delta_i = 1$ if i is aperiodic), and consider

$$\begin{aligned} \mathrm{X}_k &= \frac{1}{\delta}(\mathrm{Z}_k + \mathrm{Z}_{k+1} + \cdots + \mathrm{Z}_{k+\delta-1}), \\ \bar{\pi}_i(k) &= \frac{1}{\delta}(\pi_i(k) + \pi_i(k+1) + \cdots + \pi_i(k+\delta-1)). \end{aligned}$$

Note that

$$
\begin{aligned}
\mathrm{X}_{k+1} &= \frac{1}{\delta}(\mathbf{F}(\mathrm{Z}_k,\mathrm{U})+\mathbf{F}(\mathrm{Z}_{k+1},\mathrm{U})+\cdots+\mathbf{F}(\mathrm{Z}_{k+\delta-1},\mathrm{U})) \\
&\quad +\Sigma_k+\Sigma_{k+1}+\ldots+\Sigma_{k+\delta-1} \\
&= \mathbf{F}(\mathrm{X}_k,\mathrm{U})+\bar{\Sigma}_k
\end{aligned}
\tag{3.25}
$$

where

$$
\begin{aligned}
\bar{\Sigma}_k = \mathrm{diag}\Big(& \sum_{j\in\Theta,\ell\in\Gamma} p_{j1}q_{\ell,j}T(j,\ell)\bar{\pi}_j(k),\ldots \\
& \ldots,\sum_{j\in\Theta,\ell\in\Gamma} p_{j\mathcal{N}}q_{\ell,j}T(j,\ell)\bar{\pi}_j(k)\Big)
\end{aligned}
$$

The "time-average second moment variable" X_k has as an intrinsic advantage over $\mathcal{X}_k$ that $\bar{\pi}_i(k)$ converges (exponentially fast) to a limiting ρ (recall (2.2)) as $k\to\infty$, as opposed to $\pi_i(k)$ that fails to converge in the case when the Markov chain is periodic. When the initial distribution of $\{\theta_k\}$ coincides with the Cesaro limit, that is, $\pi_0=\rho=\lim_{k\to\infty}\bar{\pi}(k)$ it is simple to see that X_k is governed by a stationary version of (3.25):

$$
\mathrm{X}_{k+1}=\mathbf{F}(\mathrm{X}_k,\mathrm{U})+\Sigma. \tag{3.26}
$$

Hence, for any U that yields convergence of X_k, the limiting X_∞ is a fixed point solution of the above equation,

$$
\mathrm{X}_\infty=\mathbf{F}(\mathrm{X}_\infty,\mathbf{U})+\Sigma. \tag{3.27}
$$

Moreover, when $\pi_0\neq\rho$, the fact that $\bar{\pi}_i(k)$ converges exponentially fast to ρ can be used to show that the limit X_∞ is unaltered. It is worth mentioning that, even if X_k converges, both $\mathscr{X}_k$ and Z_k may fail to converge because of the periodicity of the Markov chain, preventing us from using the result given in Proposition 2.

Having established a suitable framework for the general Markov chain scenario, we rewrite the cost,

$$
\begin{aligned}
\mathcal{J} &= \check{R}+\lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\sum_{k=0}^{\mathcal{K}}\Big[\mathrm{tr}(\mathrm{QZ}_k)\Big] \\
&= \check{R}+\lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\Big\{-\delta^{-1}\mathrm{tr}\big(\mathrm{Q}(\mathrm{Z}_{\mathcal{K}+1}+\cdots+\mathrm{Z}_{\mathcal{K}+\delta-1})\big) \\
&\quad +\sum_{k=0}^{\mathcal{K}}\Big[\mathrm{tr}(\mathrm{Q}\mathbf{X_k})\Big]\Big\}.
\end{aligned}
\tag{3.28}
$$

where

$$
\check{R}=\sum_{i\in\Theta,\ell\in\Gamma} q_{i,\ell}\rho_i\,\mathrm{tr}\big[M(\ell)J(i)VJ(i)^\intercal M(\ell)^\intercal R(i)\big]
$$

For any U for which there is a corresponding limiting X_∞, it is simple to show that the terms $\mathrm{Z}_{\mathcal{K}+j}$, $0\leq j\leq\delta-1$ are bounded in the sense that exists some $\bar{Z}$ such that $\mathrm{Z}_j\leq\bar{Z}$ for large enough j (otherwise the average X_k would not converge); this yields

$$\begin{aligned}\mathcal{J} = \check{R} + \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\Big\{ -\delta^{-1}\mathrm{tr}\big(\mathrm{Q}(\mathrm{Z}_{\mathcal{K}+1}+\cdots+\mathrm{Z}_{\mathcal{K}+\delta-1})\big)+ \\ +\sum_{k=0}^{\mathcal{K}}\Big[\mathrm{tr}(\mathrm{QX}_k)\Big]\Big\} \leq \check{R} + \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\Big\{ -\delta^{-1}\mathrm{tr}\big(\mathrm{Q}(\bar{Z}+\cdots+\bar{Z})\big)+ \\ +\sum_{k=0}^{\mathcal{K}}\Big[\mathrm{tr}(\mathrm{QX}_k)\Big]\Big\} = \check{R} + \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\left\{\sum_{k=0}^{\mathcal{K}}\Big[\mathrm{tr}(\mathrm{QX}_k)\Big]\right\}\end{aligned} \tag{3.29}$$

A similar evaluation follows by using the lower bound $\mathrm{Z}_j \geq 0$, leading to, $\mathcal{J} \geq \check{R} + \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\left\{\sum_{k=0}^{\mathcal{K}}\Big[\mathrm{tr}(\mathrm{QX}_k)\Big]\right\}$; this and (3.29) provide

$$\mathcal{J} = \check{R} + \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\left\{\sum_{k=0}^{\mathcal{K}}\Big[\mathrm{tr}(\mathrm{QX}_k)\Big]\right\}. \tag{3.30}$$

Consider an arbitrary time instant $\iota \geq 0$. Rearranging terms and taking into account that the matrix $\mathbf{F}_\iota(\mathrm{X},\mathrm{U})$ converges (exponentially fast) to $\mathbf{F}(\mathrm{X},\mathrm{U})$ we write:

$$\begin{aligned}\mathcal{J} &= \check{R} + \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\Big\{\Big[\mathrm{tr}\big(\mathrm{QX}_0+\cdots+\mathrm{QX}_\iota+ \\ &+\mathrm{Q}(\mathbf{F}(\mathrm{X}_\iota,\mathrm{U})+\Sigma)+\cdots+\mathrm{Q}(\mathbf{F}^{\mathcal{K}-\iota}(\mathrm{X}_\iota,\mathrm{U})+ \\ &+\mathbf{F}^{\mathcal{K}-\iota-1}(\Sigma,\mathrm{U})+\cdots+\mathbf{F}(\Sigma,\mathrm{U})+\Sigma)\big)\Big]+O(\iota)\Big\} \\ &= \check{R} + \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\Big\{\mathrm{tr}\Big[\big(\mathrm{QX}_0+\cdots+\mathrm{QX}_\iota+ \\ &+(\mathbf{G}(\mathbf{Q},\mathrm{U})\mathrm{X}_\iota+\mathrm{Q}\Sigma)+\cdots+\mathbf{G}^{\mathcal{K}-\iota}(\mathrm{Q},\mathrm{U})\mathrm{X}_\iota+ \\ &+(\mathbf{G}^{\mathcal{K}-\iota-1}(\mathrm{Q},\mathrm{U})+\cdots+\mathbf{F}(\mathrm{Q},\mathrm{U})+\mathrm{Q})\Sigma\big)\Big]+O(\iota)\Big\}.\end{aligned} \tag{3.31}$$

We have used a compact notation for composition of operators, e.g. $\mathbf{F}^2(\cdot,\mathrm{U})$ denotes $\mathbf{F}(\mathbf{F}(\cdot,\mathrm{U}),\mathrm{U})$. Inspired on the operator $\mathbf{G}$ the self adjoint of $\mathbf{F}$, we define the "co-state" P_k with the same dimensions as X_k satisfying,

$$\mathrm{P}_{k+1} = \mathbf{G}(\mathrm{P}_k,\mathrm{U})+\mathrm{Q}. \tag{3.32}$$

with initial condition $\mathrm{P}_0 = 0$. Note that e.g. $\mathrm{P}_1 = \mathrm{Q}$ and $\mathrm{P}_2 = \mathbf{G}(\mathrm{Q},\mathrm{U})+\mathrm{Q}$ with $\mathrm{P}_k \in \mathcal{S}_d^{\mathcal{N}(0)}$. Substituting this in (3.30) and rearranging terms,

$$\begin{aligned}\mathcal{J} &= \check{R} + \lim_{\mathcal{K}\to\infty}\frac{1}{\mathcal{K}}\Big\{\mathrm{tr}\Big[\big(\mathrm{Q}(\mathrm{X}_0+\cdots+\mathrm{X}_\iota) \\ &+(\mathrm{P}_{\mathcal{K}}-\mathrm{Q})\mathrm{X}_\iota+(\mathrm{P}_1+\mathrm{P}_2+\cdots+\mathrm{P}_{\mathcal{K}-\iota})\Sigma\Big]+O(\iota)\Big\}.\end{aligned} \tag{3.33}$$

Taking ι as the dividend of $\mathcal{K}/2$ in the above, and taking limits,

$$\mathcal{J} = \check{R} + \mathrm{tr}\big((1/2)\mathrm{QX}_\infty + (1/2)\mathrm{P}_\infty\Sigma\big). \tag{3.34}$$

Different formulas can be obtained by rearranging the terms in (3.33), e.g. if ι is fixed then $\mathcal{J}(\mathrm{U}) = \check{R} + \mathrm{tr}(\mathrm{P}_\infty \Sigma)$. With $\iota = \kappa - 1$ we have $\mathcal{J}(\mathrm{U}) = \check{R} + \mathrm{tr}(\mathrm{Q}\mathrm{X}_\infty)$. We collect the main results of this section in the following theorem,

Theorem 3. Consider a controller U for which $\mathrm{X}_\infty = \lim_{k\to\infty} \mathrm{X}_k$ and $\mathrm{P}_\infty = \lim_{k\to\infty} \mathrm{P}_k$ exist. Then: (i) Both P_∞ and X_∞ are irrespective of the initial conditions π_0, x_0 and (ii) The long run average cost $\mathcal{J}$ defined in (3.20) can be written as

$$\begin{aligned}
\mathcal{J} &= \check{R} + \mathrm{tr}((1/2)\mathrm{Q}\mathrm{X}_\infty + (1/2)\mathrm{P}_\infty \Sigma) \\
&= \check{R} + \mathrm{tr}(\mathrm{P}_\infty \Sigma) \\
&= \check{R} + \mathrm{tr}(\mathrm{Q}\mathrm{X}_\infty).
\end{aligned}$$

CHAPTER

4

AN ALGORITHM FOR THE LARC

Considering the results of Theorem 3, we study the optimization problem that consists of $\min_{\mathrm{U},\mathrm{X}_\infty,\mathrm{P}_\infty} \mathcal{J}$ subject to the constraint (3.27) and its dual counterpart

$$\mathrm{P}_\infty = \mathbf{G}(\mathrm{P}_\infty, \mathrm{U}) + \mathrm{Q}. \tag{4.1}$$

This is a highly nonlinear problem in the sense that, if we expand any of the formulas given in the theorem, exposing the hidden controller matrices, we will find terms involving multiplication of up to four of the variables of the optimization problem. Moreover, the relative high dimension and the fact that both X_∞ and P_∞ are frequently ill-conditioned matrices in real world applications, makes it difficult to obtain a high precision numerical solution by means of standard methods. In some cases it is reported that X_∞ and P_∞ leave the cone of positive semi-definite matrices (DOLGOV, 2017). Another complication is that X_∞ and P_∞ are restricted to the cone of positive semi-definite matrices $\mathcal{S}_d^{\mathcal{N}(+)}$ (which is considered throughout the paper).

Fortunately, if we fix X_∞ and P_∞, then the objective function (OF) comprises multiplications of no more than two of the variables of the optimization problem. Inspired by optimization methods that sequentially use an optimization operator and an feasibility operator, for problems where both operators are computationally inexpensive (Shaikh; Caines, 2007), and taking into account what we described before, it is natural to consider a method that minimize the OF in the variable U and then projects X_∞ and P_∞ onto the solution of (3.27) / (4.1). In the next sections we construct these operator, separately.

4.1 Two stage Optimization - Feassibility method

4.2 Optimization operator

Here we focus on the problem $\min_{\mathrm{U}} \mathcal{J}$ for given, fixed $\mathrm{X}_\infty, \mathrm{P}_\infty$. Although the equations given in Theorem 3 are useful for computing the long run average cost, (specially when one computes X_∞ and P_∞ via (3.27), (4.1)) In fact, by using the Kronecker product, one can transform (3.27) in a conventional linear system in the form $\mathscr{A}x = b$, and similarly for P_∞, leading to a relatively inexpensive way to compute the cost for a given U. However, we shall modify the OF to obtain some useful properties that will be used later. First we use the equivalences in Theorem 3 to obtain $\mathrm{tr}(P_\infty\Sigma - QX_\infty) = 0$, then we subtract this term from the first equality given in the theorem, yielding

$$\mathcal{J} = \check{R} + \mathrm{tr}((3/2)QX_\infty - (1/2)P_\infty\Sigma). \tag{4.2}$$

One can check via the definition of Q and Σ that the terms of the above OF contain the blocks M and K of U, only; in a descent-like method we seek for, which updates U based on the gradient of the OF, this would cause entire blocks of U to be fixed along iterations. To avoid this, we substitute (3.27) into (4.41) to get

$$\mathcal{J}(\mathrm{U}) = \check{R} + \mathrm{tr}((3/2)QX_\infty + (1/2)P_\infty(\mathbf{F}(X_\infty,\mathrm{U}) - X_\infty).$$

The optimization problem can now be written as

$$\begin{aligned} \min_{\mathrm{U}} \ & \check{R} + \mathrm{tr}((3/2)QX_\infty + (1/2)P_\infty(\mathbf{F}(X_\infty,\mathrm{U}) - X_\infty) \\ \text{s.t.} \quad & X_\infty = \mathbf{F}(X_\infty,\mathrm{U}) + \Sigma, \\ & P_\infty = \mathbf{G}(P_\infty,\mathrm{U}) + Q. \end{aligned} \tag{4.3}$$

By expanding the OF (4.3) and introducing a convenient notation that makes the dependence on U explicit, one can write (4.4)

$$\begin{aligned} 2\mathcal{J}(\mathrm{U}) = \sum_{i\in\Theta} \Bigg\{ & 2\,tr\left(\rho(i)\begin{bmatrix} 0_{p\times n_c} & I_p \\ 0_{p\times n_c} & 0_p \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^\mathsf{T}(i) \end{bmatrix} \mathrm{U}^\mathsf{T} \begin{bmatrix} 0_{n_c\times p} & 0_{n_c\times p} \\ R(i) & 0_p \end{bmatrix}\right) + tr\Bigg[\mathcal{E}(i)\Bigg(\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} + \\ & + \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\Bigg) \mathscr{X}_\infty(i) \Bigg(\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\Bigg)^\mathsf{T}\Bigg] + \\ & + 3\,tr\left(\begin{bmatrix} Q(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \mathrm{U}^\mathsf{T} \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\right) \mathscr{X}_\infty(i) - tr\Big(\mathcal{P}_\infty(i)\mathscr{X}_\infty(i)\Big)\Bigg\} \end{aligned} \tag{4.4}$$

where $\mathscr{X}_\infty(i)$ is the $i-$th block-diagonal element of X_∞, and similarly for the $i-$th block-diagonal $\mathcal{P}_\infty(i)$ of P_∞; and $\mathcal{E}(i) = \sum_{j\in\Theta} p_{ij}\mathcal{P}_\infty(j)$ is introduced to ease notation. Note in (33) that all blocks of U are relevant. It is also interesting that no more than two variables U appear in each sum, so that the cost is "quadratic" in U. In order to avoid possible saddle points, we need a lower bound for the objective function, which follows from the fact that

$$\mathcal{J}(\mathrm{U}) + \mathrm{tr}((1/2)P_\infty X_\infty) = \check{R} + \mathrm{tr}((3/2)QX_\infty + (1/2)P_\infty\mathbf{F}(X_\infty,\mathrm{U})).$$

is a non-negative number, leading to $\mathcal{J}(\mathrm{U}) \geq -(1/2)\mathrm{tr}(P_\infty X_\infty)$ and

$$\mathcal{J}(\mathrm{U}) \geq -(1/2)n^2\|P_\infty\|\|X_\infty)\|, \tag{4.34}$$

explaining why we have adopted (4.2). We are now in position to state the main result of this section.

Lemma 4. $\mathcal{J}(\mathrm{U})$ defined in (4.4) has at least one finite norm global minimum.

Proof. Consider a matrix norm for U, denoted by $\|\mathrm{U}\|$, the set $\mathcal{B}_r = \{\mathrm{U} : \|\mathrm{U}\| = r\}$, and define the function $h : \mathcal{B}_1 \to R$ by

$$h_{\mathrm{U}}(\mathrm{U}) = \min_{\alpha} \mathcal{J}(\alpha \mathrm{U}).$$

Using (4.4) we can write $\mathcal{J}(\alpha\mathrm{U}) = \mathcal{J}(0) + a_1(\mathrm{U})\alpha + a_2(\mathrm{U})\alpha^2$ where $a_1(\mathrm{U}), a_2(\mathrm{U})$ are scalar-valued functions, continuous in U, and we know that either $a_2(\mathrm{U}) > 0$ or $a_2(\mathrm{U}) = a_1(\mathrm{U}) = 0$ in view of (4.34), so its minimizer is $\alpha = -a_1(\mathrm{U})/a_2(\mathrm{U})$ when $a_2 > 0$ or any α when $a_2 = 0$. This implies that h_{U} is well defined for each U. One can check that there exists $\kappa > 0$, uniform in $\mathcal{B}_1$ such that $|a_1(\mathrm{U})| \leq \kappa |a_2(\mathrm{U})|$ (otherwise, the smaller is $|a_2|$ when compared with $|a_1|$, the smaller is the minimum of $\mathcal{J}(\alpha\mathrm{U})$, eventually violating the lower bound expressed in (4.34). This implies that the minimizer (obtaining by differentiating the above polynome) $\alpha = -a_1(\mathrm{U})/(2a_2(\mathrm{U}))$ is such that $|\alpha| < \kappa/2$, which allows to write

$$\begin{aligned}
\inf_{\mathrm{U}} \mathcal{J}(\mathrm{U}) &= \inf_{\mathrm{U}\in\mathcal{B}_1} h_{\mathrm{U}}(\mathrm{U}) = \inf_{\mathrm{U}\in\mathcal{B}_1} \min_{|\alpha|\leq\kappa/2} \mathcal{J}(\alpha\mathrm{U}) \\
&= \inf_{\mathrm{U}\in\mathcal{B}_{\kappa/2}} \min_{|\alpha|\leq 1/2} \mathcal{J}(\alpha\mathrm{U}) = \inf_{\mathrm{U}\in\mathcal{B}_{\kappa/2}} \mathcal{J}(\mathrm{U}).
\end{aligned}$$

Finally, note from Theorem 1 that $\mathcal{J}$ is continuous in U; then it follows from Weirstrass theorem that the inf on the right hand side of the above equation is realized by some minimizer $\mathrm{U} \in \mathcal{B}_{\kappa/2}$.

Lemma 4 allows us to define the optimality operator $\mathcal{O} : \mathcal{M}_{n_{\mathrm{U}}} \to \mathcal{M}_{n_{\mathrm{U}}}$ of the method as:

$$\mathcal{O}(\mathrm{U}) = \arg\min_{\mathrm{U}} J(\mathrm{U}). \tag{4.35}$$

Since the minimizer of $\mathcal{J}(\mathrm{U})$ always exists, one can compute it by taking the partial derivatives of $\mathcal{J}$ w.r.t. U and set it to zero ($\partial_{\mathrm{U}}\, \mathcal{J}(\mathrm{U}) = \mathbf{0}$). The result is as follows.

Theorem 5. The minimizer of $\mathcal{J}(\mathrm{U})$ is the solution of the matrix equation

$$\begin{aligned}
&\sum_{i\in\Theta} \left\{ 2\,\rho(i) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} + \gamma_L(i)\,\mathrm{U}\,\gamma_M(i) \right\} = \\
&-\sum_{i\in\Theta} \left\{ \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}(i) \begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \right\}
\end{aligned} \tag{4.36}$$

where,

$$\begin{aligned}
\gamma_L(i) &= 3 \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} + \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \\
\gamma_M(i) &= \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}
\end{aligned}$$

Proof. If we consider, the partitioned control matrix U, and we explicit the control term in every matrix component of the cost, we have

$$\mathrm{A_{cl}}(i,\ell) = \begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \tag{4.37}$$

$$\mathrm{H_{cl}}(i,\ell) = \begin{bmatrix} H(i) & 0_{n\times r} \\ 0_{n_c\times l} & 0_{n_c\times r} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times l} & 0_{n_c\times r} \\ 0_{m\times l} & J(i) \end{bmatrix} \tag{4.38}$$

$$\mathrm{Q_{cl}}(i,\ell) = \begin{bmatrix} Q(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \mathrm{U}^{\mathsf{T}} \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c\times n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \tag{4.39}$$

$$\mathcal{R}(i) = tr\left(\begin{bmatrix} 0_{p\times n_c} & I_p \\ 0_{p\times n_c} & 0_{p\times p} \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} \mathrm{U}^{\mathsf{T}} \begin{bmatrix} 0_{n_c\times p} & 0_{n_c\times p} \\ R(i) & 0_p \end{bmatrix} \right) \tag{4.40}$$

Recalling the cost functional,

$$\mathcal{J} = tr(\mathbf{R}) + tr((3/2)\mathbf{Q}\mathbf{X}_\infty + (1/2)\mathbf{P}_\infty(\mathbf{F}(\mathbf{X}_\infty,\mathbf{U}) - \mathbf{X}_\infty) \tag{4.41}$$

In that sense we can rewrite the cost,

$$\begin{aligned}
\mathcal{J} = &\sum_{i\in\Theta} tr\left(\rho(i) \begin{bmatrix} 0_{p\times n_c} & I_p \\ 0_{p\times n_c} & 0_{p\times p} \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} \mathrm{U}^{\mathsf{T}} \begin{bmatrix} 0_{n_c\times p} & 0_{n_c\times p} \\ R(i) & 0_p \end{bmatrix} \right) + \\
&+(3/2)\sum_{i\in\Theta} tr\left(\begin{bmatrix} Q(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} \mathscr{X}_\infty(i) + \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \mathrm{U}^{\mathsf{T}} \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} \mathrm{U} \cdot \right. \\
&\left. \cdot \begin{bmatrix} 0_{n_c\times n} & I_{n_c\times n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \right) + \\
&+(1/2)\sum_{i\in\Theta} tr\left[\mathcal{E}_i(P_\infty) \left(\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \right) \mathscr{X}_\infty(i) \cdot \right. \\
&\left. \cdot \left(\begin{bmatrix} A^{\mathsf{T}}(i) & 0_{n_c\times n} \\ 0_{n\times n_c} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \mathrm{U}^{\mathsf{T}} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \right) \right] + \\
&-(1/2)\sum_{i\in\Theta} tr\left(P_\infty(i)\mathscr{X}_\infty(i) \right)
\end{aligned} \tag{4.42}$$

For the extended cost form, $\mathcal{J}$, we have that for an optimal U,

$$\partial_{\mathrm{U}}\, \mathcal{J}(\mathrm{U}) = 0 \tag{4.43}$$

$$
\begin{aligned}
0 = \sum_{i\in\Theta} \Bigg(& \rho(i) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} + \\
& + (3/2) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \\
& + \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \\
& + \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \Bigg)
\end{aligned}
$$

$$
\begin{aligned}
0 = \sum_{i\in\Theta} \Bigg\{ & 2\,\rho(i) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} + \\
& + \left(3 \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix} + 2 \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \right) \cdot \\
& \cdot \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \\
& + 2 \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}
\end{aligned}
\tag{4.44}
$$

We take $R(i)$ as a symmetric positive definite matrix, in order to compare it with (DRAGAN; COSTA, 2016a) MJLS dynamic output feedback Riccatti approach.

$$
\begin{aligned}
0 = \sum_{i\in\Theta} \Bigg\{ & 2\,\rho(i) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} \mathrm{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} + \\
& + \left(3 \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} + \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \right) \cdot \\
& \cdot \mathrm{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \\
& + \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}
\end{aligned}
$$

We rename,

$$
\gamma_L(i) = 3 \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} + \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \tag{4.45}
$$

$$
\gamma_M(i) = \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \tag{4.46}
$$

Leading to the following matrix equation,

$$\sum_{i\in\Theta}\left\{2\,\rho(i)\begin{bmatrix}0_{n_c} & 0_{n_c\times p}\\ 0_{p\times n_c} & R(i)\end{bmatrix}\mathrm{U}\begin{bmatrix}0_{n_c} & 0_{n_c\times m}\\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i)\end{bmatrix}+\gamma_L(i)\,\mathrm{U}\,\gamma_M(i)\right\}=$$
$$=-\sum_{i\in\Theta}\left\{\begin{bmatrix}0_{n_c\times n} & I_{n_c}\\ B^{\mathsf{T}}(i) & 0_{p\times n_c}\end{bmatrix}\mathcal{E}_i(P_\infty)\begin{bmatrix}A(i) & 0_{n\times n_c}\\ 0_{n_c\times n} & 0_{n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix}0_{n\times n_c} & C^{\mathsf{T}}(i)\\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}\right\} \tag{4.47}$$

4.2.1 Matricial Equation for U

The equation presented in Theorem 5 can be transformed into a linear system in the form $\mathscr{A}_{\mathcal{O}}\cdot x=b$, with $\mathscr{A}_{\mathcal{O}}\in\mathcal{S}_d^N$ by using stack-column operators and the Kronecker product properties. In this system, the solution matrix is of dimension r_{U}. Another alternative is to write it as a general Sylvester equation and use a gradient-iterative or a least-square iterative method e.g (DING; CHEN, 2005a; DING; CHEN, 2005b).

We can expand the matrix equation obtained in (4.47), by using the partitioned form of U,$\mathscr{X}_\infty(i)$ and $P_\infty(i)$,

Recalling, that we defined for a partitionated $P_\infty(i)$ and $\mathscr{X}_\infty(i)$,

$$\mathscr{X}_\infty(i)=\begin{bmatrix}\mathscr{X}_1(i) & \mathscr{X}_{12}(i)\\ (\mathscr{X}_{12}(i))^T & \mathscr{X}_2(i)\end{bmatrix}\text{ and }P_\infty(i)=\begin{bmatrix}P_1(i) & P_{12}(i)\\ (P_{12}(i))^T & P_2(i)\end{bmatrix} \tag{4.48}$$

We define the operator,

$$\mathcal{E}_i(P_\infty)=\sum_{j\in\Theta}p_{ij}P_\infty(j) \tag{4.49}$$

and his partitioned version,

$$\mathcal{E}_i(P_\infty)=\begin{bmatrix}\mathcal{E}_1(i) & \mathcal{E}_{12}(i)\\ (\mathcal{E}_{12}(i))^T & \mathcal{E}_2(i)\end{bmatrix} \tag{4.50}$$

In that order we start with γ_L (4.45)

$$\gamma_L(i)=\begin{bmatrix}\mathcal{E}_2(i) & \mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\\ B^{\mathsf{T}}(i)\mathcal{E}_{12}(i) & \gamma_l(i)\end{bmatrix} \tag{4.51}$$

where,

$$\gamma_l(i)=3R(i)+B^{\mathsf{T}}(i)\mathcal{E}_1B(i) \tag{4.52}$$

and γ_M (4.46) form,

$$\gamma_M(i)=\begin{bmatrix}\mathscr{X}_2(i) & \mathscr{X}_{12}^{\mathsf{T}}(i)C^{\mathsf{T}}(i)\\ C(i)\mathscr{X}_{12}(i) & \gamma_m\end{bmatrix} \tag{4.53}$$

where

$$\gamma_m = C(i)\mathscr{X}_1(i)C^\intercal(i) \tag{4.54}$$

With the partitioned U, (4.51) and (4.53) replacing in (4.47)

$$\sum_{i\in\Theta}\left\{2\,\rho(i)\begin{bmatrix}0_n & 0_{n\times m}\\ 0_{p\times n} & R(i)MJ(i)VJ^\intercal(i)\end{bmatrix}+\begin{bmatrix}\mathcal{E}_2(i) & \mathcal{E}_{12}^\intercal(i)B(i)\\ B^\intercal(i)\mathcal{E}_{12}(i) & \gamma_l(i)\end{bmatrix}\mathrm{U}\begin{bmatrix}\mathscr{X}_2(i) & \mathscr{X}_{12}^\intercal(i)C^\intercal(i)\\ C(i)\mathscr{X}_{12}(i) & \gamma_m\end{bmatrix}\right\}=$$

$$=-\sum_{i\in\Theta}\left\{\begin{bmatrix}\mathcal{E}_{12}^\intercal(i)A(i)\mathscr{X}_{12}(i) & \mathcal{E}_{12}^\intercal(i)A(i)\mathscr{X}_1(i)C^\intercal(i)\\ B^\intercal(i)\mathcal{E}_1^\intercal(i)A(i)\mathscr{X}_{12}(i) & B^\intercal(i)\mathcal{E}_1(i)A(i)\mathscr{X}_1(i)C^\intercal(i)\end{bmatrix}\right\}$$

$$-\sum_{i\in\Theta}\left\{\begin{bmatrix}\mathcal{E}_{12}^\intercal(i)A(i)\mathscr{X}_{12}(i) & \mathcal{E}_{12}^\intercal(i)A(i)\mathscr{X}_1(i)C^\intercal(i)\\ B^\intercal(i)\mathcal{E}_1^\intercal(i)A(i)\mathscr{X}_{12}(i) & B^\intercal(i)\mathcal{E}_1(i)A(i)\mathscr{X}_1(i)C^\intercal(i)\end{bmatrix}\right\}=\sum_{i\in\Theta}\left\{\begin{bmatrix}0_n & 0_{n\times m}\\ 0_{p\times n} & 2\,\rho(i)R(i)MJ(i)VJ^\intercal(i)\end{bmatrix}+\right.$$

$$+\left[\begin{matrix}\mathcal{E}_2(i)F\mathscr{X}_2(i)+\mathcal{E}_{12}^\intercal(i)B(i)L\mathscr{X}_2(i)+\mathcal{E}_2(i)KC(i)\mathscr{X}_{12}(i)+\mathcal{E}_{12}^\intercal(i)B(i)MC(i)\mathscr{X}_{12}(i)\\ \\ B^\intercal(i)\mathcal{E}_{12}(i)F\mathscr{X}_2(i)+\gamma_l(i)L\mathscr{X}_2(i)+B^\intercal(i)\mathcal{E}_{12}(i)KC(i)\mathscr{X}_{12}(i)+\gamma_l(i)MC(i)\mathscr{X}_{12}(i)\end{matrix}\right.$$

$$\left.\left.\begin{matrix}\mathcal{E}_2(i)F\mathscr{X}_{12}^\intercal(i)C^\intercal(i)+\mathcal{E}_{12}^\intercal(i)B(i)L\mathscr{X}_{12}^\intercal(i)C^\intercal(i)+\mathcal{E}_2(i)K\gamma_m(i)+\mathcal{E}_{12}^\intercal(i)B(i)M\gamma_m(i)\\ \\ B^\intercal(i)\mathcal{E}_{12}(i)F\mathscr{X}_{12}^\intercal(i)C^\intercal(i)+\gamma_l(i)L\mathscr{X}_{12}^\intercal(i)C^\intercal(i)+B^\intercal(i)\mathcal{E}_{12}(i)K\gamma_m(i)+\gamma_l(i)M\gamma_m(i)\end{matrix}\right]\right\}$$

Leading to the following coupled set of matrix equations,

$$\begin{aligned}
&\sum_{i\in\Theta}\Big\{\mathcal{E}_2(i)\underline{F}\mathscr{X}_2(i)+\mathcal{E}_2(i)\underline{K}C(i)\mathscr{X}_{12}(i)+\mathcal{E}_{12}^\intercal(i)B(i)\underline{L}\mathscr{X}_2(i)+\\
&\qquad+\mathcal{E}_{12}^\intercal(i)B(i)\underline{M}C(i)\mathscr{X}_{12}(i)+\mathcal{E}_{12}^\intercal(i)A(i)\mathscr{X}_{12}(i)\Big\}=0_n\\
&\sum_{i\in\Theta}\Big\{\mathcal{E}_2(i)\underline{F}\mathscr{X}_{12}^\intercal(i)C^\intercal(i)+\mathcal{E}_2(i)\underline{K}\gamma_m(i)+\mathcal{E}_{12}^\intercal(i)B(i)\underline{L}\mathscr{X}_{12}^\intercal(i)C^\intercal(i)+\\
&\qquad+\mathcal{E}_{12}^\intercal(i)B(i)\underline{M}\gamma_m(i)+\mathcal{E}_{12}^\intercal(i)A(i)\mathscr{X}_1(i)C^\intercal(i)\Big\}=0_{n\times m}\\
&\sum_{i\in\Theta}\Big\{B^\intercal(i)\mathcal{E}_{12}(i)\underline{F}\mathscr{X}_2(i)+B^\intercal(i)\mathcal{E}_{12}(i)\underline{K}C(i)\mathscr{X}_{12}(i)+\gamma_l(i)\underline{L}\mathscr{X}_2(i)+\\
&\qquad+\gamma_l(i)\underline{M}C(i)\mathscr{X}_{12}(i)+B^\intercal(i)\mathcal{E}_1(i)A(i)\mathscr{X}_{12}(i)\Big\}=0_{p\times n}\\
&\sum_{i\in\Theta}\Big\{B^\intercal(i)\mathcal{E}_{12}(i)\underline{F}\mathscr{X}_{12}^\intercal(i)C^\intercal(i)+B^\intercal(i)\mathcal{E}_{12}(i)\underline{K}\gamma_m(i)+\gamma_l(i)\underline{L}\mathscr{X}_{12}^\intercal(i)C^\intercal(i)+\\
&\qquad+\gamma_l(i)\underline{M}\gamma_m(i)+2\,\rho(i)R(i)\underline{M}J(i)VJ^\intercal(i)+B^\intercal(i)\mathcal{E}_1(i)A(i)\mathscr{X}_1(i)C^\intercal(i)\Big\}=0_{p\times m}
\end{aligned}\tag{4.55}$$

Using vec operator and kronecker product properties we can factorize F,K,L,M,

$$\begin{bmatrix} \Gamma_{FF} & \Gamma_{FK} & \Gamma_{FL} & \Gamma_{FM} \\ \Gamma_{FK}^{\mathsf{T}} & \Gamma_{KK} & \Gamma_{KL} & \Gamma_{KM} \\ \Gamma_{FL}^{\mathsf{T}} & \Gamma_{KL}^{\mathsf{T}} & \Gamma_{LL} & \Gamma_{LM} \\ \Gamma_{FM}^{\mathsf{T}} & \Gamma_{KM}^{\mathsf{T}} & \Gamma_{LM}^{\mathsf{T}} & \Gamma_{MM} \end{bmatrix} \begin{bmatrix} vec(F) \\ vec(K) \\ vec(L) \\ vec(M) \end{bmatrix} + \begin{bmatrix} \Gamma_{FC} \\ \Gamma_{KC} \\ \Gamma_{LC} \\ \Gamma_{MC} \end{bmatrix} = 0_{4n} \tag{4.56}$$

where,

$$\Gamma_{FF} = \sum_{i\in\Theta} \left(\mathscr{X}_2(i) \otimes \mathcal{E}_2(i)\right) \qquad \Gamma_{FK} = \sum_{i\in\Theta} \left(\mathscr{X}_{12}^{\mathsf{T}}(i) C^{\mathsf{T}}(i) \otimes \mathcal{E}_2(i)\right)$$

$$\Gamma_{FL} = \sum_{i\in\Theta} \left(\mathscr{X}_2(i) \otimes \mathcal{E}_{12}^{\mathsf{T}}(i) B(i)\right) \qquad \Gamma_{FM} = \sum_{i\in\Theta} \left(\mathscr{X}_{12}^{\mathsf{T}}(i) C^{\mathsf{T}}(i) \otimes \mathcal{E}_{12}^{\mathsf{T}}(i) B(i)\right)$$

$$\Gamma_{FC} = vec\Big(\sum_{i\in\Theta} \left(\mathcal{E}_{12}^{\mathsf{T}}(i) A(i) \mathscr{X}_{12}(i)\right) \Big)$$

$$\Gamma_{KK} = \sum_{i\in\Theta} \left(\gamma_m(i) \otimes \mathcal{E}_2(i)\right) \qquad \Gamma_{KL} = \sum_{i\in\Theta} \left(C(i) \mathscr{X}_{12}(i) \otimes \mathcal{E}_{12}^{\mathsf{T}}(i) B(i)\right)$$

$$\Gamma_{KM} = \sum_{i\in\Theta} \left(\gamma_m(i) \otimes \mathcal{E}_{12}^{\mathsf{T}}(i) B(i)\right)$$

$$\Gamma_{KC} = vec\Big(\sum_{i\in\Theta} \left(\mathcal{E}_{12}^{\mathsf{T}}(i) A(i) \mathscr{X}_1(i) C^{\mathsf{T}}(i)\right) \Big)$$

$$\Gamma_{LL} = \sum_{i\in\Theta} \left(\mathscr{X}_2(i) \otimes \gamma_l(i)\right) \qquad \Gamma_{LM} = \sum_{i\in\Theta} \left(\mathscr{X}_{12}^{\mathsf{T}}(i) C^{\mathsf{T}}(i) \otimes \gamma_l(i)\right)$$

$$\Gamma_{LC} = vec\Big(\sum_{i\in\Theta} \left(B^{\mathsf{T}}(i) \mathcal{E}_1(i) A(i) \mathscr{X}_{12}(i)\right) \Big)$$

$$\Gamma_{MM} = \sum_{i\in\Theta} \left(\gamma_m(i) \otimes \gamma_l(i) + 2\,\rho(i) J(i) V J^{\mathsf{T}}(i) \otimes R(i)\right)$$

$$\Gamma_{MC} = vec\Big(\sum_{i\in\Theta} \left(B^{\mathsf{T}}(i) \mathcal{E}_1(i) A(i) \mathscr{X}_1(i) C^{\mathsf{T}}(i)\right) \Big)$$

And the solution can be obtained solving the linear system (4.56),

$$X_{sys} = (\Gamma)^{+}(-C) \tag{4.57}$$

4.3 Feasibility operator and the basics of the algorithm

Aiming at a numerically inexpensive feasibility operator, and taking into account that the fixed point equations (3.27), (4.1) may be written as conventional linear systems in the form

$\mathscr{A}_{\mathcal{F}} \cdot x = b$ by using stack-column operators and the Kronecker product properties, we start defining $\mathcal{F} : \{(x,p) \in \mathcal{S}_{r\mathcal{N}}^{\mathcal{N}(+)} \times \mathcal{S}_{r\mathcal{N}}^{\mathcal{N}(+)}\}$

$$\mathcal{F}(\mathrm{X}_\infty, \mathrm{P}_\infty) = \underset{\mathrm{X}_\infty,\mathrm{P}_\infty}{\mathrm{feasb}}\{(\mathrm{X}_\infty, \mathrm{P}_\infty) : \text{ (3.27) and (4.1) hold true}\}. \tag{4.58}$$

An inconvenience is that we can not show existence of solution to the above feasibility problem. This motivated us to "relax" the use of $\mathcal{F}$, as described next. We need some additional notation: we write the OF in (4.35) as $\mathcal{J}(\mathrm{U}, \mathrm{X}_\infty, \mathrm{P}_\infty)$ to emphasize the dependence on its arguments. Also, we denote $\mathrm{U}^\eta, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta$ the solution at the η-th iteration of the algorithm. Given $\mathrm{U}^\eta, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta$ one computes

$$\mathrm{U}^{\eta+1} = \mathcal{O}(\mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta),$$

providing $\mathrm{U}^{\eta+1}, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta$. If the following conditions hold:

$$\text{(i) } (\mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1}) = \mathcal{T}(\mathrm{U}^\eta) \text{ exists,}$$

and

$$\text{(ii) } \mathcal{J}(\mathrm{U}^{\eta+1}, \mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1}) \leq \mathcal{J}(\mathrm{U}^\eta, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta) \tag{4.59}$$

then set $\eta = \eta + 1$ and go back to the optimization step. If any of (i)/(ii) fails, or both fail, compute

$$\begin{aligned} \mathrm{X}_{\mathrm{aux}} &= \mathbf{F}(\mathrm{X}_\infty^\eta, \mathrm{U}^{\eta+1}) + \Sigma, \\ \mathrm{P}_{\mathrm{aux}} &= \mathbf{G}(\mathrm{P}_\infty^\eta, \mathrm{U}^{\eta+1}) + \mathrm{Q}. \end{aligned} \tag{4.60}$$

Now, if $\mathcal{J}(\mathrm{U}^{\eta+1}, \mathrm{X}_{\mathrm{aux}}, \mathrm{P}_{\mathrm{aux}}) \leq \mathcal{J}(\mathrm{U}^\eta, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta)$ then set $\eta = \eta + 1$ and go back to the optimization step.; otherwise, find the largest $\rho, 0 \leq \rho < 1$ such that

$$(\mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1}) := \rho\left(\mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta\right) + (1-\rho)\left(\mathrm{X}_{\mathrm{aux}}, \mathrm{P}_{\mathrm{aux}}\right)$$

satisfies (4.59). Such a ρ always exists because, in view of Lemma 4, $\mathcal{J}(\mathrm{U}^{\eta+1}, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta) \leq \mathcal{J}(\mathrm{U}^\eta, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta)$.

4.4 The algorithm

The algorithm explained in Section 4.3 is formalized in the sequel.

Algorithm 1 – Two Stage $\mathcal{O}$-$\mathcal{F}$

Result: U
Initialize: $\mathrm{U}^0, \mathrm{X}_\infty^0, \mathrm{P}_\infty^0, maxIter$
$\eta \leftarrow 0$
while $\eta < maxIter$ **do**
  $\mathrm{U}^{\eta+1} \leftarrow \mathcal{O}(\mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta)$
  if $\mathcal{T}(\mathrm{U}^{\eta+1})$ **then**
    $(\mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1}) \leftarrow \mathcal{T}(\mathrm{U}^\eta)$
    $\mathcal{J}^\eta \leftarrow \mathcal{J}(\mathrm{U}^\eta, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta)$
    $\mathcal{J}^{\eta+1} \leftarrow \mathcal{J}(\mathrm{U}^{\eta+1}, \mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1})$
    if $\mathcal{J}^{\eta+1} > \mathcal{J}^\eta$ **then**
      $(\mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1}) \leftarrow \mathcal{R}(\mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta, \mathrm{U}^{\eta+1}, \mathrm{U}^\eta, \mathcal{J}^\eta)$
    end
  else
    $(\mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1}) \leftarrow \mathcal{R}(\mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta, \mathrm{U}^{\eta+1}, \mathrm{U}^\eta, \mathcal{J}^\eta)$
  end
  $\eta + 1 \leftarrow \eta$
end

`procedure` $\mathcal{R}(\mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta, \mathrm{U}^{\eta+1}, \mathrm{U}^\eta, \mathcal{J}^\eta)$
$(\mathrm{X}_{aux}, \mathrm{P}_{aux}) \leftarrow \mathcal{T}(\mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta, \mathrm{U}^{\eta+1})$
$\mathcal{J}^{\eta+1} \leftarrow \mathcal{J}(\mathrm{U}^{\eta+1}, \mathrm{X}_{\mathrm{aux}}, \mathrm{P}_{\mathrm{aux}})$
while $\mathcal{J}^{\eta+1} > \mathcal{J}^\eta$ **do**
  $(\mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1}) \leftarrow \rho\left(\mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta\right) + (1-\rho)\left(\mathrm{X}_{\mathrm{aux}}, \mathrm{P}_{\mathrm{aux}}\right)$
  $\rho \leftarrow \rho + 0.01$
end

Theorem 6. Algorithm 1 converges.

Proof. The proof is immediate from the facts that Lemma 4 ensures $\mathcal{J}(\mathrm{U}^{\eta+1}, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta) \leq \mathcal{J}(\mathrm{U}^\eta, \mathrm{X}_\infty^\eta, \mathrm{P}_\infty^\eta)$ and that $\mathrm{X}_\infty^{\eta+1}, \mathrm{P}_\infty^{\eta+1}$ satisfy the condition in (4.59).
For indirect variational methods initialization plays a huge role, in the sense that it can influence the quality of convergence and its domain, for our method (Algorithm 1) we initialize first, that is, we initialize with a starting feasibility set: $(\mathrm{X}_\infty^0, \mathrm{P}_\infty^0)$ that by definition is bounded from below (by its positive definiteness); with this set we encounter a minimizer U^0 for $\mathcal{O}(\mathrm{X}_\infty^0, \mathrm{P}_\infty^0)$. In our experience fast convergence is achieved for small random initialization values of $\mathrm{X}_\infty^0, \mathrm{P}_\infty^0$. Two main conclusions can be drawn for the analysis of the initialization of Algorithm 1, the first one is for the convergence of a static controller

Corollary 1. For $\mathcal{T}(\mathrm{X}_\infty^0, \mathrm{P}_\infty^0) := \{(\alpha I_{nN}, \beta I_{nN}) | \alpha, \beta \in \mathbb{R}\}$ Algorithm ?? converges and $\mathcal{O}(\mathrm{U}^\eta)$

is a static controller defined by the solution of the matrix equation:

$$\sum_{i\in\Theta}\left\{2\,\rho(i)R(i)MJ(i)VJ^{\mathsf{T}}(i)+\gamma_l(i)\;M\;\gamma_m(i)\right\}= \\ =-\sum_{i\in\Theta}\left\{B^{\mathsf{T}}(i)\mathcal{E}_1(i)A(i)\mathscr{X}_1(i)C^{\mathsf{T}}(i)\right\} \tag{4.61}$$

where,

$$\gamma_l(i)=3R(i)+B^{\mathsf{T}}(i)\mathcal{E}_1(i)B(i)$$
$$\gamma_m=C(i)\mathscr{X}_1(i)C^{\mathsf{T}}(i)$$

Proof. The proof is straightforward, we calculate with the feasibility set $\{\mathrm{X}^0_\infty,\mathrm{P}^0_\infty\}:=(\alpha\ \mathrm{diag}(I^N_n),\beta\ \mathrm{diag}(I^N_n)$ the optimization operator for the initial minimizer $\mathcal{O}(\mathrm{U}^0)$ whom is of the form $\mathrm{U}^1=\begin{bmatrix}0&0\\0&M_0\end{bmatrix}$ and leads to a new feasibility set that is of the form $\{\mathrm{X}^1_\infty,\mathrm{P}^1_\infty\}:=\left\{\mathrm{diag}\left(\begin{bmatrix}\mathscr{X}_1(1)&0\\0&0\end{bmatrix},\ldots,\begin{bmatrix}\mathscr{X}_1(N)&0\\0&0\end{bmatrix}\right),\right.$ $\left.\mathrm{diag}\left(\begin{bmatrix}P_1(1)&0\\0&0\end{bmatrix},\ldots,\begin{bmatrix}P_1(N)&0\\0&0\end{bmatrix}\right)\right\}$ deriving in to a new minimizer $\mathcal{O}(\mathrm{U}^1)$ for the matrix equation (4.61).

Remark 2. When the feasibility set, is set with $(\mathrm{X}^0_\infty,\mathrm{P}^0_\infty)$ with quadratic dimension $n+n_c$ with $0<n_c<n$, we will obtain a "reduced-order" controller in the sense that the $\hat{x}$ will have a reduction in dimension with respect to the system state x, this fact comes from (3.5) and (3.6). We can say that Algorithm **??**, can converge also to a detector based controller in the sense that a sequence of controllers can be obtained following a cluster of modes.

CHAPTER

5

NUMERICAL SIMULATION

We applied the proposed control strategy for 10^3 instances of the system (3.1) and for our cost function, which were randomly created. We obtained the parameters for the controller (3.2) using the proposed Algorithm $\mathcal{O}-\mathcal{F}$, and also using a genetic algorithm (GA). The implemented GA follows the classical concepts proposed by Holland (HOLLAND, 1975), having the steps of *fitness, selection, crossover, mutation* and *replacement*. We denote by $\mathcal{J}_{\mathcal{OF}}$ the OF obtained with Algorithm $\mathcal{O}-\mathcal{F}$, and $\mathcal{J}_{\text{AG}}$ the OF via the AG.

Table 1 shows the proportion of instances / methods yielding $\mathcal{J} < 10^{10}$, and Figure 4 gives a comparison of the costs using both algorithms, when the cost is smaller than $\mathcal{J} < 10^{10}$. As we can see, the proposed Algorithm $\mathcal{O}-\mathcal{F}$ clearly outperforms the GA: the percentage of instances with costs below 10^{10} is much higher ($55.47\% + 28.53\% = 84\%$ via $\mathcal{O}-\mathcal{F}$ versus $8.61\% + 28.53\% = 37.14\%$ via AG) and Figure 4 contains more points above and far from the line $\mathcal{J}_{\mathcal{OF}} = \mathcal{J}_{\text{AG}}$.

Regarding the instance generator, the dimension of the state variable, output, noise processes and Markov state have uniform distribution with $2 \leq n_x \leq 5$, $2 \leq n_v \leq 5$, $1 \leq n_h \leq 5$, $1 \leq n_j \leq 5$, $1 \leq n_c \leq 5$, and $2 \leq \mathcal{N} \leq 4$.

Some features of the modes are also random, e.g. (A_1, B_1) has 80% of chance of being controllable.

For each instance we run the Riccati-based algorithm in (DRAGAN; COSTA, 2016b) (that is, in the scenario of perfect observation of θ) and we discard the instance when the cost is higher than 10^{100}, for it is likely not to be stabilizable in the scenario we are dealing with. Around 50% of instances were discarded, indicating that the generator creates hard problems. For the remaining 10^3 instances, 80% are not stable - that is, when all matrices $B_i = 0$, there is no solution for (22) with $\Sigma = I$, again indicating that a significant share of the instances are hard to handle. This also means that 20% of the instances are easy (stable), which makes the 37.14% success rate of the AG less significant when compared with the 84% of our algorithm.

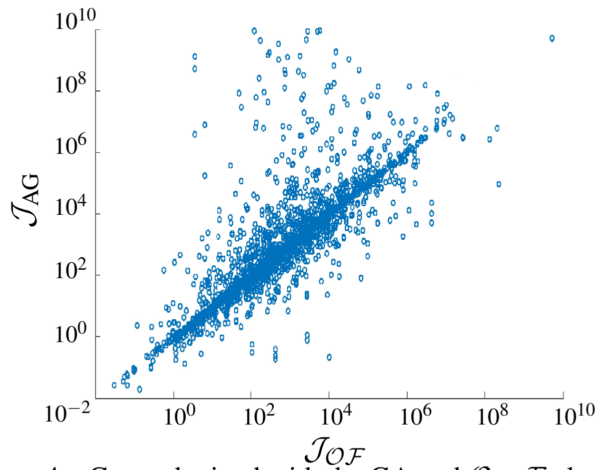


Figure 4 – Costs obtained with the GA and $\mathcal{O}-\mathcal{F}$ algorithm.

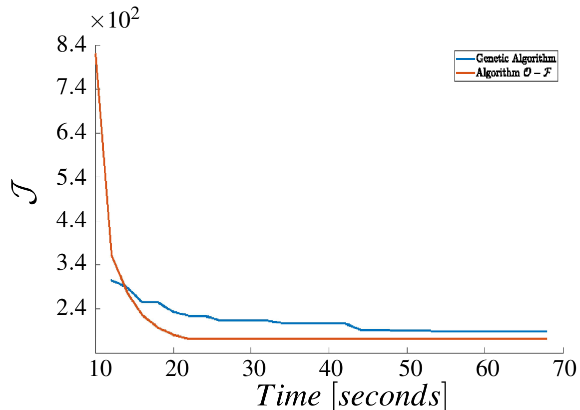


Figure 5 – Costs obtained with the GA and $\mathcal{O}-\mathcal{F}$ algorithm along execution time for instance number 1.

Table 1 – Percentages of instances / methods with costs smaller than 10^{10} ("success rate" of the methods).

	%
$\mathcal{J}_{\mathrm{AG}} < 10^{10}$, $\mathcal{J}_{\mathcal{OF}} \geq 10^{10}$	8.61%
$\mathcal{J}_{\mathrm{AG}} \geq 10^{10}$, $\mathcal{J}_{\mathcal{OF}} < 10^{10}$	55.47%
$\mathcal{J}_{\mathrm{AG}} < 10^{10}$, $\mathcal{J}_{\mathcal{OF}} < 10^{10}$	28.53%
$\mathcal{J}_{\mathrm{AG}} \geq 10^{10}$, $\mathcal{J}_{\mathcal{OF}} \geq 10^{10}$	7.39%

CHAPTER

6

CONCLUSIONS AND DISCUSSION

This thesis has studied discrete-time linear systems and stationary dynamic output-feedback controllers, featuring selectable dimension n_x, additive noise, jump parameters governed by a Markov chain that is not necessarily ergodic (periodic and transient states allowed), and detector-based observation of θ". A compact formulation for the optimization problem was obtained, relying on an operator-based approach, which is the starting point for us to develop the two stage $\mathcal{O}-\mathcal{F}$ (optimality - feasibility) algorithm. Numerical experiments with randomly created plants and weighting matrices indicate that our method finds solutions with "feasible cost" ($\mathcal{J} < 10^{10}$) in 84% of the 10^3 random instances, versus only 37.14% by the genetic algorithm. Future research may further explore the complexity of the OF, similarly to what was done in Section IV, and include weights in the OF to modulate the relative importance of the terms in the partial derivatives of the optimization step.

BIBLIOGRAPHY

ÅSTRÖM, K. J. **Introduction to Stochastic Control Theory**. [S.l.]: Academic Press, 1970. Citation on page 18.

BERTSEKAS, D. **Dynamic Programming and Stochastic Control**. [S.l.]: Athena Scientific, 1987. Citation on page 17.

BORTOLIN, D. **Métodos numéricos para o controle linear quadrático com saltos e observação parcial de estado**. Master's Thesis (Master's Thesis) — Universidade de São Paulo, 2012. Citation on page 20.

CHIZECK, H. J.; WILLSKY, A. S. Discrete-time markovian jump linear quadratic optimal control. **International Journal of Control**, v. 43, n. 1, p. 213–231, 1986. Citation on page 19.

COSTA, E.; VARGAS, J. d. V. A. Quadratic costs and second moments of jump linear systems with general markov chain. **Math. Control Signals Syst**, v. 23, n. 1, p. 141–157, 2011. Citation on page 19.

COSTA, O.; ARAUJO, M. A generalized multi-period mean-variance portfolio optimiza- tion with markov switching parameters. **Automatica**, v. 44, n. 10, p. 2487–2497, 2008. Citation on page 18.

COSTA, O. L. V.; FRAGOSO, M. D.; MARQUES, R. P. **Discrete-Time Markov Jump Linear Systems**. 1st. ed. [S.l.]: Springer-Verlag London, 2005. (Probability and Its Applications). Citations on pages 18, 19, and 25.

COSTA, O. L. V.; FRAGOSO, M. D.; TODOROV, M. **Continuous-Time Markov Jump Linear Systems**. 1. ed. [S.l.]: Springer-Verlag Berlin Heidelberg, 2013. (Probability and Its Applications). Citation on page 19.

COSTA, O. L. V.; FRAGOSO, M. D.; TODOROV, M. G. A detector-based approach for the H_2 control of markov jump linear systems with partial information. **IEEE Transactions on Automatic Control**, v. 60, n. 5, p. 1219 – 1233, 2015. Citation on page 20.

de Oliveira, A. M.; Costa, O. L. V.; Daafouz, J. Design of stabilizing dynamic output feedback controllers for hidden markov jump linear systems. **IEEE Control Systems Letters**, v. 2, n. 2, p. 278–283, April 2018. ISSN 2475-1456. Citations on pages 20 and 24.

DING, F.; CHEN, T. Gradient based iterative algorithms for solving a class of matrix equations. **IEEE Transactions on Automatic Control**, v. 50, n. 8, p. 1216 – 1221, 2005. Citation on page 38.

____. Iterative least-squares solutions of coupled sylvester matrix equations. **Systems & Control Letters**, n. 54, p. 95–107, 2005. Citation on page 38.

DOLGOV, M. **Approximate Stochastic Optimal Control of Smooth Nonlinear Systems and Piecewise Linear Systems**. Phd Thesis (PhD Thesis) — Fakultät für Informatik des Karlsruher Instituts für Technologie (KIT), May 2017. Citation on page 33.

DOLGOV, M.; CHLEBEK, C.; HANEBECK, U. D. Dynamic compensation of markov jump linear systems without mode observation. In: **Proceedings of the 2016 European Control Conference (ECC)**. [S.l.: s.n.], 2016. p. 2060–2065. Citations on pages 19, 20, and 24.

DOLGOV, M.; HANEBECK, U. D. Infinite - horizon linear optimal control of markov jump systems without mode observation via state feedback. **arXiv preprint: Systems and Control (cs.SY)**, July 2015. Citation on page 20.

____. Static output- feedback control of markov jump linear systems without mode observation. **IEEE Transactions in Automatic Control**, v. 62, n. 10, p. 5401 – 5406, October 2017. Citation on page 20.

DOOB, J. **Stochastic processes**. [S.l.]: Wiley, 1953. (Wiley publications in statistics). Citation on page 22.

DRAGAN, V.; COSTA, E. F. Optimal stationary dynamic output-feedback controllers for discrete-time linear systems with markovian jumping parameters and additive white noise perturbations. **IEEE Transactions on Automatic Control**, v. 61, n. 12, p. 3912 – 3924, December 2016. Citations on pages 18, 19, 20, 24, and 37.

____. Optimal stationary dynamic output-feedback controllers for discrete-time linear systems with markovian jumping parameters and additive white noise perturbations. **IEEE Transactions on Automatic Control**, v. 61, n. 12, p. 3912–3924, Dec 2016. ISSN 0018-9286. Citation on page 45.

ELLIOTT T. K. SIU, L. C. R. J.; LAU, J. W. Pricing options under a generalized markov-modulated jump-diffusion model. **Stochastic Analysis and Applications,**, v. 25, n. 4, p. 821–843, 2007. Citation on page 18.

HESPANHA, J. P.; NAGHSHTABRIZI, P.; XU, Y. A survey of recent results in networked control systems. **Proceedings of the IEEE**, v. 95, n. 1, p. 138–162, January 2007. Citation on page 18.

HOLLAND, J. H. **Adaptation in natural and artificial systems**. [S.l.]: University of Michigan Press, 1975. Citation on page 45.

KRASOVSKII, N. N.; LIDSKII, E. A. Analytic design of controller in systems with random attributes - part i. **Automation and Remote Control**, v. 22, p. 1021–1025, 1961. Citation on page 18.

MESKIN, N.; KHORASANI, K. Fault detection and isolation of discrete - time markovian jump linear systems with application to a network of multi-agent systems having imperfect communication channels,. **Automatica**, v. 45, n. 9, p. 2032–2040, 2009. Citation on page 18.

OLIVEIRA, L. T. **Método variacional com atualização múltipla de ganhos para controle de sistemas lineares com parâmetros sujeitos a saltos Markovianos não observados.** Master's Thesis (Master's Thesis) — Universidade de São Paulo, 2014. Citation on page 20.

POLDERMAN, J. W.; WILLEMS, J. C. **Introduction to Mathematical Systems Theory**. [S.l.]: Springer International, 2010. Citation on page 17.

Shaikh, M. S.; Caines, P. E. On the hybrid optimal control problem: Theory and algorithms. **IEEE Transactions on Automatic Control**, v. 52, n. 9, p. 1587–1603, Sep. 2007. ISSN 2334-3303. Citation on page 33.

SILVA, C. A. **Algoritmos para o custo médio a longo prazo de sistemas com saltos markovianos parcialmente observados**. Phd Thesis (PhD Thesis) — Universidade de São Paulo, 2012. Citation on page 20.

SIQUEIRA, A.; TERRA, M. A fault tolerant manipulator robot based on h2, h-infinity, and mixed h2/h-infinity markovian controls. **IEEE/ASME Trans. Mechatron.**, v. 14, n. 2, p. 257–263, 2009. Citation on page 19.

SONTAG, E. D. **Mathematical Control Theory: Deterministic Finite Dimensional Systems**. [S.l.]: Springer International, 1998. Citation on page 17.

TODOROV, M. G.; FRAGOSO, M. D.; COSTA, O. L. V. Detector-based H_∞ results for discrete-time markov jump linear systems with partial observations. **Automatica**, n. 91, p. 159–172, 2018. Citation on page 20.

VAL, J. B. R. do; BASAR, T. Receding horizon control of jump linear systems and a macroeconomic policy problem,. **Journal of Economic Dynamics and Control**, v. 23, n. 8, 1999. Citation on page 18.

VARGAS, A. N.; ACHO, L.; PUJOL, G.; COSTA, E. F.; ISHIHARA, J. a. Y.; VAL, J. a. B. R. do. Output feedback of Markov jump linear systems with no mode observation: An automotive throttle application. **International Journal of Robust and Nonlinear Control**, v. 26, n. 9, p. 1980–1993, 2016. ISSN 1099-1239. Rnc.3393. Available: <http://dx.doi.org/10.1002/rnc.3393>. Citation on page 20.

VARGAS, A. N.; COSTA, E. F.; VAL, J. B. R. do. On the control of markov jump linear systems with no mode observation: Application to a dc motor device. **International Journal of Robust and Nonlinear Control**, v. 23, n. 10, p. 1136–1150, 2013. Citations on pages 18, 19, and 20.

____. **Advances in the Control of Markov Jump Linear Systems with No Mode Observation**. 1st. ed. [S.l.]: Springer International Publishing, 2016. (SpringerBriefs in Electrical and Computer Engineering - Control, Automation and Robotics). Citation on page 20.

APPENDIX

A

APPENDIX

A.1 Matrix Derivatives

If $f : \mathbb{M}^{r\times d} \to \mathbb{R}$ is a differentiable function on the domain $\mathbb{M}^{r\times d}$, we denote the partial derivative $\partial f(X)/\partial X$ as $\partial_X f(X)$, whenever $X, A, B, C \in \mathbb{M}^{r\times d}$.

$$\begin{gathered} \partial_X\, tr(AXB) = A^\intercal B^\intercal \qquad \partial_X\, tr(AX^\intercal B) = BA \\ \partial_X\, tr(AXCX^\intercal B) = A^\intercal C^\intercal X B^\intercal + CAXB \end{gathered} \tag{A.1}$$

A.1.1 Derivatives

Given the cost functional

$$\begin{aligned}
\mathcal{J} = \sum_{i\in\Theta} tr\Bigg(\rho(i) &\begin{bmatrix} 0_{p\times n_c} & I_p \\ 0_{p\times n_c} & 0_p \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^\intercal(i) \end{bmatrix} \mathcal{U}^\intercal \begin{bmatrix} 0_{n_c\times p} & 0_{n_c\times p} \\ R(i) & 0_p \end{bmatrix}\Bigg) + \\
+(3/2)\sum_{i\in\Theta} tr\Bigg(&\begin{bmatrix} Q(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} \mathcal{X}_\infty(i) + \begin{bmatrix} 0_{n\times n_c} & C^\intercal(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \mathcal{U}^\intercal \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix} \mathcal{U} \cdot \\
&\cdot \begin{bmatrix} 0_{n_c\times n} & I_{n_c\times n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathcal{X}_\infty(i)\Bigg) + \\
+(1/2)\sum_{i\in\Theta} tr\Bigg[\mathcal{E}_i(P_\infty)&\Bigg(\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\Bigg) \mathcal{X}_\infty(i) \cdot \\
&\cdot \Bigg(\begin{bmatrix} A^\intercal(i) & 0_{n_c\times n} \\ 0_{n\times n_c} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & C^\intercal(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \mathcal{U}^\intercal \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\intercal(i) & 0_{p\times n_c} \end{bmatrix}\Bigg)\Bigg] + \\
-(1/2)\sum_{i\in\Theta} tr\Big(&P_\infty(i)\mathcal{X}_\infty(i)\Big)
\end{aligned} \tag{A.2}$$

Derivation with respect to $\mathcal{U}$ matrix, to ease the derivation we will denote every term as follows,

$$\mathcal{J}(\mathcal{U}) = \mathcal{J}_1(\mathcal{U}) + \mathcal{J}_2(\mathcal{U}) + \mathcal{J}_3(\mathcal{U}), \tag{A.3}$$

with,

$$\mathcal{J}_1(\mathcal{U}) = \sum_{i\in\Theta} tr\left(\rho(i)\begin{bmatrix} 0_{p\times n_c} & I_p \\ 0_{p\times n_c} & 0_p \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix}\mathcal{U}^{\mathsf{T}}\begin{bmatrix} 0_{n_c\times p} & 0_{n_c\times p} \\ R(i) & 0_p \end{bmatrix}\right)$$

$$\mathcal{J}_2(\mathcal{U}) = (3/2)\sum_{i\in\Theta} tr\left(\begin{bmatrix} Q(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix}\mathscr{X}_\infty(i) + \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}\mathcal{U}^{\mathsf{T}}\begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix}\mathcal{U}\cdot\right.$$
$$\left.\cdot\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\mathscr{X}_\infty(i)\right)$$

$$\mathcal{J}_3(\mathcal{U}) = (1/2)\sum_{i\in\Theta} tr\left[\mathcal{E}_i(P_\infty)\left(\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\right)\mathscr{X}_\infty(i)\cdot\right.$$
$$\left.\cdot\left(\begin{bmatrix} A^{\mathsf{T}}(i) & 0_{n_c\times n} \\ 0_{n\times n_c} & 0_{n_c} \end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}\mathcal{U}^{\mathsf{T}}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix}\right)\right]$$

For $\partial_{\mathcal{U}}(J(\mathcal{U}))$ we obtain the derivative for each term:

$$\partial_{\mathcal{U}}(J_1(\mathcal{U})) = \sum_{i\in\Theta}\left(\rho(i)\begin{bmatrix} 0_{n_c\times p} & 0_{n_c\times p} \\ I_p & 0_p \end{bmatrix}\begin{bmatrix} 0_{p\times n_c} & R^{\mathsf{T}}(i) \\ 0_{p\times n_c} & 0_p \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} + \right.$$
$$\left. + \begin{bmatrix} 0_{n_c\times p} & 0_{n_c\times p} \\ R(i) & 0_p \end{bmatrix}\begin{bmatrix} 0_{p\times n_c} & I_p \\ 0_{p\times n_c} & 0_p \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix}\right)$$

$$\partial_{\mathcal{U}}(J_1(\mathcal{U})) = \sum_{i\in\Theta}\left(\rho(i)\begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix}\right)$$

$$\partial_{\mathcal{U}}(J_2(\mathcal{U})) = (3/2)\sum_{i\in\Theta}\partial_{\mathcal{U}}\left\{tr\left(\begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}\mathcal{U}^{\mathsf{T}}\begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\mathscr{X}_\infty(i)\right)\right\}$$
$$= (3/2)\sum_{i\in\Theta}\partial_{\mathcal{U}}\left\{tr\left(\begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}\mathcal{U}^{\mathsf{T}}\right)\right\}$$

$$= (3/2)\sum_{i\in\Theta}\left(\begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i) \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \right.$$
$$\left. + \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R^{\mathsf{T}}(i) \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}\right)$$

$$\partial_{\mathcal{U}}(J_2(\mathcal{U})) = (3/2)\sum_{i\in\Theta}\left(\begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix}\right)$$

$$\partial_{\mathcal{U}}(J_3(\mathcal{U})) = (1/2)\sum_{i\in\Theta} \partial_{\mathcal{U}}\Bigg\{ tr\Bigg[\mathcal{E}_i(P_\infty)\Bigg(\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c}\end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p}\end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c}\end{bmatrix}\Bigg)\cdot$$
$$\cdot\,\mathscr{X}_\infty(i)\Bigg(\begin{bmatrix} A^\mathsf{T}(i) & 0_{n_c\times n} \\ 0_{n\times n_c} & 0_{n_c}\end{bmatrix} + \begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}\mathcal{U}^\mathsf{T}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\Bigg)\Bigg]\Bigg\}$$

$$= (1/2)\sum_{i\in\Theta} \partial_{\mathcal{U}}\Bigg\{ tr\Bigg(\mathcal{E}_i(P_\infty)\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} A^\mathsf{T}(i) & 0_{n_c\times n} \\ 0_{n\times n_c} & 0_{n_c}\end{bmatrix}\Bigg)+$$
$$+tr\Bigg(\mathcal{E}_i(P_\infty)\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}\mathcal{U}^\mathsf{T}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\Bigg)+$$
$$+tr\Bigg(\mathcal{E}_i(P_\infty)\begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p}\end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} A^\mathsf{T}(i) & 0_{n_c\times n} \\ 0_{n\times n_c} & 0_{n_c}\end{bmatrix}\Bigg)+$$
$$+tr\Bigg(\mathcal{E}_i(P_\infty)\begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p}\end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}\mathcal{U}^\mathsf{T}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\Bigg)\Bigg\}$$

$$= (1/2)\sum_{i\in\Theta}\Bigg(\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\mathcal{E}_i(P_\infty)\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}+$$
$$+\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\mathcal{E}_i(P_\infty)\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}+$$
$$+\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\mathcal{E}_i(P_\infty)\begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p}\end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}+$$
$$+\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\mathcal{E}_i(P_\infty)\begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p}\end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}\Bigg)$$

$$\partial_{\mathcal{U}}(J_3(\mathcal{U})) = (1/2)\sum_{i\in\Theta}\Bigg(2\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\mathcal{E}_i(P_\infty)\begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}+$$
$$+2\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^\mathsf{T}(i) & 0_{p\times n_c}\end{bmatrix}\mathcal{E}_i(P_\infty)\begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p}\end{bmatrix}\mathcal{U}\begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c}\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix} 0_{n\times n_c} & C^\mathsf{T}(i) \\ I_{n_c} & 0_{n_c\times m}\end{bmatrix}\Bigg)$$

Therefore,

$$
\begin{aligned}
\partial_{\mathcal{U}}(J(\mathcal{U})) = \sum_{i\in\Theta} \Bigg\{ & \rho(i) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} + \\
& + (3/2) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \\
& + (1/2) \Bigg(2 \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \\
& + 2 \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \Bigg) \Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
\partial_{\mathcal{U}}(J(\mathcal{U})) = \sum_{i\in\Theta} \Bigg\{ & \rho(i) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_{n_c} & 0_{n_c\times m} \\ 0_{m\times n_c} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} + \\
& + (3/2) \begin{bmatrix} 0_{n_c} & 0_{n_c\times p} \\ 0_{p\times n_c} & R(i)+R^{\mathsf{T}}(i) \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \\
& + \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} A(i) & 0_{n\times n_c} \\ 0_{n_c\times n} & 0_{n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} + \\
& + \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ B^{\mathsf{T}}(i) & 0_{p\times n_c} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} 0_{n\times n_c} & B(i) \\ I_{n_c} & 0_{n_c\times p} \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_{n_c\times n} & I_{n_c} \\ C(i) & 0_{m\times n_c} \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_{n\times n_c} & C^{\mathsf{T}}(i) \\ I_{n_c} & 0_{n_c\times m} \end{bmatrix} \Bigg\}
\end{aligned} \tag{A.4}
$$

A.1.2 Observations on FKLM and U

$$
\begin{aligned}
& \sum_{i\in\Theta} \Bigg\{ 2\,\rho(i) \begin{bmatrix} 0_n & 0_{n\times p} \\ 0_{p\times n} & R(i) \end{bmatrix} \mathcal{U} \begin{bmatrix} 0_n & 0_{n\times m} \\ 0_{m\times n} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} + \\
& + \begin{bmatrix} \mathcal{E}_2(i) & \mathcal{E}_{12}^{\mathsf{T}}(i)B(i) \\ B^{\mathsf{T}}(i)\mathcal{E}_{12}(i) & \gamma_l(i) \end{bmatrix} \mathcal{U} \begin{bmatrix} \mathscr{X}_2(i) & \mathscr{X}_{12}^{\mathsf{T}}(i)C^{\mathsf{T}}(i) \\ C(i)\mathscr{X}_{12}(i) & \gamma_m \end{bmatrix} \Bigg\} = \\
& = -\sum_{i\in\Theta} \Bigg\{ \begin{bmatrix} \mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_{12}(i) & \mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_1(i)C^{\mathsf{T}}(i) \\ B^{\mathsf{T}}(i)\mathcal{E}_1^{\mathsf{T}}(i)A(i)\mathscr{X}_{12}(i) & B^{\mathsf{T}}(i)\mathcal{E}_1(i)A(i)\mathscr{X}_1(i)C^{\mathsf{T}}(i) \end{bmatrix} \Bigg\}
\end{aligned}
$$

Applying vec operator and kronecker properties,

$$\sum_{i\in\Theta}\left\{2\,\rho(i)\begin{bmatrix}0_n & 0_{n\times m}\\ 0_{m\times n} & J(i)VJ^{\mathsf{T}}(i)\end{bmatrix}\otimes\begin{bmatrix}0_n & 0_{n\times p}\\ 0_{p\times n} & R(i)\end{bmatrix}+\right.$$
$$\left.+\begin{bmatrix}\mathscr{X}_2(i) & \mathscr{X}_{12}^{\mathsf{T}}(i)C^{\mathsf{T}}(i)\\ C(i)\mathscr{X}_{12}(i) & \gamma_m\end{bmatrix}\otimes\begin{bmatrix}\mathcal{E}_2(i) & \mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\\ B^{\mathsf{T}}(i)\mathcal{E}_{12}(i) & \gamma_l(i)\end{bmatrix}\right\}vec(\mathcal{U})=$$
$$=-vec\left(\sum_{i\in\Theta}\left\{\begin{bmatrix}\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_{12}(i) & \mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_1(i)C^{\mathsf{T}}(i)\\ B^{\mathsf{T}}(i)\mathcal{E}_1^{\mathsf{T}}(i)A(i)\mathscr{X}_{12}(i) & B^{\mathsf{T}}(i)\mathcal{E}_1(i)A(i)\mathscr{X}_1(i)C^{\mathsf{T}}(i)\end{bmatrix}\right\}\right)$$

A.1.3 Coupled matrix linear equations for one mode

$$\begin{aligned}
&\sum_{i\in\Theta}\Big\{\mathcal{E}_2(i)\underline{F}\mathscr{X}_2(i)+\mathcal{E}_2(i)\underline{K}C(i)\mathscr{X}_{12}(i)+\mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{L}\mathscr{X}_2(i)+\\
&\qquad+\mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{M}C(i)\mathscr{X}_{12}(i)+\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_{12}(i)\Big\}=0_n\\
&\sum_{i\in\Theta}\Big\{\mathcal{E}_2(i)\underline{F}\mathscr{X}_{12}^{\mathsf{T}}(i)C^{\mathsf{T}}(i)+\mathcal{E}_2(i)\underline{K}\gamma_m(i)+\mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{L}\mathscr{X}_{12}^{\mathsf{T}}(i)C^{\mathsf{T}}(i)+\\
&\qquad+\mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{M}\gamma_m(i)+\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_1(i)C^{\mathsf{T}}(i)\Big\}=0_{n\times m}\\
&\sum_{i\in\Theta}\Big\{B^{\mathsf{T}}(i)\mathcal{E}_{12}(i)\underline{F}\mathscr{X}_2(i)+B^{\mathsf{T}}(i)\mathcal{E}_{12}(i)\underline{K}C(i)\mathscr{X}_{12}(i)+\gamma_l(i)\underline{L}\mathscr{X}_2(i)+\\
&\qquad+\gamma_l(i)\underline{M}C(i)\mathscr{X}_{12}(i)+B^{\mathsf{T}}(i)\mathcal{E}_1(i)A(i)\mathscr{X}_{12}(i)\Big\}=0_{p\times n}\\
&\sum_{i\in\Theta}\Big\{B^{\mathsf{T}}(i)\mathcal{E}_{12}(i)\underline{F}\mathscr{X}_{12}^{\mathsf{T}}(i)C^{\mathsf{T}}(i)+B^{\mathsf{T}}(i)\mathcal{E}_{12}(i)\underline{K}\gamma_m(i)+\gamma_l(i)\underline{L}\mathscr{X}_{12}^{\mathsf{T}}(i)C^{\mathsf{T}}(i)+\\
&\qquad+\gamma_l(i)\underline{M}\gamma_m(i)+2\,\rho(i)R(i)\underline{M}J(i)VJ^{\mathsf{T}}(i)+B^{\mathsf{T}}(i)\mathcal{E}_1(i)A(i)\mathscr{X}_1(i)C^{\mathsf{T}}(i)\Big\}=0_{p\times m}
\end{aligned}\tag{A.5}$$

Solving the system for one mode, $i\in\Theta:\{1\}$, can be approached in $\mathcal{O}(n^3)$ using Gaussian-elimination analytically,

$$\begin{aligned}
\mathcal{E}_2(i)\underline{F}\mathscr{X}_2(i)+\mathcal{E}_2(i)\underline{K}C(i)\mathscr{X}_{12}(i)+\mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{L}\mathscr{X}_2(i)+\mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{M}C(i)\mathscr{X}_{12}(i)&=-\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_{12}(i)\\
\mathcal{E}_2(i)\underline{K}\gamma_{oM}+0+(\mathcal{E}_{12}^{\mathsf{T}}(i)B(i))\underline{M}\gamma_{oM}&=-\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\beta_{oM}\\
\gamma_{oL}\underline{L}\mathscr{X}_2(i)+\gamma_{oL}\underline{M}(C(i)\mathscr{X}_{12}(i))&=-\beta_{oL}A(i)\mathscr{X}_{12}(i)\\
\gamma_{oL}\underline{M}\gamma_{oM}+2\,\rho(i)R(i)\underline{M}J(i)VJ^{\mathsf{T}}(i)&=-\beta_{oL}A(i)\beta_{oM}
\end{aligned}\tag{A.6}$$

with,

$$\begin{aligned}
\gamma_{oM} &= \left(\gamma_m(i) - C(i)\mathscr{X}_{12}(i)\mathscr{X}_2^{-1}(i)\mathscr{X}_{12}^{\mathsf{T}}(i)C^{\mathsf{T}}(i)\right)\\
\gamma_{oL} &= \left(\gamma_l(i) - B^{\mathsf{T}}(i)\mathcal{E}_{12}(i)\mathcal{E}_2^{-1}(i)\mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\right)\\
\beta_{oM} &= \left(\mathscr{X}_1(i) - \mathscr{X}_{12}(i)\mathscr{X}_2^{-1}\mathscr{X}_{12}^{\mathsf{T}}(i)\right)C^{\mathsf{T}}(i)\\
\beta_{oL} &= B^{\mathsf{T}}(i)\left(\mathcal{E}_1(i) - B^{\mathsf{T}}(i)\mathcal{E}_{12}(i)\mathcal{E}_2^{-1}(i)\mathcal{E}_{12}^{\mathsf{T}}(i)\right)
\end{aligned}$$

Replacing (4.54) and (4.52),

$$\gamma_{oM} = C(i)\left(\mathscr{X}_1(i) - \mathscr{X}_{12}(i)\mathscr{X}_2^{-1}(i)\mathscr{X}_{12}^{\mathsf{T}}(i)\right)C^{\mathsf{T}}(i) \tag{A.7}$$

$$\gamma_{oL} = 3R(i) + B^{\mathsf{T}}(i)\left(\mathcal{E}_1 - \mathcal{E}_{12}(i)\mathcal{E}_2^{-1}(i)\mathcal{E}_{12}^{\mathsf{T}}(i)\right)B(i) \tag{A.8}$$

Therefore we have,

$$\begin{aligned}
\mathcal{E}_2(i)\underline{F}\mathscr{X}_2(i) + \mathcal{E}_2(i)\underline{K}C(i)\mathscr{X}_{12}(i) + \mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{L}\mathscr{X}_2(i) + \mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{M}C(i)\mathscr{X}_{12}(i) &= -\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_{12}(i)\\
\mathcal{E}_2(i)\underline{K}C(i) + (\mathcal{E}_{12}^{\mathsf{T}}(i)B(i))\underline{M}C(i) &= -\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\\
\left(3R(i) - B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}B(i)\right)\left(\underline{L}\mathscr{X}_2(i) + \underline{M}(C(i)\mathscr{X}_{12}(i))\right) &= -B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}A(i)\mathscr{X}_{12}(i)\\
\left(3R(i) - B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}B(i)\right)\underline{M}C(i) &= -B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}A(i)
\end{aligned} \tag{A.9}$$

with

$$\gamma_{o\mathcal{E}} = \left(\mathcal{E}_1(i) - \mathcal{E}_{12}(i)\mathcal{E}_2^{-1}(i)\mathcal{E}_2^{\mathsf{T}}(i)\right) \tag{A.10}$$

Operating,

$$\begin{aligned}
\mathcal{E}_2(i)\underline{F}\mathscr{X}_2(i) + \mathcal{E}_2(i)\underline{K}C(i)\mathscr{X}_{12}(i) + \mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{L}\mathscr{X}_2(i) + \mathcal{E}_{12}^{\mathsf{T}}(i)B(i)\underline{M}C(i)\mathscr{X}_{12}(i) &= -\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\mathscr{X}_{12}(i)\\
\mathcal{E}_2(i)\underline{K}C(i) + (\mathcal{E}_{12}^{\mathsf{T}}(i)B(i))\underline{M}C(i) &= -\mathcal{E}_{12}^{\mathsf{T}}(i)A(i)\\
\left(3R(i) - B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}B(i)\right)\underline{L}\mathscr{X}_2(i) - B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}A(i)\mathscr{X}_{12}(i) &= -B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}A(i)\mathscr{X}_{12}(i)\\
\left(3R(i) - B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}B(i)\right)\underline{M}C(i) &= -B^{\mathsf{T}}(i)\gamma_{o\mathcal{E}}A(i)
\end{aligned} \tag{A.11}$$

A.1.4 Iterative general Sylvester matrix equations

We present two ways of solving the matrix equation, both of them reducing the equation into a general sylvester equation form.

First we show the existence of the solution,

$$\left(\sum_{i\in\Theta}\left\{2\,\rho(i)\begin{bmatrix}0_n & 0_{n\times m}\\ 0_{m\times n} & J(i)VJ^{\mathsf{T}}(i)\end{bmatrix}\otimes\begin{bmatrix}0_n & 0_{n\times p}\\ 0_{p\times n} & R(i)\end{bmatrix} + \gamma_M(i)\,\otimes\,\gamma_L(i)\right\}\right) vec(\mathcal{U}) =$$

$$= -\,vec\left(\sum_{i\in\Theta}\left\{\begin{bmatrix}0_n & I_n\\ B^{\mathsf{T}}(i) & 0_{p\times n}\end{bmatrix}\mathcal{E}_i(P_\infty)\begin{bmatrix}A(i) & 0_n\\ 0_n & 0_n\end{bmatrix}\mathscr{X}_\infty(i)\begin{bmatrix}0_n & C^{\mathsf{T}}(i)\\ I_n & 0_{n\times m}\end{bmatrix}\right\}\right)$$

renaming,

$$\begin{aligned} D &= -\, vec\left(\sum_{i\in\Theta} \left\{ \begin{bmatrix} 0_n & I_n \\ B^{\mathsf{T}}(i) & 0_{p\times n} \end{bmatrix} \mathcal{E}_i(P_\infty) \begin{bmatrix} A(i) & 0_n \\ 0_n & 0_n \end{bmatrix} \mathscr{X}_\infty(i) \begin{bmatrix} 0_n & C^{\mathsf{T}}(i) \\ I_n & 0_{n\times m} \end{bmatrix} \right\} \right) \\ \Gamma &= \left\{ \sum_{i\in\Theta} \left(2\,\rho(i) \begin{bmatrix} 0_n & 0_{n\times m} \\ 0_{m\times n} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} \otimes \begin{bmatrix} 0_n & 0_{n\times p} \\ 0_{p\times n} & R(i) \end{bmatrix} + \gamma_M(i) \,\otimes\, \gamma_L(i) \right\} \end{aligned} \tag{A.12}$$

We have,

$$\Gamma\; vec(\mathcal{U}) = D \tag{A.13}$$

So if Γ is non-singular $vec(\mathcal{U})$ exists, and it is the solution of,

$$vec(\mathcal{U}) = \Gamma^{-1} D \tag{A.14}$$

As Γ is almost always near to singular so instead of using the inverse we use the generalized Moore-Penrose pseudoinverse,

$$vec(\mathcal{U}) = \Gamma^{+} D \tag{A.15}$$

We use a gradient iterative solution for based on **??**, in that sense we accommodate our matrix equation to the form of a general sylvester equation,

$$\sum_{i\in\Theta} \tilde{A}_j X \tilde{B}_j = \tilde{F} \tag{A.16}$$

$$\tilde{A}_1 X \tilde{B}_1 + \tilde{A}_2 X \tilde{B}_2 + \cdots + \tilde{A}_N X \tilde{B}_N = \tilde{F} \tag{A.17}$$

for $\tilde{A}, \tilde{B} \in \mathbb{M}^{r\times d}$ with j the index of Θ.
If we rename,

$$R_\Gamma(i) = 2\,\rho(i) \begin{bmatrix} 0_n & 0_{n\times p} \\ 0_{p\times n} & R(i) \end{bmatrix}; \quad J_\Gamma(i) = \begin{bmatrix} 0_n & 0_{n\times m} \\ 0_{m\times n} & J(i)VJ^{\mathsf{T}}(i) \end{bmatrix} \tag{A.18}$$

Then we have,

$$\begin{aligned} &\Big(\gamma_M(1)\,\mathcal{U}\;\gamma_L(1)\Big) + \Big(\gamma_M(2)\,\mathcal{U}\;\gamma_L(2)\Big) + \cdots + \Big(\gamma_M(N)\,\mathcal{U}\;\gamma_L(N)\Big) + \\ &+ \Big(R_\Gamma(1)\,\mathcal{U}\;J_\Gamma(1)\Big) + \Big(R_\Gamma(2)\,\mathcal{U}\;J_\Gamma(2)\Big) + \cdots + \Big(R_\Gamma(N)\,\mathcal{U}\;J_\Gamma(N)\Big) = D \end{aligned} \tag{A.19}$$

We see that γ_L and R_Γ, γ_M and J_Γ have the same dimensions respectively, considering that each matrix is indexed by $i \in \Theta$ whose cardinality is $\mathcal{N}$, we can group them in new indexing sets with cardinality of $\in\mathcal{N}$,

$$\tilde{A} = \left\{\gamma_L(1),\dots,\gamma_L(N),R_\Gamma(1),\dots,R_\Gamma(\mathcal{N})\right\} \tag{A.20}$$

$$\tilde{B} = \left\{\gamma_M(1),\dots,\gamma_M(N),J_\Gamma(1),\dots,J_\Gamma(\mathcal{N})\right\} \tag{A.21}$$

and considering a new index set, $\Upsilon := 1,\dots,2N$ We have,

$$\sum_{q\in\Upsilon} \tilde{A}(q)\,\mathcal{U}\,\tilde{B}(q) = D \tag{A.22}$$

And the gradient iterative solution is given by,

Algorithm 2 – GradIt for Grl Sylvester equation

Result: $\mathcal{U}$

Initialize $\mathcal{U}_0(i)$

$\mu = 1/[\sum_{i=1}^{2N}\|A\tilde{(i)}\|^2\|B\tilde{(i)}\|^2]$

while $k < maxIter$ **do**

$\mathcal{U}_k(i) = \mathcal{U}_{k-1}(i) + \mu\tilde{A}^\intercal(i)\left[D - \sum_{j=1}^{2N}\tilde{A}(j)\mathcal{U}_{k-1}\tilde{B}(j)\right]\tilde{B}^\intercal(i)$

$\mathcal{U} = [\mathcal{U}_k(1) + \mathcal{U}_k(2) + \cdots + \mathcal{U}_k(2N)]/2N$

if $|\mathcal{U}_k - \mathcal{U}_{k-1}| < minTol$ **then**

$k{+}{+}$

BREAK

end

$k{+}{+}$

end

ICMC USP
SÃO CARLOS