Approximate Stochastic Optimal Control of Smooth Nonlinear Systems and Piecewise Linear Systems

zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

der Fakultät für Informatik
des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Maxim Dolgov
aus St. Petersburg

Tag der mündlichen Prüfung:	05. Mai 2017
Erster Gutachter:	Prof. Dr.-Ing. Uwe D. Hanebeck
Zweiter Gutachter:	Prof. Dr. Daniel E. Quevedo

Acknowledgment

This thesis contains research carried out at the Intelligent Sensor–Actuator–Systems Laboratory (ISAS) at the Karlsruhe Institute of Technology under the advisory of Prof. Dr.-Ing. Uwe D. Hanebeck. I would like to thank Prof. Hanebeck for giving me the opportunity to work at his laboratory and for his support during that time. Also, I thank Prof. Dr. Daniel E. Quevedo for being my co-advisor. My work was supported by the German Research Foundation (DFG) as part of the Research Training Group 1194 (GRK 1194) on "Self-organizing Sensor-Actuator-Networks".

During my time at ISAS I had the opportunity to work in a great atmosphere created by my colleagues Marcus Baum, Christof Chlebek, Florian Faion, Jörg Fischer, who deserves a special thanks for being my mentor during my first year at ISAS, Igor Gilitschenski, Achim Hekler, Gerhard Kurz, Martin Pander, Florian Pfaff, Marc Reinhardt, Florian Rosenthal, Jannik Steinbring, and Antonio Zea. The morning coffee and the late afternoon breaks were filled with jokes and laughter, and sometimes even with scientific discussions.

Finally, I want to thank my wife Yasmin and my children Mila and Leo for their patience and understanding when I had to finish a paper or when my body was back from the lab but my mind was not. Many of the ideas that I had during my PhD time sparked in times when I disconnected from work with you.

Renningen, April 2018 Maxim Dolgov

Contents

Glossary and Notation

Glossary

BMI	Bilinear Matrix Inequality
CKF	Cubature Kalman Filter
DDP	Differential Dynamic Programming
DP	Dynamic Programming
EKF	Extended Kalman Filter
GP	Gaussian Process
KL	Kullback–Leibler divergence
LCD	Localized Cumulative Distribution
LQG	Linear Quadratic Gaussian
LMI	Linear Matrix Inequality
LTI	Linear Time-Invariant (system)
mCvMd	modified Cramér–von Mises distance
MJS	Markov Jump System
MJLS	Markov Jump Linear System
MPC	Model Predictive Control
OMAT	Optimal MAss Transfer
OSPA	Optimal SubPattern Assignment
POMDP	Partially-Observable Markov Decision Process
PWLC	PieceWise Linear Continuous
RUKF	Randomized Unscented Kalman Filter
SDP	Semi-Definite Program
S2KF	Smart Sampling Kalman Filter
UKF	Unscented Kalman Filter

General Notation

x	scalar
$\boldsymbol{x}$	random scalar
$\underline{x}$	vector
$\underline{\boldsymbol{x}}$	random vector
$\mathbf{X}$	matrix
$\mathbf{X}^\top$	transpose of $\mathbf{X}$
$\sqrt{\mathbf{X}}$	matrix square root of $\mathbf{X}$
$\mathbf{X}^\dagger$	pseudoinverse of $\mathbf{X}$
$\mathbf{I}_n$	$n \times n$ identity matrix
$\mathbf{0}_{m\times n}, \mathbf{0}_n$	$m \times n$ zero matrix, $n \times n$ zero matrix
$a_{1:n}$	abbreviation for a sequence $a_1, a_2, \ldots, a_n$
$\mathrm{tr}\,[\mathbf{X}]$	trace of $\mathbf{X}$
$\mathrm{diag}\,[\underline{x}]$	diagonal matrix with $\underline{x}$ on the diagonal
$\mathrm{diag}\,[\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n]$	block diagonal matrix of matrices $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$
$\mathrm{vec}\,[\mathbf{X}]$	vectorization of $\mathbf{X}$
$\mathrm{eig}\,(\mathbf{X})$	eigenvalues of $\mathbf{X}$
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A} > \mathbf{B}$	the matrix $\mathbf{A} - \mathbf{B}$ is positive definite
$\mathbf{A} \geq \mathbf{B}$	the matrix $\mathbf{A} - \mathbf{B}$ is positive semi-definite
$\|\underline{x}\|$, $\|\mathbf{A}\|$	norm of vector $\underline{x}$ (Euclidean) or matrix $\mathbf{A}$ (Frobenius)
$\mathbb{N}$	natural numbers
$\mathbb{R}$	real numbers
∇	del (nabla) operator
$p(a)$	probability density function of a
$p(a\|b)$	conditional probability density function of a given b
$\mathrm{E}\,\{A\}$	expected value of A
$\mathrm{E}\,\{A\|\,B\}$	expected value of A conditioned on B
$\mathcal{N}(\underline{\mu}, \mathbf{\Sigma})$	Gaussian with mean $\underline{\mu}$ and covariance Σ
$\mathcal{N}(\underline{x}\|\underline{\mu}, \mathbf{\Sigma})$	Gaussian with mean $\underline{\mu}$ and covariance Σ evaluated at $\underline{x}$
$\underline{\boldsymbol{x}}_k$	system state at time step k
$\underline{\boldsymbol{y}}_k$	measurement at time step k
$\underline{u}_k$	control input at time step k
$\underline{\boldsymbol{w}}_k$	process noise at time step k
$\underline{\boldsymbol{v}}_k$	measurement noise at time step k
$\mathcal{I}_k$	information set available to the controller at time step k
$\mathcal{J}$	cost function

Smooth Nonlinear Systems

$\underline{a}_k(\cdot)$	system dynamics at time step k
$\underline{h}_k(\cdot)$	measurement equation at time step k
$\delta(\cdot)$	Dirac deltat distribution
$p_k^x(\underline{x})$	probability distribution of the state at time step k
$\bar{p}_k^x(\underline{x})$	state probability distribution at time step k of the reference trajectory
$\tilde{p}_k^x(\underline{x}_k)$	predicted probability distribution of the state at time step k
$\mathbf{C}_k$	covariance of a distribution at time step k
$\mathcal{C}_k(\cdot)$	cost function at time step k
$V_k(\cdot)$	value function at time step k
$\underline{\pi}_k(\cdot)$	control policy at time step k
$\bar{\underline{u}}_k$	reference control input at time step k

$k(\cdot,\cdot)$	covariance function
$\mathbf{X}$	training inputs of a GP
$\mathbf{\Sigma}$	GP kernel
$\underline{x}_*$	vector-valued test input of a GP
d_*	test in put to a GP provided in terms of a probability distribution
y_*	GP output at a test input
$D(d_i, d_j)$	distance measure between the probability distributions d_i and d_j

Markov Jump Linear Systems

$\boldsymbol{\theta}_k$	mode at time step k
μ_k^i	probability of being in mode i at time step k
p_{ij}	transition probability from mode i into mode j
$\mathbf{T}$	transition probability matrix
$\rho(\mathbf{A})$	spectral radius of matrix $\mathbf{A}$
$\mathbb{1}_A$	indicator function of event A
$\mathbf{A}_{\boldsymbol{\theta}_k}$	value of the matrix $\mathbf{A}$ for the mode $\boldsymbol{\theta}_k$
$\mathbf{X}_k^i$	closed-loop second moment of the state at time step k
$\mathbf{X}_\infty^i$	steady-state value of the closed-loop second moment of the state
$\tilde{\mathbf{X}}_k^i$	closed-loop second moment of the augmented state at time step k
$\mathbf{X}_{m,k}^i$	partition m of closed-loop second moment of the augmented state at time step k
$\mathcal{H}$	Hamiltonian of a dynamic system

Kurzfassung

Ein Agent, sei es ein Mensch, ein Tier oder ein künstliches autonomes System, der mit seiner Umgebung interagiert, kann als ein dynamisches System modelliert werden, dessen Zustand sowohl den gegenwärtigen Zustand des Agenten als auch den seiner Umgebung beinhaltet. Zum Beispiel kann der Zustand eines Roboters, welcher sich in einem Gebäude bewegt, aus der Pose des Roboters sowie einer Karte des Gebäudes in Form von Landmarken bestehen. Normalerweise kennt der Agent seinen Zustand nicht genau und muss seine Handlungsentscheidungen basierend auf Beobachtungen treffen. Diese Beobachtungen, welche in Form von Messungen vorliegen, liefern in der Regel kein vollständiges Bild der Umgebung oder sind verrauscht. Zudem können die Modelle des Agenten und seiner Umwelt ungenau sein, was zu zusätzlichen Unsicherheiten führt. Im stochastischen Formalismus werden sowohl das Messrauschen als auch die Modellunsicherheiten als Zufallsvariablen modelliert, deren Wahrscheinlichkeitsverteilungen entweder bekannt sind oder geschätzt werden müssen werden.

Das beschriebene Problem der Entscheidungsfindung unter Unsicherheit wird in der stochastischen Regelungstheorie behandelt. Entsprechend den Prinzipien der Systemtheorie kann die Interaktion des Agenten mit seiner Umwelt als eine Regelungsschleife modelliert werden, in der die Entscheidungen über das Verhalten des Agenten von einem Regler getroffen werden, welcher selbst ein Teilsystem des Agenten darstellt. Basierend auf verfügbaren Informationen, welche aus dem Systemmodell sowie den Statistiken der Unsicherheiten bestehen, kann der Regler eine Schätzung des Systemzustands in Form einer Wahrscheinlichkeitsverteilung vorhalten, die mittels nichtlinearer Filtermethoden aus den empfangenen Messungen gewonnen wird. Diese Zustandsschätzung wird anschließend dazu verwendet, Stellwerte zu berechnen, welche die Handlungsentscheidungen darstellen. Um optimale Entscheidungen treffen zu können, die ein Gütekriterium in Form einer Kostenfunktion[1] optimieren, muss ein stochastischer Regler nicht nur die bisher verfügbaren Informationen, sondern auch sein zukünftiges Handeln, welches auf zukünftiger Information basieren wird, berücksichtigen. Solche optimalen Regler werden im Englischen als Closed-Loop Regler bezeichnet. Zudem findet eine Unterteilung der Regler in Regler mit einem endlichen und einem unendlichen Planungshorizont statt. Nur um einige wenige Beispiele zu nennen, wird die optimale stochastische Regelung dank dem hohen mathematischen Abstraktionsniveau in Robotik, bei automatisiertem Fahren und Fahrerassistenzsystemen, bei der Steuerung von unbemannten Fahrzeugen, in chemischen Anlagen, und bei Betrachtung von biologischen Systemen angewandt.

Die vorliegende Arbeit befasst sich mit der optimalen stochastischen Regelung von zeitdiskreten Systemen mit wertekontinuierlichen Zuständen und Mess- und Stellwerten. Insbesondere werden zwei Arten von stochastischen Systemen betrachtet: (1) nichtlineare Systeme mit stetiger

[1] In der Theorie der partiell beobachtbaren Markov'schen Entscheidungsprozesse (engl. POMDPs) findet anstelle einer Kostenminimierung eine Maximierung einer Belohnungsfunktion statt.

Dynamik und (2) Systeme mit nichtstetiger Dynamik. Wie bereits erwähnt, muss bei optimaler stochastischer Regelung der Einfluss von aktuellen Entscheidungen auf die Qualität zukünftiger Zustandsschätzung berücksichtigt werden, da diese ihrerseits die zukünftige Entscheidungsfindung beeinflusst. Diese Verkoppelung von Entscheidungsfindung und Zustandsschätzung führt dazu, dass das betrachtete Regelungsproblem außer in wenigen, sehr speziellen Fällen unlösbar wird. Aus diesem Grund sind Approximationen notwendig. Ein wichtiger Ansatz für die erste Klasse der betrachteten Systeme stellt die Trajektorienoptimierung dar, bei der die erwartete Trajektorie des Systems so geformt wird, dass die Kostenfunktion des konkreten Regelungsproblems ein Minimum annimmt. In der Regel läuft die Trajektorienoptimierung iterativ ab. In jedem Iterationsschritt wird zunächst die Kostenfunktion entlang der Trajektorie approximiert. Basierend auf dieser Approximation wird anschließend eine neue, bessere Trajektorie berechnet. Im ersten Teil dieser Arbeit wird ein neuartiges Approximationsverfahren vorgestellt, welches eine statistische Taylorreihenentwicklung zweiter Ordnung der Kostenfunktion entlang der aktuellen Trajektorie durchführt. Dieser Ansatz kann als eine Weiterentwicklung des Stands der Forschung angesehen werden. Zum einen nehmen verwandte Verfahren an, dass der geschätzte Systemzustand normalverteilt sei. Diese Einschränkung muss im vorgeschlagenen Verfahren nicht getroffen werden. Vielmehr können Zustandsschätzungen in Form von beliebigen, mittels Partikeln repräsentierten Wahrscheinlichkeitsverteilungen vorliegen. Und zum anderen ist die Approximation der Kostenfunktion robuster gegenüber Fehlern bei der Zustandsschätzung als bei Methoden aus der Literatur. Die Performanz des vorgeschlagenen Verfahrens wird in einer Simulation gezeigt.

Als Alternative zur Approximation der Kostenfunktion mittels statistischer Taylorreihenentwicklung zweiter Ordnung, wird im weiteren Verlauf der Arbeit ein nichtparametrisches Verfahren vorgeschlagen, dass auf Gauß-Prozessen (GP) basiert. Die klassische Definition eines GP erlaubt nur deterministische vektorwertige Eingänge. Allerdings muss im vorliegenden Szenario die Kostenfunktion entlang Trajektorien approximiert werden, welche als eine Menge von allgemeinen Wahrscheinlichkeitsdichten vorliegen. Aus diesem Grund wird die klassische Definition eines GP erweitert und eine neuartige Formulierung vorgeschlagen, welche den GP über dem Raum von Wahrscheinlichkeitsdichten definiert. Die Grundidee des vorgeschlagenen Ansatzes besteht darin, die Kovarianzfunktionen, welche einen GP beschreiben, als Funktionen des Abstands zwischen den Wahrscheinlichkeitsdichten, die als Eingänge dienen, zu definieren. Im Rahmen der vorliegenden Arbeit, werden ausgewählte, anwendbare Abstandsmaße diskutiert und das vorgeschlagene Verfahren mit einer Methode verglichen, die auf sogenannten mittleren Kovarianzfunktionen basiert. Außerdem ist es wichtig anzumerken, dass der vorgeschlagene Ansatz, Eingänge, welche in Form von Dichten gegeben sind, über ein Abstandsmaß auf nichtnegative reelle Zahlen abzubilden sich nicht nur auf GP beschränkt, sondern auch in anderen nichtparametrischen stochastischen Funktionsapproximationsmethoden angewandt werden kann.

Die bisher genannten Beiträge aus dem ersten Teil der Arbeit zielen auf eine Anwendung in übergeordneten Regelungsschleifen ab, wie sie zum Beispiel in Robotik vorliegen. Im zweiten Teil der vorliegenden Arbeit wird die Regelung auf tieferen Systemebenen betrachtet. Insbesondere liegt der Fokus auf Systemen, deren Dynamik sich abrupt ändert. Solche Systeme

werden durch eine endliche Menge von kontinuierlichen Dynamiken modelliert, zwischen denen entsprechend dem Wert einer diskretwertigen Zustandsvariable umgeschaltet wird. Da solche Systeme sowohl kontinuierliche als auch diskrete Zustandsvariablen besitzen, werden sie als hybride Systeme bezeichnet. In dieser Arbeit nehmen wir an, dass das Umschalten zwischen den kontinuierlichen Dynamiken, oder Moden des Systems, unabhängig vom Wert der kontinuierlichen Zustandsvariablen ist und durch eine Markov-Kette modelliert werden kann. Die Klasse solcher Systeme, bei denen zudem noch die kontinuierlichen Dynamiken linear sind, wird in Literatur als Markov Jump Linear Systems (MJLS) bezeichnet. Eine Modellierung mittels MJLS kommt bei der Betrachtung von Systemen mit Komponentenausfällen, Regelung über Kommunikationsnetzwerke u.v.m. in Frage. Stochastische optimale Regelung von MJLS mit bekannter Mode ist gelöst, da die Zustandsschätzung und Entscheidungsfindung entkoppelt sind, d.h. es gilt das sogenannte Separationsprinzip. Allerdings ist die Regelung von MJLS ohne Modenkenntnis immer noch ein aktiver Forschungszweig, da aufgrund fehlender Separation das Regelungsproblem unlösbar ist (s. Diskussion weiter oben). Die vorliegende Arbeit trägt zu dieser Forschung bei, indem ein linearer Regler für statische Regelung mit Ausgangsrückführung (engl. static output feedback) und unendlichem Planungshorizont sowie jeweils ein Regler für dynamische Regelung mit Ausgangsrückführung (eng. dynamic output feedack) mit endlichem und unendlichen Planungshorizont hergeleitet werden, die robust gegenüber der Mode sind, d.h. die Regler sind modenunabhängig. Die Bezeichnung statische Ausgangsrückführung bezieht sich darauf, dass der Regler keinen internen Zustand besitzt und Messwerte direkt zur Entscheidungsfindung verwendet. Im Gegensatz dazu wird bei dynamischer Ausgangsrückführung ein solcher Zustand vom Regler vorgehalten und auf Basis der gesammelten Messwerte aktualisiert. Im ersten Fall werden die Regler daher als statisch und im zweiten als dynamisch bezeichnet. Um die hergeleiteten Regler zu evaluieren, werden numerische Simulationen durchgeführt.

Aufgrund der Relevanz der betrachteten Probleme ist die vorliegende Arbeit sowohl von praktischer als auch von theoretischer Sicht interessant. Darüber hinaus bildet sie eine solide Grundlage für zukünftige Forschung.

Abstract

An agent, be it a human, an animal, or an artificial autonomous system, that interacts with its environment can be modeled as a dynamic system, whose state incorporates the current state of the agent and the state of its relevant surroundings. For example, the state of a robot that moves inside a building can consist of the robot's pose and the map of the building provided, e.g., in form of landmarks. Usually, this state cannot be observed and the agent has to make its decisions based on observations. Often, these observations in form of measurements do not provide a complete picture of the world or they are subject to noise. Moreover, the agent's and the world's models may be imprecise, which introduces additional uncertainty. In stochastic formalism, the measurement noise and model uncertainties are represented as random variables with known probability distributions or probability distributions that have to be estimated.

The described problem of decision making under uncertainty is addressed within the framework of stochastic control. Applying the principles of systems theory, the agent's interaction with the environment can be modeled as a control loop and the decisions are made by the controller, which is a subsystem of the agent. Based on the available knowledge that consists of the dynamic system model and the uncertainty statistics, the controller can maintain an estimate of the system state in terms of a probability distribution obtained from received measurements using nonlinear filtering methods. This estimate is then used to determine control actions, i.e., make decisions. In order to be able to make optimal decisions, i.e., decisions that optimize a performance criterion defined in terms of a cost function[2], a stochastic controller not only has to consider the information available to it at the current time step, but also has to take into account its future decision making based on the information that will become available later. Such optimal controllers are referred to as stochastic closed-loop optimal controllers. Depending on the length of the planning horizon, these controllers are divided into finite- and infinite-horizon controllers. Due to the high level of abstraction stochastic optimal control can be applied in robotics, assisted and autonomous driving, unmanned vehicles, process control, finance, and biological systems, to name only a fraction of possible applications.

In this thesis, we address stochastic closed-loop optimal control of discrete-time dynamic systems with continuous-valued states, measurements, and control actions. In particular, we consider two classes of stochastic systems: (1) nonlinear systems with smooth dynamics and (2) systems with abruptly changing dynamics. As mentioned above, closed-loop optimal control of stochastic systems requires consideration of the influence of current decision making on the quality of future state estimation, because this in turn affects future decision making. This coupling of decision making and state estimation makes the considered control problem intractable unless in a few very special cases. For this reason, approximate but tractable approaches are of interest.

[2] In the framework of partially observable Markov decision processes, we have to maximize a reward function instead of minimizing a cost function.

An important approach for the first class of considered systems is *trajectory optimization*, where the expected trajectory of the system is designed such that the cost function associated with the considered problem is minimized. This approach usually has to be performed iteratively. In each iteration step, the cost function is approximated along the current trajectory and a new better trajectory is computed using this approximation. In the first part of this thesis, we present a novel approximation scheme that performs *second-order statistical Taylor expansion* of the cost function along the current expected system trajectory. This approach extends state-of-the-art approaches that require the agent's state estimates to be Gaussian distributions to arbitrary probability distributions that are represented using particles. Additionally, the cost function approximation of the presented approach is more robust to errors in state estimation compared to methods available in literature. The proposed algorithm is demonstrated in a numerical example.

As an alternative to cost function approximation based on the second-order statistical Taylor expansion, we present a non-parametric approximation method that is based on Gaussian Processes (GPs). The classical GP formulation allows only for deterministic vector-valued inputs. However, in the considered scenario, the cost function has to be approximated along trajectories given in terms of a set of probability distributions. For this reason, we extend the classical GP formulation and describe a novel framework for *GPs defined over the space of probability distributions*. The main notion of our approach is to define the covariance functions that determine the GP as functions of the distance between the probability distributions provided as inputs. In the course of this thesis, we discuss a selection of admissible distance measures and compare our approach with a method based on the notion of mean covariance functions. Please note that the presented idea to map inputs that are provided in terms of probability distributions to non-negative real numbers is not limited to GPs but can be applied in other non-parametric Bayesian function approximation methods.

The methods presented in the first part of the thesis aim at high-level control, e.g., in robotic applications. In the second part of the thesis, we address control on lower levels and consider systems with abruptly changing dynamics. These systems are assumed to be modeled using a finite set of continuous dynamics and a discrete-valued variable whose value corresponds to the currently active continuous dynamic. Such systems are usually referred to as *hybrid systems* because they comprise both continuous- and discrete-valued dynamics. Furthermore, we assume that the switching between the different continuous-valued dynamics, or *modes* of the systems, is independent of its continuous-valued state and can be modeled using a Markov chain. In literature, this particular system class is referred to as the *Markov Jump Linear Systems* (MJLSs), if additionally the dynamics of each individual mode are linear. MJLS can be encountered, e.g., in finance, control of systems with component failures, control over data networks, and many more. Stochastic optimal control of MJLS with known mode was successfully solved, because state estimation and decision making are decoupled, i.e., the *separation principle* holds. However, control of MJLS with unknown mode is an active research field, because there is no such separation and the control problem becomes intractable (see discussion above). In this thesis, we contribute to this research and derive a linear infinite-horizon static output-feedback controller, and a finite- and an infinite-horizon linear dynamic

output-feedback control laws that are robust to the mode, i.e., mode-independent. The term static output feedback refers to controllers that do not maintain an internal state estimate of the agent and its surrounding, and operate solely on the measurements. On the other hand, dynamic output-feedback controllers maintain an internal state estimate using the received measurements. In the latter case, the controllers are denoted dynamic and in the former they are referred to as static. Numerical examples are used to demonstrate the derived controllers.

We think that the results presented in this thesis are relevant both from the theoretical and the practical points of view and believe that they lay a solid foundation for future research.

CHAPTER 1

Introduction

An agent, be it a human, an animal, or an artificial autonomous system, that interacts with its environment pursues some sort of a goal. To achieve this goal, the agent makes decisions based on its current internal state and the observations of its surroundings. For example, the state of a robot that moves inside a building could consist of the robot's pose and the map of the building provided, e.g., in form of landmarks, while observations could be the measured distances to the landmarks and measurements of the robot's movement. Another example is a batch reactor whose state comprises the volume of the processed material, its temperature at selected positions, and the speed of the rotating agitator, whereas only the temperature at a single position and the power consumption by the agitator and the heater are measured.

Following the principles of systems theory, the agent's decision making process can be formalized as a control loop (see Fig. 1.1). In this loop, the *controller* represents the subsystem of the agent that makes the decisions, the *process* represents the remaining part of the agent and the relevant part of the environment, the *sensor* models a system that generates observations or measurements of the agent's and the environment's states, and the *actuator* translates controller's decisions or control inputs into physical quantities that affect the states of the agent and the environment. Often, the actuator is omitted in block diagrams, because it is modeled as a part of the process.

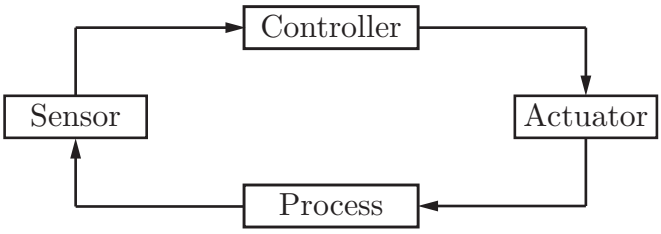


Figure 1.1: Block diagram of a general control loop.

Finally, the controller requires some sort of performance criterion that provides a mathematical formulation of the agent's goal. In optimal decision making, this performance criterion is represented as a cost function to be minimized or as a reward function that must be maximized[1].

[1] Without loss of generality, we will consider only the cost function minimization, because the two concepts are interchangeable.

Additionally, the controller can use the dynamic models of the agent and the environment in order to anticipate future system behavior. In this case, the optimization is denoted model-based or model-predictive. Furthermore, the controllers are divided into finite- and infinite-horizon controllers according to the length of the planning horizon.

In practical applications, the measurements received by the controller do not provide the complete state of the process or they are subject to measurement noise. Moreover, the agent's and the world's models used by the controller may be imprecise, which introduces additional uncertainty. Finally, the agent can be disturbed by external forces such as friction or wind. In stochastic formalism, these uncertainties are modeled as random variables with known, estimated, or assumed probability distributions. The controller's task then consists in computing optimal control inputs that take these uncertainties into account, because ignorance towards them results in poor performance in the best case as demonstrated by Witsenhausen in [172] or leads to instability in the worst case.

The described problem of optimal decision making under uncertainty is almost omnipresent and can be found, e.g., in robotics, assisted and autonomous driving, unmanned vehicles, process control, finance, and biological systems to name only a small fraction of possible applications [7, 133]. This problem can be addressed within the frameworks of stochastic optimal control and partially-observable Markov decision processes (POMDPs). As the name indicates, stochastic optimal control emerged from the field of control systems engineering, while POMDP theory has its origin in computer science. From the mathematical point of view, both frameworks are equivalent and the main difference is that stochastic optimal control usually assumes a model in terms of differential or difference equations that can be obtained from a physical model, whereas POMDPs usually assume a given probabilistic system model that is determined by state transition probability distribution and measurement likelihood. Without loss of generality, we will describe and address the considered problems from the point of view of stochastic optimal control in this thesis.

During operation, the controller continuously receives measurements. As mentioned above, these measurements are noisy and they do not necessarily provide information about the entire system state. Furthermore, storing all the measurements is impossible for long and infinite operation times. However, using the available knowledge that consists of the dynamic model of the process and the uncertainty statistics, the controller can maintain an estimate of the process' state in terms of a probability distribution obtained from collected measurements using (non-)linear filtering methods [137]. This estimate constitutes sufficient statistics of the control problem, because it provides all the information that is necessary for controller's decision making[2] [15]. In this way, a model-based controller for stochastic systems consists of an estimator or filter that generates the state estimates and a regulator that maps these state estimates to control inputs (see Fig. 1.2).

In order to be truly optimal, a controller not only has to make its decisions based on the information available up to the current time step. But it also has to take into account the influence of its decisions on the future decision making that will be based on the information that

[2]Please note that this is a rather theoretical concept because estimating and maintaining general non-parametrized probability distributions is intractable.

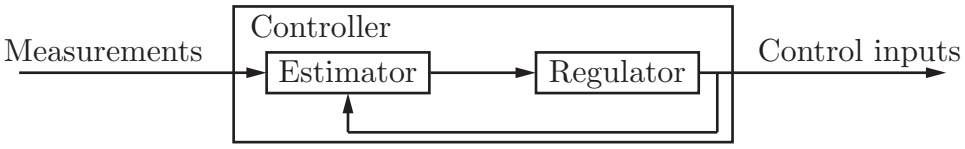


Figure 1.2: Structure of a stochastic optimal controller.

will become available later [11]. This coupling between decision making and information gain is referred to as the *dual effect* and controllers that take it into account are denoted *closed-loop*. The presence of the dual effect is exploited in adaptive control in order to improve the quality of the state or parameter estimate. This exploitation is referred to as active learning [156, 155]. However, the dual effect makes the calculation of closed-loop optimal stochastic controllers intractable except in a few important but very special cases. For this reason, approximate closed-loop approaches that achieve at least a local minimum are of interest [16].

1.1 Considered Problems and Contribution

In this thesis, we address stochastic closed-loop optimal control of discrete-time dynamic systems with continuous-valued states, measurements, and control inputs. In particular, we consider two classes of stochastic systems: (1) nonlinear systems with smooth dynamics and (2) systems with abruptly switching dynamics that are linear between the switches.

1.1.1 Control of Nonlinear Systems with Smooth Dynamics

As mentioned above, the coupling of control and estimation makes stochastic control problems intractable except in a few very special cases. This issue demands approximate but tractable approaches. An important approximate approach to finite-horizon stochastic optimal control is the *trajectory optimization*, where the expected trajectory of the system is designed such that the cost function associated with the considered problem is minimized. This approach usually has to be performed iteratively. In each iteration step, the cost function is approximated along the current trajectory and a new, better trajectory is computed using this approximation.

In this thesis, we present two novel cost function approximation schemes that can be used for trajectory optimization. The first scheme performs second-order statistical Taylor expansion of the cost function along the current expected system trajectory. This approach extends state-of-the-art approaches that require the agent's state estimates to be Gaussian distributions to arbitrary probability distributions that are represented using particles. Additionally, the cost function approximation of the presented approach is more robust to errors in state estimation compared to methods available in literature. It turns out that the proposed second-order statistical Taylor approximation of the cost function induces a linear controller, whose derivation is presented and whose performance is demonstrated in a numerical example.

Even in the vicinity of a reference trajectory, a quadratic approximation and the induced linear controller may not always be sufficiently good. For this reason, we propose a second approximation scheme for the costs along a reference trajectory that uses a non-parametric approximation method based on Gaussian Processes (GPs). The classical GP formulation allows only for deterministic vector-valued inputs. However, in the considered scenario, the cost function has to be approximated along trajectories given in terms of a set of probability distributions. For this reason, we extend the classical GP formulation and describe a novel framework for GPs defined over the space of probability distributions. The main notion of our approach is to define the covariance functions that determine the GP as functions of the distances between the probability distributions provided as inputs. In the course of this thesis, we discuss a selection of admissible distance measures and compare the presented approximation method with a state-of-the-art approach in a numerical example. The presented ideas are not limited to GPs but can be extended to other non-parametric Bayesian function approximation methods.

1.1.2 Control of Systems with Abruptly Changing Dynamics

In the second part of the thesis, we consider systems with abruptly changing dynamics. These systems are assumed to be modeled using a finite set of continuous dynamics and a discrete-valued variable whose value corresponds to the currently active continuous dynamic. Such systems are usually referred to as *hybrid systems*, because they comprise both continuous- and discrete-valued dynamics. Furthermore, we assume that the switching between the different continuous-valued dynamics, or *modes* of the systems, is independent of its continuous-valued state and can be modeled using a Markov chain. In literature, systems that belong to this particular system class are referred to as the *Markov Jump Linear Systems* (MJLSs) if additionally the dynamics of each individual mode are linear.

Research on control of MJLS with observed mode is well-advanced. However, if the mode is not observed, there is dual effect, which requires approximations. In this thesis, we consider derivation of linear controllers that are robust to the mode, i.e., the controller parameters are independent of the mode. Under this assumption, we first address infinite-horizon control of MJLS where a noise-free mapping of the continuous-valued state is fed back to the controller that are then directly used to compute control inputs without a prior state estimation. We show that the resulting control problem is non-convex, but it can be convexified by introducing additional constraints. Unfortunately, these constraints depend on the particular state space representation and can be very conservative. Thus, we propose an iterative algorithm for computation of the controller parameters that yields better control performance. The presented approach is compared with state-of-the-art methods in a numerical example.

Next, we consider finite- and infinite-horizon control of MJLS without mode observation, where noisy measurements of the continuous-valued state are fed back to the controller. In order to derive iterative algorithms that compute the controller parameters (time-invariant for finite horizon and time-variant for the infinite horizon), we bound the considered problems from above. In the finite-horizon case, we additionally are able to show that the proposed algorithm

converges to a (local) minimum. We compare the proposed methods with the optimal controller that requires mode feedback.

1.2 Outline

This thesis is divided into two main parts according to the classification of the considered systems described above. The first part addresses optimal stochastic closed-loop control of nonlinear discrete-time systems with smooth dynamics. It begins with a motivation in Sec. 2.1 and states the mathematical problem formulation in Sec. 2.2. Then, in Sec. 2.3, we briefly describe the general solution approach and review related work in Sec. 2.4. The derivation of a linear controller based on statistical second-order Taylor expansion of the costs along a reference trajectory is presented and demonstrated by means of simulations in Sec. 2.5. The first part of the thesis finishes with Sec. 2.6, where we present a framework for Gaussian processes over inputs that are provided in terms of arbitrary probability distributions. The proposed framework can be used to approximate the so-called value function that encodes the optimal costs and/or the policy in stochastic control problems.

The second part of the thesis is devoted to linear mode-independent control of MJLS without mode observation. We begin with a motivation for MJLS in Sec. 3.1 and provide a brief introduction to this class of dynamic systems in Sec. 3.2. The important topic of stochastic stability of MJLS is presented in Sec. 3.3. A review of related work and a discussion on intractability of optimal control of MJLS with non-observed mode in Sec. 3.4 conclude the introductory part of the chapter on stochastic optimal control of MJLS without mode observation. In Sec. 3.5, we then address static output-feedback control, where we first provide sufficient feasibility conditions for control of MJLS with and without mode observation via static output feedback. Because these conditions are not invariant to the choice of the particular state space representation, we propose an iterative algorithm for computation of the control law parameters and evaluate it by means of simulations in a comparison with a state-of-the-art approach. Next, we address infinite-horizon dynamic output-feedback control of MJLS without mode observation in Sec. 3.6 and its finite-horizon counterpart in Sec. 3.7. The performance of the controllers computed using the proposed algorithms is demonstrated in a numerical example.

Finally, in the last chapter, we summarize the contribution of the presented work in Sec. 4.1 and conclude the thesis with an outlook on possible future research direction in Sec. 4.2.

CHAPTER 2

Control of Nonlinear Systems with Smooth Dynamics

Contents

In this chapter, we consider finite-horizon stochastic optimal control of nonlinear systems with smooth dynamics via trajectory optimization. The presented approach relies on local approximation of the cumulative costs along a reference trajectory. This trajectory is iteratively optimized in order to achieve a local minimum of the costs. Two approximation methods are discussed: (1) a method that performs a sample-based local Taylor expansion of the cumulative costs and (2) a Gaussian process regression approach. We begin with a motivating example in the next section and provide a formal problem definition in Sec. 2.2. The general solution approach is discussed in Sec. 2.4, where we also review the related work. Then, the two proposed approaches are presented in Secs. 2.5 and 2.6. The material presented in this section is original and has not been published elsewhere.

2.1 Motivation

As discussed in the introduction, nonlinear stochastic optimal control problems are subject to the dual effect that makes the calculation of optimal control inputs intractable except in a few special cases. The simplest approaches to address this issue are: (i) to neglect the uncertainty and solve the resulting deterministic control problem and (ii) to linearize the system at the operation point and use well-understood methods from linear control. However, ignorance towards stochastic system disturbances results in worse performance and can even destabilize the control loop. On the other hand, linearization of an inherently nonlinear system is reasonable only as long as the system state stays close to the linearization point, i.e., only for small-scale behavior where the control signals and disturbances are sufficiently small.

The two approaches, neglecting the uncertainty and linearization around a single operation point, are not suitable in control problems, where closed-loop control with transient system behavior is required, such as in high-level control of robots and unmanned vehicles. The robotic scenario depicted in Fig. 2.1 demonstrates the difference between a closed-loop optimal controller and an ignorant controller. In the depicted scenario, the robot has the task to reach a target. For localization, it uses distance measurements to a house that serves as a landmark. The noise that affects the measurements is state-dependent. In particular, it is small near the house and large far from it (indicated by a shadow in Fig. 2.1). Consequently, the robot's localization is more precise near the landmark and less precise far from it. We assume that the robot's position is represented as a Gaussian distribution and use ellipses to depict it. The mean of a Gaussian corresponds to the center of the ellipse and the covariance to ellipse's size. In this scenario, linearization is not an option, because the robot dynamics are highly nonlinear in its orientation and the path of the robot may include many different turns. Furthermore, for large covariances, a linearization around the mean of the Gaussian may result in a bad approximation quality as we will discuss in Sec. 2.5.

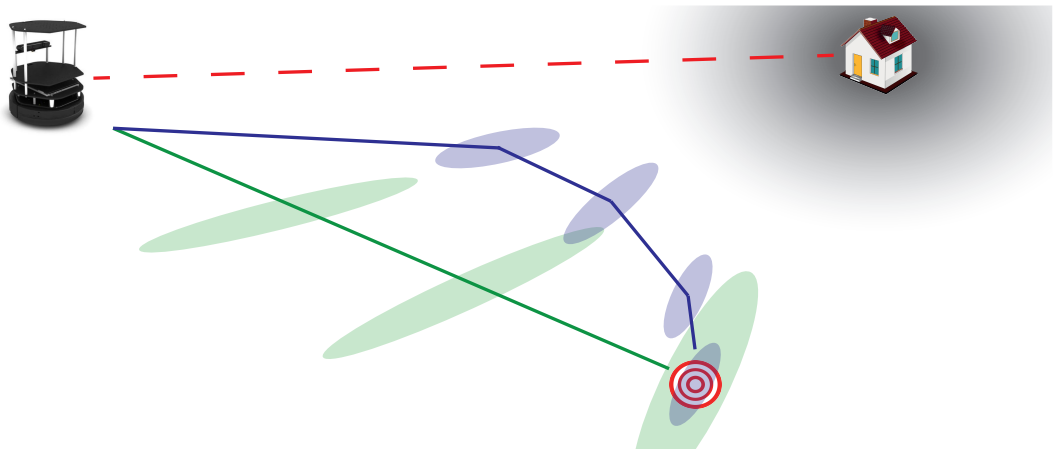


Figure 2.1: A robotic path planning example with state-dependent measurement noise.

In the considered scenario, an ignorant controller that does not take the localization quality into account would most probably plan the green path as shown in Fig. 2.1. By doing so, its localization would be poor, which is indicated by green ellipses. Although the mean position of the robot remains on the path, there is a large probability that it deviates from it. This issue can become critical at the end of the path, because the probability is small that the robot will actually be within the target area. On the other hand, a closed-loop optimal controller would plan a path that is similar to the blue path in Fig. 2.1. By doing so, it can obtain a much better quality of localization, which is indicated by blue ellipses that are much smaller than the green ellipses, which represent the localization quality of the ignorant controller. In this way, we can expect that the robot more likely will be within the target area at the end of the path. Other scenarios with a closed-loop optimal controller being the right choice include problems with multi-modal probability distributions of the state, where it is desirable to eliminate one of the modes through the choice of appropriate control inputs, or problems, where observability of some of the state components, i.e., the possibility to infer the probability distribution of these components from measurements, depends on the state.

2.2 Problem Formulation

According to the structure of the control loop depicted in Fig. 1.1, the individual components of the loop are assumed to be as follows. We consider processes modeled by general nonlinear discrete-time stochastic systems with smooth dynamics

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k, \underline{\boldsymbol{w}}_k) \ , \tag{2.1}$$

where the time-variant nonlinear function $\underline{a}_k$ maps the current state of the system $\underline{\boldsymbol{x}}_k \in \mathbb{R}^{n_x}$, the control input $\underline{u}_k \in \mathbb{R}^{n_u}$, and a realization of the independent process noise $\underline{\boldsymbol{w}}_k \in \mathbb{R}^{n_w}$ to the next system state $\underline{\boldsymbol{x}}_{k+1}$. The process noise is assumed to be modeled as a time-variant stochastic process with probability distribution $p_k^w(\underline{w}_k)$. The measurements received by the controller are obtained by the sensor according to

$$\underline{\boldsymbol{y}}_k = \underline{h}_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{v}}_k) \ , \tag{2.2}$$

where the state $\underline{\boldsymbol{x}}_k$ and the realization of the measurement noise $\underline{\boldsymbol{v}}_k \in \mathbb{R}^{n_v}$ are mapped to the measurement $\underline{\boldsymbol{y}}_k \in \mathbb{R}^{n_y}$ via a time-variant measurement function $\underline{h}_k$. The probability distribution $p_k^v(\underline{v}_k)$ of the independent measurement noise is known. In this chapter, we demand that we can draw samples from the probability distributions $p_k^w(\underline{w}_k)$ and $p_k^v(\underline{v}_k)$. However, most related state-of-the-art approaches are forced to make stricter assumptions.

For the system (2.1)–(2.2), we seek a closed-loop controller that minimizes the finite-horizon cost function

$$\mathcal{J} = \mathrm{E}\left\{ \mathcal{C}_K(\underline{\boldsymbol{x}}_K) + \sum_{k=0}^{K-1} \mathcal{C}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k) \right\} \ , \tag{2.3}$$

where $K \in \mathbb{N}$ denotes the length of the planning horizon and $\mathcal{C}_K(\cdot)$ and $\mathcal{C}_k(\cdot,\cdot)$ are functions that compute the costs at the time steps K and k for $k = 0, 1, \ldots, K-1$, respectively.

As mentioned in the introduction, in order to be truly optimal, the controller needs to consider the influence of the current decision making on future decision making that is based on the information that will become available later. For this reason, rather than being interested in particular control inputs $\underline{u}_k$, we seek control laws or policies $\underline{\pi}_k(\mathcal{I}_k)$ that map the information sets $\mathcal{I}_k$ available to the controller at time steps k to particular control inputs $\underline{u}_k$, i.e., $\underline{u}_k = \underline{\pi}_k(\mathcal{I}_k)$. The information available to the controller at time step k consists of the process model (2.1), the measurement model (2.2), the cost function (2.3), the noise statistics in form of $p_k^w(\underline{w}_k)$ and $p_k^v(\underline{v}_k)$, the probability distribution $p_0^x(\underline{x}_0)$ that describes the initial system state $\underline{\boldsymbol{x}}_0$, the measurements $\underline{\boldsymbol{y}}_{1:k}$, and the control inputs $\underline{u}_{0:k-1}$. For brevity, we will not explicitly mention the models and the noise statistics in the information sets $\mathcal{I}_k$. Thus, we can define

$$\mathcal{I}_0 = \{p_0^x(\underline{x}_0)\} \qquad \text{and} \qquad \mathcal{I}_k = \{\underline{\boldsymbol{y}}_k, \underline{u}_{k-1}, \mathcal{I}_{k-1}\} \; . \tag{2.4}$$

Please note that alternatively, the considered control problem can be addressed in the receding-horizon framework, where the control law is periodically recomputed, e.g., at each time step. In this way, the control approach proposed in this chapter can be extended to infinite operation time at cost of additional computational burden.

To conclude this section, we want to emphasize that the described problem is referred to as the *partial-information* stochastic control problem [15] because the state $\underline{\boldsymbol{x}}_k$ is not available to the controller. However, by exploiting the Markov property[1], it can be cast into the so-called *full-information* problem by reformulating it in terms of the estimate $p_k^x(\underline{x}_k)$ of the system state $\underline{\boldsymbol{x}}_k$ that condenses the information set $\mathcal{I}_k$. Then, the policy can also be written in terms of the state estimate $p_k^x(\underline{x}_k)$ according to $\underline{u}_k = \underline{\pi}_k(p_k^x(\underline{x}_k))$. Because the probability distribution of the state provides all the information that is required for decision making, it is referred to as *sufficient statistic* of the control problem. Additionally, this reformulation is more convenient than maintaining the information set $\mathcal{I}_k$ that grows with time. However, the described concept of reformulation of partial-information problems into full-information problems is only theoretical, because estimating and storing general non-parametrized state estimates $p_k^x(\underline{x}_k)$ is intractable.

Please note that we do not consider the problem of state estimation in this chapter. We assume that an admissible filter has been selected beforehand that maintains the state estimate in form of a Dirac mixture

$$p_k^x(\underline{x}_k) = \sum_{i=1}^{M} \alpha_k^i \delta(\underline{x}_k - \underline{x}_k^i) \; , \tag{2.5}$$

which is a discrete distribution with $M \in \mathbb{N}$ particles at $\underline{x}_i \in \mathbb{R}^{n_x}$ and weights $\alpha_k^i \in \mathbb{R}_+$ for $i = 1, \ldots, M$, $\sum_{i=1}^{M} \alpha_k^i = 1$ over a continuous domain. Such filters are referred to as particle filters [6]. Furthermore, it is possible to employ filters that do not explicitly maintain the state estimate in form of a Dirac mixture, but where Dirac mixtures are generated as a byproduct during state estimation. These filters are, e.g., the Unscented Kalman Filter (UKF) [101], the Randomized UKF (RUKF) [56], the Cubature Kalman Filter (CKF) [5], and the Smart

[1] The Markov property holds when the current system state depends only on the previous state and not the states before that.

Sampling Kalman Filter (S2KF) [145]. Please keep in mind that although the non-parametric representation of an arbitrary state distribution using samples is optimal if the number of samples is infinite, its practical implementations with finite numbers of samples constitute an approximation. Also, the choice of a particular state estimator such as the UKF over the particle filter can limit the class of representable probability distributions, which affects the overall performance of the controlled system.

2.3 General Approach

The problem of minimizing (2.3) under the information structure (2.4) and subject to the model (2.1)–(2.2) can be addressed within the framework of *Dynamic Programming* (DP) [15]. The main notion of DP is to exploit the *Bellman's Principle of Optimality* [12] by reformulating the minimization of (2.3) as the recursion

$$\begin{aligned} V_K(\mathcal{I}_K) &= \mathrm{E}\left\{\mathcal{C}_K(\underline{\boldsymbol{x}}_K) \middle| \mathcal{I}_K\right\} \ , \\ V_k(\mathcal{I}_k) &= \inf_{\underline{u}_k} \mathrm{E}\left\{\mathcal{C}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k) + V_{k+1}(\mathcal{I}_{k+1}) \middle| \mathcal{I}_k\right\} \ , \end{aligned} \tag{2.6}$$

where $V_k(\mathcal{I}_k)$ is referred to as the optimal *costs-to-go* or the *value function*[2] at time step $k = 0, \ldots, K$. Recursion (2.6) is also called *value iteration*. The optimal policy at time step k can be recovered according to

$$\underline{\pi}_k(\mathcal{I}_k) = \arg\inf_{\underline{u}_k} \mathrm{E}\left\{\mathcal{C}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k) + V_{k+1}(\mathcal{I}_{k+1}) \middle| \mathcal{I}_k\right\} \ .$$

Following the discussion in Sec. 2.2, recursion (2.6) can be written in terms of the state estimates $p_k^x(\underline{x}_k)$ that are constructed from the information sets $\mathcal{I}_k$ using stochastic filtering techniques [137]. Consequently, after resolving the expectation operators, we obtain

$$\begin{aligned} V_K(p_K^x(\underline{x}_K)) &= \int_{\mathbb{R}^{n_x}} \mathcal{C}_K(\underline{x}_K) p_K^x(\underline{x}_K) \, \mathrm{d}\underline{x}_K \ , \\ V_k(p_k^x(\underline{x}_k)) &= \inf_{\underline{u}_k} \int_{\mathbb{R}^{n_x}} \left[\mathcal{C}_k(\underline{x}_k, \underline{u}_k) + \int_{\mathbb{R}^{n_y}} p(\underline{y}_{k+1} | \underline{x}_k, \underline{u}_k) V_{k+1}(\widetilde{p}_{k+1}^x(\underline{x}_{k+1} | \underline{y}_{k+1}, \underline{u}_k)) \, \mathrm{d}\underline{y}_{k+1} \right] p_k^x(\underline{x}_k) \, \mathrm{d}\underline{x}_k \ , \end{aligned} \tag{2.7}$$

where $p(\underline{y}_{k+1} | \underline{x}_k, \underline{u}_k)$ is the probability distribution of observing the measurement $\underline{y}_{k+1}$ given a particular state $\underline{x}_k$ and a control input $\underline{u}_k$, i.e.,

$$p(\underline{y}_{k+1} | \underline{x}_k, \underline{u}_k) = \int_{\mathbb{R}^{n_x}} p(\underline{y}_{k+1} | \underline{x}_{k+1}) p(\underline{x}_{k+1} | \underline{x}_k, \underline{u}_k) \, \mathrm{d}\underline{x}_{k+1} \ .$$

[2] Please note that in the theory of partially-observable Markov decision processes, we maximize the expected reward rather than minimizing the costs. Therefore, the value function is maximized and the term costs-to-go is more appropriate in our context. Nevertheless, we will denote $V_k(\mathcal{I}_k)$ as value function.

The probability distribution $\bar{p}^x_{k+1}(\underline{x}_{k+1}|\underline{y}_{k+1})$ can be computed using the Bayes' law according to

$$\widetilde{p}^x_{k+1}(\underline{x}_{k+1}|\underline{y}_{k+1},\underline{u}_k) = \frac{p(\underline{y}_{k+1}|\underline{x}_{k+1}) \int\limits_{\mathbb{R}^{n_x}} p(\underline{x}_{k+1}|\underline{x}_k,\underline{u}_k) p^x_k(\underline{x}_k)\ \mathrm{d}\underline{x}_k}{\int\limits_{\mathbb{R}^{n_x}} p(\underline{y}_{k+1}|\underline{x}_k,\underline{u}_k) p^x_k(\underline{x}_k)\ \mathrm{d}\underline{x}_k}$$

for a given measurement $\underline{y}_{k+1}$, i.e., for each measurement $\underline{y}_{k+1}$ of the outer integral in (2.7).

Recursion (2.6) can be solved optimally only in a few very special scenarios, e.g., if the spaces of states, control inputs, and measurements are finite and sufficiently small [141, 142]. Another important example of a tractable DP recursion is the Linear Quadratic Gaussian (LQG) control [3], where the model (2.1)–(2.2) is linear with deterministic or white parameters[3] [44, 15], i.e.,

$$\begin{aligned} \underline{\boldsymbol{x}}_{k+1} &= \mathbf{A}_k \underline{\boldsymbol{x}}_k + \mathbf{B}_k \underline{u}_k + \mathbf{H}_k \underline{\boldsymbol{w}}_k\ , \\ \underline{\boldsymbol{y}}_k &= \mathbf{C}_k \underline{\boldsymbol{x}}_k + \mathbf{J}_k \underline{\boldsymbol{v}}_k\ , \end{aligned}$$

the noises $\underline{\boldsymbol{w}}_k$ and $\underline{\boldsymbol{v}}_k$ are Gaussian and independent and identically distributed, and the cost function is quadratic

$$\mathcal{J} = \mathrm{E}\left\{ \underline{\boldsymbol{x}}_K^\top \mathbf{Q}_K \underline{\boldsymbol{x}}_K + \sum_{k=0}^{K-1} \underline{\boldsymbol{x}}_k^\top \mathbf{Q}_k \underline{\boldsymbol{x}}_k + \underline{u}_k^\top \mathbf{R}_k \underline{u}_k \right\}$$

with deterministic or white positive semi-definite matrices $\mathbf{Q}_k$ and positive definite matrices $\mathbf{R}_k$. The tractability of the recursion is maintained because in the LQG control, the separation between control and estimation holds. Then, the optimal control law consists of a Kalman filter [102] that computes the state estimate and a linear regulator that maps the mean of the state estimate to control inputs.

However, if the control problem does not belong to any of the two cases mentioned above, recursion (2.6) is not tractable especially due to the following issues. First, an arbitrary, non-parameterized state distribution $p^x_k(\underline{x}_k)$ cannot be maintained. Second, in order to maintain the value function $V_k(p^x_k(\underline{x}_k))$, we need to store its value for every possible probability distribution $p^x_k(\underline{x}_k)$, which is not possible for non-parameterizable functions $V_k(p^x_k(\underline{x}_k))$. And finally, an efficient computation of the infimum of $V_k(p^x_k(\underline{x}_k))$ with respect to $\underline{u}_k$ may not be possible, for example, because $V_k(p^x_k(\underline{x}_k))$ is not convex in $\underline{u}_k$ but even more severe, because at each step of the algorithm that performs the optimization, we need to solve an inner integral with respect to $\underline{\boldsymbol{y}}_{k+1}$.

2.4 Related Work

In its general formulation, recursion (2.6) is intractable. And thus, approximate but tractable approaches are of interest. These approaches can be distinguished into approaches that perform

[3] The parameters are independent of the states, the measurements, and the control inputs and do not form a Markov chain, i.e., they are independent and identically distributed.

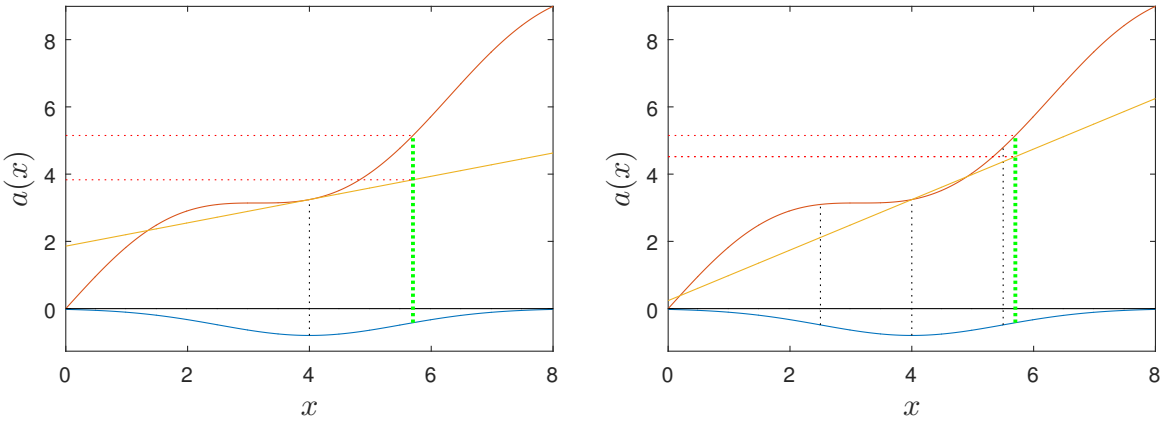
an explicit and approaches that perform an implicit approximation of the value function [16]. The approaches in the latter class usually simplify the considered problem or make assumptions regarding the control law. On the other hand, the approaches that belong to the former class either bound the value function or employ function approximation methods in order to approximate the value function (so-called fitted approximation). However, often an unambiguous classification is not possible, e.g., if the approximation methods from both classes are combined in some way. Thus, our review of related work is organized into approaches that maintain a global approximation of the value function and approaches that rely on its local approximation.

Most approaches to DP that aim at maintaining a global value function approximation belong to the class of point-based value iteration methods [130], where the costs-to-go are maintained for a set of possible state estimates. The costs-to-go at other state estimates are then approximated using function approximation methods. A popular approach to the stochastic optimal control problem where the value function is maintained globally was published in [149]. In this work, the authors use a randomly sampled Dirac mixture as the representation of the state estimate and approximate the value function using a nearest-neighbor method that uses the Kullback–Leibler (KL) divergence [136] in order to evaluate the "distance" between state estimates. Please note that the KL divergence does not satisfy the axioms for distance measures. Furthermore, it is not defined for Dirac mixtures. Therefore, the authors of [149] use a form of Parzen–window smoothing [128] to be able to apply the KL divergence. In order to handle continuous control inputs, the authors of [149] use a Q-learning approach [147], where they estimate a function of the costs-to-go for state estimates and control inputs instead of maintaining only the value function as in the classical value iteration. An important theoretical contribution to the theory of global DP approaches with a continuous state space was presented in [132], where the authors generalize the results from [141] and [142] for problems with a discrete state space. In particular, the authors of [132] demonstrate that the notion of α-vectors that encode the value function and the optimal policy in stochastic control problems with a discrete-valued state space can be extended to so-called α-functions. Moreover they show that if the spaces of the control inputs and measurements, and the planning horizon are finite, the optimal value function is Piece-Wise Linear Continuous (PWLC) in the state estimates. If these assumptions do not hold, the value function can be approximated arbitrarily close with PWLC functions of state estimates. Using these results, the authors propose a control approach, where the state estimates and the α-functions are represented with Gaussian mixtures, and an approach, where the state estimates are maintained in form of Dirac mixtures and α-functions in form of Gaussian mixtures. Representing both the state estimates and α-functions using Dirac mixtures is not possible. The results from [132] are combined with a state-aggregation heuristic in [24] to reduce the number of state estimates that have to be stored for approximation of the value function, and with a control law representation that uses a finite-state controller [84] in [9]. While the approaches so far maintain the state estimates in form of Dirac or Gaussian mixtures, the authors of [25] propose to restrict the state estimates to be Gaussian. By doing so, their approach becomes computationally far less expensive than the approaches from [149, 132, 24, 9]. However, due to this strict assumption, the method from [25] is only applicable for selected problem classes.

The approaches to stochastic control reviewed above have in common that they maintain a point-based global estimate of the value function for the considered state space. On the other hand, local approaches perform DP in a tube around a reference trajectory. To this end, they proceed as follows. First, a reference trajectory is generated using an initial control law. Then, the value iteration runs from the end of the planning horizon to its beginning and at each iteration step, the value function is locally approximated at the current state estimate. This approximation is used to improve the control law, which in the next DP run yields a better reference trajectory. This process is repeated until convergence of the costs (2.3). By doing so, the issue with maintaining an approximation of the value function as in global DP is traded for additional computational requirements. In the last decade, local DP approaches also referred to as *trajectory optimization* methods have gained in popularity. Especially, approaches that are based on extension of the LQG control to control of nonlinear systems have received an increased research interest. These approaches use second-order Taylor expansion of the value function and linearize the system dynamics. For the policy, they assume an affine control law. For this reason, they can be seen as a special case of Differential Dynamic Programming (DDP) [118]. A version of extended LQG for stochastic optimal control problems where the state is available to the controller was first presented in [150]. The approach from [150] is extended to problems with noisy measurements in [110]. In this work, the state estimate, and the measurement and the process noises are assumed to be Gaussian. As a state estimator the authors of [110] use an Extended Kalman Filter (EKF) [10] and assume a control law that is affine in the mean of the state estimate. Their approach to approximation of the value function consists in using a second-order Taylor expansion around the mean of the state estimate. We will refer to this approximation scheme as *EKF-based*. To compute the estimator gains and the gains of the control law, the algorithm from [110] performs an interleaved optimization with respect to these parameters. For the same setup, i.e., Gaussian state estimate, and process and measurement noises, and the usage of an EKF for state estimation, the authors of [158, 159] present a trajectory optimization method, where they perform Taylor expansion around the Gaussian state estimate (mean and square root of the covariance stacked in a vector) instead of only around the mean. To this end, they formulate the closed-loop system dynamics in terms of the Gaussian state estimate. Furthermore, they propose to use rapidly-exploring belief trees [26] for generation of the initial reference trajectory in order to address problems with obstacles. In contrast to the approaches from [110] and [158, 159], the methods proposed in [58] and [131] consider Maximum Likelihood measurements, i.e., the predicted measurements correspond to the mean of the Gaussian measurement distribution. By doing so, the dynamics of the state estimate are rendered deterministic.

2.5 Local Value Function Approximation

In this section, we present a method for approximation of the value function $V_k(p_k^x(\underline{x}_k))$ along a reference system trajectory using a statistical second-order Taylor expansion and derive a control law that employs this approximation scheme. Our approach can be seen as an extension of the EKF-based approximation from [110] and [159], where the costs-to-go are



(a) Approximation using the EKF-based approach (yellow) at the mean $x = 4$ (black dotted line) produces an absolute error of 1.3173.

(b) Approximation using a statistical approach with three sigma points (yellow) at $x = \{4, 2.5, 5.5\}$ (black dotted lines) produces an absolute error of 0.6289.

Figure 2.2: Linearization error of the function $a(x) = \sin(x) + x$ (red) at the true input $x = 5.7$ for the Gaussian distribution $\mathcal{N}(4, 1.5)$ (blue).

approximated only at the mean of the state estimate that is assumed to be Gaussian, to statistical sample-based expansion at the positions of the particles of the Dirac mixtures that are used to represent the state estimates $p_k^x(\underline{x}_k)$. Statistical approximation methods are widely applied in nonlinear Kalman filtering [109], because they tend to be more robust than EKF-based approximations. The latter observation is demonstrated in Fig. 2.2, where the function $a(x) = \sin(x) + x$ is linearized for the Gaussian probability distribution $\mathcal{N}(4, 1.5)$. The EKF-based linear approximation at the mean $x = 4$ produces an absolute error of 1.3173, while the sample-based linearization at the three samples $x = \{4,\ 2.5,\ 5.5\}$ yields an absolute error of 0.6289. Due to the advantage of sample-based approximation methods, we believe that the presented control approach constitutes an important contribution to the theory of stochastic optimal control.

2.5.1 Preliminaries

Before we present the derivation of the proposed control approach in the next section, we need to introduce some preliminaries. First, we show how to compute the empirical covariance of a Dirac mixture in Lemma 2.1 and second, we define a notion of statistical Taylor expansion in Definition 2.2.

Lemma 2.1. *For jointly Gaussian state and measurement, the empirical covariance of the probability distribution* $p_{k+1}^x(\underline{x}_{k+1})$ *can be computed from the state estimate* $p_k^x(\underline{x}_k) = \sum_{i=1}^M \alpha_k^i \delta(\underline{x}_k - \underline{x}_k^i)$ *using the model* (2.1)–(2.2)*, the control input* $\underline{u}_k$*, and sampled process and measurement noises* $p_k^w(\underline{w}_k) = \sum_{m=1}^R \beta_k^m \delta(\underline{w}_k - \underline{w}_k^m)$ *and* $p_{k+1}^v(\underline{v}_{k+1}) = \sum_{n=1}^S \gamma_{k+1}^n \delta(\underline{v}_{k+1} - \underline{v}_{k+1}^n)$ *according to*

$$\mathbf{C}_{k+1}^{xx|y} = \mathbf{C}_{k+1}^{xx} - \mathbf{C}_{k+1}^{xy}(\mathbf{C}_{k+1}^{yy})^{-1}(\mathbf{C}_{k+1}^{xy})^\top \ ,$$

where

$$\mathbf{C}_{k+1}^{xx} = \sum_{i=1}^{M}\sum_{m=1}^{R} \alpha_k^i \beta_k^m (\underline{x}_k^{[im]} - \hat{\underline{x}}_{k+1})(\underline{x}_k^{[im]} - \hat{\underline{x}}_{k+1})^\top \ ,$$
$$\mathbf{C}_{k+1}^{xy} = \sum_{i=1}^{M}\sum_{m=1}^{R}\sum_{n=1}^{S} \alpha_k^i \beta_k^m \gamma_{k+1}^n (\underline{x}_k^{[im]} - \hat{\underline{x}}_{k+1})(\underline{y}_{k+1}^{[imn]} - \hat{\underline{y}}_{k+1})^\top \ ,$$
$$\mathbf{C}_{k+1}^{yy} = \sum_{i=1}^{M}\sum_{m=1}^{R}\sum_{n=1}^{S} \alpha_k^i \beta_k^m \gamma_{k+1}^n (\underline{y}_{k+1}^{[imn]} - \hat{\underline{y}}_{k+1})(\underline{y}_{k+1}^{[imn]} - \hat{\underline{y}}_{k+1})^\top \ ,$$

with

$$\underline{x}_k^{[im]} = \underline{a}_k(\underline{x}_k^i, \underline{u}_k, \underline{w}_k^m) \ , \qquad \hat{\underline{x}}_{k+1} = \sum_{i=1}^{M}\sum_{m=1}^{R} \alpha_k^i \beta_k^m \underline{x}_{k+1}^{[im]} \ ,$$
$$\underline{y}_{k+1}^{[imn]} = \underline{h}_{k+1}(\underline{x}_{k+1}^{[im]}, \underline{v}_{k+1}^n) \ , \qquad \hat{\underline{y}}_{k+1} = \sum_{i=1}^{M}\sum_{m=1}^{R}\sum_{n=1}^{S} \alpha_k^i \beta_k^m \gamma_{k+1}^n \underline{y}_{k+1}^{[imn]} \ .$$

The above result can be found, e.g., in [168]. It is generally more convenient and less computationally demanding to use samples from the joint probability distribution of $\underline{\boldsymbol{x}}_k$, $\underline{\boldsymbol{w}}_k$, and $\underline{\boldsymbol{v}}_k$. Please observe that the covariance $\mathbf{C}_{k+1}$ is independent of $\underline{\boldsymbol{y}}_{k+1}$. Therefore, it remains the same before and after the measurement $\underline{\boldsymbol{y}}_{k+1}$ is fused into $p_{k+1}^x(\underline{x}_{k+1})$. This observation will be useful during the derivation of the proposed control law in the next section.

Next, we present our notion of statistical Taylor expansion for functions whose arguments are probability distributions that are given in terms of particles. Please note that statistical or Monte-Carlo evaluation of integrals, in particular of expected values, is a well-established technique. However, the notion of statistical Taylor expansion at a probability distribution given in terms of particles is a novel concept.

Definition 2.2. *Let a function a that maps a probability distribution $p^x(\underline{x})$ on $\mathbb{R}^{n_x}$ to a scalar value be of the form $a(p^x(\underline{x})) = \mathrm{E}\{g(\underline{\boldsymbol{x}})\}$, where $g : \mathbb{R}^{n_x} \to \mathbb{R}$. Then, the second-order Taylor expansion of $a(\cdot)$ at a Dirac mixture $\bar{p}^x(\underline{x}) = \sum_{i=1}^{M} \alpha^i \delta(\underline{x} - \underline{x}^i)$, $M \in \mathbb{N}$, $\underline{x}^i \in \mathbb{R}^{n_x}$, $\alpha^i \in (0, 1]$, $i = 1, \ldots, M$, $\sum_{i=1}^{M} \alpha^i = 1$ is given by*

$$a(p^x(\underline{x})) \approx \mathrm{E}\left\{\sum_{i=1}^{M} \alpha^i \left[g(\underline{x}^i) + \nabla g(\underline{x}^i)^\top (\underline{\boldsymbol{x}} - \underline{x}^i) + \frac{1}{2}(\underline{\boldsymbol{x}} - \underline{x}^i)^\top \nabla^2 g(\underline{x}^i)(\underline{\boldsymbol{x}} - \underline{x}^i)\right]\right\} \ , \qquad (2.8)$$

if the distributions $p^x(\underline{x})$ and $\bar{p}^x(\underline{x})$ are independent and close[4].

[4]The closeness of the probability distributions $p^x(\underline{x})$ and $\bar{p}^x(\underline{x})$ can be evaluated either using a distance measure for probability distributions [178] or in terms of the difference $a(p^x(\underline{x})) - a(\bar{p}^x(\underline{x}))$.

The result of (2.8) can be obtained by following the lines of the computation of the expected value of a function of random variable using Taylor expansion. It holds

$$\begin{aligned} a(p^x(\underline{x})) &= \mathrm{E}\left\{g(\underline{\boldsymbol{x}})\right\} \\ &= \mathrm{E}\left\{g(\underline{\widetilde{\boldsymbol{x}}} + (\underline{\boldsymbol{x}} - \underline{\widetilde{\boldsymbol{x}}}))\right\} \\ &= \mathrm{E}\left\{g(\underline{\widetilde{\boldsymbol{x}}}) + \nabla g(\underline{\widetilde{\boldsymbol{x}}})^\top(\underline{\boldsymbol{x}} - \underline{\widetilde{\boldsymbol{x}}}) + \frac{1}{2}(\underline{\boldsymbol{x}} - \underline{\widetilde{\boldsymbol{x}}})^\top \nabla^2 g(\underline{\widetilde{\boldsymbol{x}}})(\underline{\boldsymbol{x}} - \underline{\widetilde{\boldsymbol{x}}}) + R(\underline{\boldsymbol{x}}, \underline{\widetilde{\boldsymbol{x}}})\right\} \\ &= \int\int \left[g(\underline{\widetilde{x}}) + \nabla g(\underline{\widetilde{x}})^\top(\underline{x} - \underline{\widetilde{x}}) + \frac{1}{2}(\underline{x} - \underline{\widetilde{x}})^\top \nabla^2 g(\underline{\widetilde{x}})(\underline{x} - \underline{\widetilde{x}}) + R(\underline{x}, \underline{\widetilde{x}})\right] p(\underline{x}, \underline{\widetilde{x}}) \, \mathrm{d}\underline{\widetilde{x}} \, \mathrm{d}\underline{x} \, , \end{aligned}$$

where R is the residuum and $p(\underline{x}, \underline{\widetilde{x}})$ is the joint distribution of $\underline{\boldsymbol{x}}$ and $\underline{\widetilde{\boldsymbol{x}}}$. Finally, ignoring the higher-order terms, assuming that $\underline{\boldsymbol{x}}$ and $\underline{\widetilde{\boldsymbol{x}}}$ are independent, and using that $p(\underline{\widetilde{x}})$ is the Dirac mixture $\bar{p}^x(\underline{x})$, we obtain

$$\begin{aligned} a(p^x(\underline{x})) &\approx \int\int \left[g(\underline{\widetilde{x}}) + \nabla g(\underline{\widetilde{x}})^\top(\underline{x} - \underline{\widetilde{x}}) + \frac{1}{2}(\underline{x} - \underline{\widetilde{x}})^\top \nabla^2 g(\underline{\widetilde{x}})(\underline{x} - \underline{\widetilde{x}})\right] p^x(\underline{x}) \bar{p}^x(\underline{\widetilde{x}}) \, \mathrm{d}\underline{\widetilde{x}} \, \mathrm{d}\underline{x} \\ &= \int \sum_{i=1}^{M} \alpha^i \left[g(\underline{x}^i) + \nabla g(\underline{x}^i)^\top(\underline{x} - \underline{x}^i) + \frac{1}{2}(\underline{x} - \underline{x}^i)^\top \nabla^2 g(\underline{x}^i)(\underline{x} - \underline{x}^i)\right] p^x(\underline{x}) \, \mathrm{d}\underline{x} \\ &= \mathrm{E}\left\{\sum_{i=1}^{M} \alpha^i \left[g(\underline{x}^i) + \nabla g(\underline{x}^i)^\top(\underline{\boldsymbol{x}} - \underline{x}^i) + \frac{1}{2}(\underline{\boldsymbol{x}} - \underline{x}^i)^\top \nabla^2 g(\underline{x}^i)(\underline{\boldsymbol{x}} - \underline{x}^i)\right]\right\} \, . \end{aligned}$$

With Lemma 2.1 and Definition 2.2, we are able to derive the proposed control approach.

2.5.2 Control Algorithm

In this section, we will derive an algorithm that can be applied in problems described in Sec. 2.2 to compute a policy. We will proceed as follows. First, we discuss how to generate reference trajectories, along which we will apply DP. Then, we present a scheme for value function approximation that relies on statistical Taylor expansion along a reference trajectory in Theorems 2.4 and 2.5. Finally, we formulate an algorithm that computes a control law for the problem from Sec. 2.2 that is affine in the mean of the state estimate.

The local DP approach presented later in this section performs statistical Taylor expansion along a reference trajectory that consists of a sequence of reference state estimates $\bar{p}^x_{0:K}(\underline{x}_{0:K})$ maintained in form of Dirac mixtures and a sequence of control inputs $\bar{\underline{u}}_{0:K-1}$. In order to be able to generate the trajectory $\{\bar{p}^x_{0:K}(\underline{x}_{0:K}), \bar{\underline{u}}_{0:K-1}\}$, we require an initial policy $\underline{\pi}_k(p^x_k(\underline{x}_k))$ for $k = 0, \ldots, K-1$. Then, the reference trajectory can be computed as presented in Algorithm 2.1.

In Algorithm 2.1, we used samples from $\bar{p}^x_k(\underline{x}_k)$, $p^w_k(\underline{w}_k)$, and $p^v_k(\underline{v}_k)$ for simulation. However, it is also possible to use the nominal trajectory of the current control policy. To this end, we simply use the means $\mathrm{E}\left\{\underline{\boldsymbol{x}}_k\right\}$, $\mathrm{E}\left\{\underline{\boldsymbol{w}}_k\right\}$, and $\mathrm{E}\left\{\underline{\boldsymbol{v}}_k\right\}$ in Algorithm 2.1.

The next two theorems give us a statistical second-order Taylor expansion of the value function along a reference trajectory. This approximation scheme induces a closed-loop control law that is affine in the mean of $\underline{\boldsymbol{x}}_k$. For the results in Theorems 2.4 and 2.5 to be well-defined, we need

Algorithm 2.1 Computation of a reference trajectory.

- *Step 1:* Initialize $\bar{p}_0^x(\underline{x}_0) = p_0^x(\underline{x}_0)$ and set $k = 0$.
- *Step 2:* Compute the reference control input $\bar{\underline{u}}_k = \underline{\pi}_k(\bar{p}_k^x(\underline{x}_k))$.
- *Step 3:* Compute a simulated system state $\underline{x}_{k+1}$ using (2.1) with a sample from the reference state estimate $\bar{p}_k^x(\underline{x}_k)$, the control input $\bar{\underline{u}}_k$, and a sample from the probability distribution $p_k^w(\underline{w}_k)$ of the process noise.
- *Step 4:* Use $\underline{x}_{k+1}$ and a sample from the probability distribution $p_{k+1}^v(\underline{v}_{k+1})$ of the measurement noise in order to compute a simulated measurement $\underline{y}_{k+1}$.
- *Step 5:* Compute the reference state estimate $\bar{p}_{k+1}^x(\underline{x}_{k+1})$ based on $\bar{p}_k^x(\underline{x}_k)$, $\bar{\underline{u}}_k$, and $\underline{y}_{k+1}$.
- *Step 6:* Stop the algorithm if $k = K$. Otherwise, set $k = k + 1$ and return to *Step 2*.

Assumption 2.3. This assumption will ensure that the local second-order Taylor approximation of the costs is positive definite and thus has a unique minimum.

Assumption 2.3. *The cost functions $\mathcal{C}_K(\underline{x}_K)$ and $\mathcal{C}_k(\underline{x}_k, \underline{u}_k)$, $k = 1, \ldots, K-1$ satisfy*

$$\begin{aligned} &\nabla_{xx}\mathcal{C}_k(\underline{x}_k, \underline{u}_k) \geq 0 \ , \quad \nabla_{uu}\mathcal{C}_k(\underline{x}_k, \underline{u}_k) > 0 \ , \\ &\begin{bmatrix} \nabla_{xx}\mathcal{C}_k(\underline{x}_k, \underline{u}_k) & \nabla_{xu}\mathcal{C}_k(\underline{x}_k, \underline{u}_k) \\ \nabla_{ux}\mathcal{C}_k(\underline{x}_k, \underline{u}_k) & \nabla_{uu}\mathcal{C}_k(\underline{x}_k, \underline{u}_k) \end{bmatrix} \geq 0 \ . \end{aligned} \tag{2.9}$$

Now, assume that we have a reference trajectory $\{\bar{p}_{0:K}^x(\underline{x}_{0:K}), \bar{\underline{u}}_{0:K-1}\}$. Then, the following result holds.

Theorem 2.4. *Given the Dirac mixture $\bar{p}_K^x(\underline{x}_K) = \sum_{i=1}^M \bar{\alpha}_K^i \delta(\underline{x}_K - \bar{\underline{x}}_K^i)$ of the reference trajectory $\{\bar{p}_{0:K}^x(\underline{x}_{0:K}), \bar{\underline{u}}_{0:K-1}\}$, a second-order approximation of the value function $V_K(p_K^x(\underline{x}_K))$ according to Definition 2.2 is given by*

$$V_K(p_K^x(\underline{x}_K)) \approx \mathrm{E}\left\{ s_K + \underline{\mathrm{v}}_K^\top \boldsymbol{\underline{x}}_K + \frac{1}{2}\boldsymbol{\underline{x}}_K^\top \mathbf{V}_K \boldsymbol{\underline{x}}_K + \frac{1}{2}\mathrm{tr}\left[\mathbf{S}_K \mathbf{C}_K\right] \right\} \ ,$$

where

$$\begin{aligned} s_K &= \sum_{i=1}^M \bar{\alpha}_K^i \left[\mathcal{C}_K(\bar{\underline{x}}_K^i) - \nabla\mathcal{C}_K(\bar{\underline{x}}_K^i)^\top \bar{\underline{x}}_K^i + \frac{1}{2}(\bar{\underline{x}}_K^i)^\top \nabla^2 \mathcal{C}_K(\bar{\underline{x}}_K^i) \bar{\underline{x}}_K^i \right] \ , \\ \underline{\mathrm{v}}_K &= \sum_{i=1}^M \bar{\alpha}_K^i \left[\nabla\mathcal{C}_K(\bar{\underline{x}}_K^i) - \nabla^2 \mathcal{C}_K(\bar{\underline{x}}_K^i) \bar{\underline{x}}_K^i \right] \ , \\ \mathbf{V}_K &= \sum_{i=1}^M \bar{\alpha}_K^i \nabla^2 \mathcal{C}_K(\bar{\underline{x}}_K^i) \ , \\ \mathbf{S}_K &= \mathbf{0}_{n_x} \ . \end{aligned}$$

Proof. The result of Theorem 2.4 follows directly from applying the approximation scheme proposed in Definition 2.2 to $V_K(p_K^x(\underline{x}_K))$. □

Next, we provide an expression for second-order approximation of the value function at time step $k = 0, \ldots, K-1$.

Theorem 2.5. *For a state estimate $\bar{p}_k^x(\underline{x}_k) = \sum_{i=1}^M \bar{\alpha}_k^i \delta(\underline{x}_k - \bar{\underline{x}}_k^i)$ that is part of the reference trajectory $\{\bar{p}_{0:K}^x(\underline{x}_{0:K}), \bar{\underline{u}}_{0:K-1}\}$, a second-order approximation of the value function $V_k(p_k^x(\underline{x}_k))$ according to Definition 2.2 for $k = 0, \ldots, K-1$ is given by*

$$V_k(p_k^x(\underline{x}_k)) \approx \mathrm{E}\left\{ s_k + \underline{\mathrm{v}}_k^\top \underline{\boldsymbol{x}}_k + \frac{1}{2}\underline{\boldsymbol{x}}_k^\top \mathbf{V}_k \underline{\boldsymbol{x}}_k + \frac{1}{2}\mathrm{tr}\left[\mathbf{S}_k \mathbf{C}_k\right] \right\} , \tag{2.10}$$

with

$$\begin{aligned}
s_k &= \left(\frac{\epsilon_k^2}{2} - \epsilon_k\right) \widetilde{\underline{r}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{r}}_k - \frac{1}{2}\widetilde{\underline{p}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{p}}_k + \widetilde{\underline{r}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{p}}_k \\
&\quad + \sum_{i=1}^M \bar{\alpha}_k^i \left[\varphi_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k) - (\underline{p}_k^i)^\top \bar{\underline{x}}_k^i + \frac{1}{2}(\bar{\underline{x}}_k^i)^\top \mathbf{Q}_k^i \bar{\underline{x}}_k^i \right] , \\
\underline{\mathrm{v}}_k &= \widetilde{\mathbf{P}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{p}}_k - \widetilde{\mathbf{P}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{r}}_k + \sum_{i=1}^M \bar{\alpha}_k^i \left[\underline{p}_k^i - \mathbf{Q}_k^i \bar{\underline{x}}_k^i \right] , \\
\mathbf{V}_k &= -\widetilde{\mathbf{P}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\mathbf{P}}_k + \sum_{i=1}^M \bar{\alpha}_k^i \mathbf{Q}_k^i , \\
\mathbf{S}_k &= \widetilde{\mathbf{P}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\mathbf{P}}_k ,
\end{aligned} \tag{2.11}$$

where $\epsilon_k \in (0, 1]$ is a parameter and

$$\begin{aligned}
&\begin{bmatrix} \underline{p}_k^i \\ \underline{r}_k^i \end{bmatrix} = \nabla \varphi_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k) , \qquad \begin{bmatrix} \mathbf{Q}_k^i & (\mathbf{P}_k^i)^\top \\ \mathbf{P}_k^i & \mathbf{R}_k^i \end{bmatrix} = \nabla^2 \varphi_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k) , \\
&\widetilde{\underline{p}}_k = \sum_{i=1}^M \bar{\alpha}_k^i \mathbf{P}_k^i \bar{\underline{x}}_k^i , \quad \widetilde{\underline{r}}_k = \sum_{i=1}^M \bar{\alpha}_k^i \underline{r}_k^i , \quad \widetilde{\mathbf{P}}_k = \sum_{i=1}^M \bar{\alpha}_k^i \mathbf{P}_k^i , \quad \widetilde{\mathbf{R}}_k = \sum_{i=1}^M \bar{\alpha}_k^i \mathbf{R}_k^i ,
\end{aligned} \tag{2.12}$$

with

$$\begin{aligned}
\varphi_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k) = \sum_{m=1}^R \bar{\beta}_k^m \Big[&\mathcal{C}_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k) + s_{k+1} + \underline{\mathrm{v}}_{k+1}^\top \underline{a}_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k, \bar{\underline{w}}_k^m) \\
&+ \frac{1}{2}\underline{a}_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k, \bar{\underline{w}}_k^m)^\top \mathbf{V}_{k+1} \underline{a}_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k, \bar{\underline{w}}_k^m) + \frac{1}{2}\mathrm{tr}\left[\mathbf{S}_{k+1}\mathbf{C}_{k+1}\right] \Big] ,
\end{aligned} \tag{2.13}$$

where R is the number of samples $\bar{\underline{w}}_k^m$, $m = 1, \ldots, R$, from the probability distribution $p_k^w(\underline{w})$ of the process noise.

Proof. The result of the above theorem can be shown by induction. Assume that (2.10) holds for $V_{k+1}(p^x_{k+1}(\underline{x}_{k+1}))$, $k = 0, \ldots, K-1$. Then, at time step k, we have

$$\begin{aligned} V_k(p^x_k(\underline{x}_k)) &= \inf_{\underline{u}_k} \mathrm{E}\left\{\mathcal{C}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k) + V_{k+1}(\underline{\boldsymbol{x}}_{k+1})\right\} \\ &\approx \inf_{\underline{u}_k} \mathrm{E}\left\{\mathcal{C}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k) + s_{k+1} + \underline{\mathrm{v}}^\top_{k+1}\underline{\boldsymbol{x}}_{k+1} + \frac{1}{2}\underline{\boldsymbol{x}}^\top_{k+1}\mathbf{V}_{k+1}\underline{\boldsymbol{x}}_{k+1} + \frac{1}{2}\mathrm{tr}\left[\mathbf{S}_{k+1}\mathbf{C}_{k+1}\right]\right\} \\ &= \inf_{\underline{u}_k} \mathrm{E}\Big\{\mathcal{C}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k) + s_{k+1} + \underline{\mathrm{v}}^\top_{k+1}\underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k, \underline{\boldsymbol{w}}_k) \\ &\qquad + \frac{1}{2}\underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k, \underline{\boldsymbol{w}}_k)^\top\mathbf{V}_{k+1}\underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k, \underline{\boldsymbol{w}}_k) + \frac{1}{2}\mathrm{tr}\left[\mathbf{S}_{k+1}\mathbf{C}^p_{k+1}\right]\Big\} \ , \end{aligned} \tag{2.14}$$

where we exploited the linearity of the expectation and the fact that s_{k+1}, $\underline{\mathrm{v}}_{k+1}$, $\mathbf{V}_{k+1}$, and $\mathbf{S}_{k+1}$ are independent of $\underline{\boldsymbol{x}}_k$, $\underline{u}_k$, and $\underline{\boldsymbol{w}}_k$. In (2.14), we write $\mathbf{C}^p_{k+1}$ to indicate that the covariance is computed using $p^x_k(\underline{\boldsymbol{x}}_k)$ instead of $p^x_{k+1}(\underline{x}_{k+1})$ as in $V_{k+1}(p^x_{k+1}(\underline{x}_{k+1}))$. If the probability distribution $p^x_k(\underline{x}_k)$ is given in form of a Dirac mixture, we can use Lemma 2.1 in order to obtain $\mathbf{C}^p_{k+1}$. Next, we perform second-order Taylor expansion of $V_k(p^x_k(\underline{x}_k))$ at $\bar{p}^x_k(\underline{x}_k) = \sum_{i=1}^M \bar{\underline{x}}^i_k \delta(\underline{x}_k - \bar{\underline{x}}^i_k)$ according to Definition 2.2. It holds

$$\begin{aligned} &V_k(p^x_k(\underline{x}_k)) \\ &\quad\approx \inf_{\underline{u}_k} \mathrm{E}\left\{\sum_{i=1}^M \bar{\underline{x}}^i_k \left[\varphi_k(\bar{\underline{x}}^i_k, \bar{\underline{u}}_k) + \begin{bmatrix}\underline{p}^i_k \\ \underline{r}^i_k\end{bmatrix}^\top \begin{bmatrix}\underline{\boldsymbol{x}}_k - \bar{\underline{x}}^i_k \\ \underline{u}_k - \bar{\underline{u}}_k\end{bmatrix} + \frac{1}{2}\begin{bmatrix}\underline{\boldsymbol{x}}_k - \bar{\underline{x}}^i_k \\ \underline{u}_k - \bar{\underline{u}}_k\end{bmatrix}^\top \begin{bmatrix}\mathbf{Q}^i_k & (\mathbf{P}^i_k)^\top \\ \mathbf{P}^i_k & \mathbf{R}^i_k\end{bmatrix}\begin{bmatrix}\underline{\boldsymbol{x}}_k - \bar{\underline{x}}^i_k \\ \underline{u}_k - \bar{\underline{u}}_k\end{bmatrix}\right]\right\} \\ &\quad= \inf_{\underline{u}_k} \mathrm{E}\Bigg\{\sum_{i=1}^M \bar{\underline{x}}^i_k \Big[\varphi_k(\bar{\underline{x}}^i_k, \bar{\underline{u}}_k) + (\underline{p}^i_k)^\top\underline{\boldsymbol{x}}_k - (\underline{p}^i_k)^\top\bar{\underline{x}}^i_k + (\underline{r}^i_k)^\top\Delta\underline{u}_k + \frac{1}{2}\underline{\boldsymbol{x}}^\top_k\mathbf{Q}^i_k\underline{\boldsymbol{x}}_k - \underline{\boldsymbol{x}}^\top_k\mathbf{Q}^i_k\bar{\underline{x}}^i_k \\ &\qquad + \frac{1}{2}(\bar{\underline{x}}^i_k)^\top\mathbf{Q}^i_k\bar{\underline{x}}^i_k + \underline{\boldsymbol{x}}^\top_k(\mathbf{P}^i_k)^\top\Delta\underline{u}_k - (\bar{\underline{x}}^i_k)^\top(\mathbf{P}^i_k)^\top\Delta\underline{u}_k + \Delta\underline{u}^\top_k\mathbf{R}^i_k\Delta\underline{u}_k\Big]\Bigg\} \ , \end{aligned} \tag{2.15}$$

where we used the definitions (2.12) and (2.13), and substituted $\Delta\underline{u}_k = \underline{u}_k - \bar{\underline{u}}_k$. Evaluation of the necessary optimality condition

$$\frac{\partial V_k(p^x_k(\underline{x}_k))}{\partial \underline{u}_k} = \underline{0}$$

yields

$$\underline{u}_k = -\widetilde{\mathbf{R}}^{-1}_k\widetilde{\mathbf{P}}_k\mathrm{E}\left\{\underline{\boldsymbol{x}}_k\right\} + \widetilde{\mathbf{R}}^{-1}_k\widetilde{\underline{p}}_k - \widetilde{\mathbf{R}}^{-1}_k\widetilde{\underline{r}}_k + \bar{\underline{u}}_k \ . \tag{2.16}$$

Following the argumentation in [159], at this point, we introduce the parameter $\epsilon_k \in (0, 1]$ such that we obtain

$$\begin{aligned} \underline{u}_k &= -\widetilde{\mathbf{R}}^{-1}_k\widetilde{\mathbf{P}}_k\mathrm{E}\left\{\underline{\boldsymbol{x}}_k\right\} + \widetilde{\mathbf{R}}^{-1}_k\widetilde{\underline{p}}_k - \epsilon_k\widetilde{\mathbf{R}}^{-1}_k\widetilde{\underline{r}}_k + \bar{\underline{u}}_k & (2.17) \\ &= \mathbf{L}_k\mathrm{E}\left\{\underline{\boldsymbol{x}}_k\right\} + \underline{d}_k \ . & (2.18) \end{aligned}$$

Later, this parameter will be used in order to ensure the convergence of the proposed algorithm for computation of the policy parameters $\mathbf{L}_k$ and $\underline{d}_k$ for $k = 0, \ldots, K-1$. Plugging (2.18) into (2.15) yields

$$V_k(p_k^x(\underline{x}_k)) \approx \mathrm{E}\left\{ \sum_{i=1}^{M} \bar{\underline{x}}_k^i \left[\varphi_k(\bar{\underline{x}}_k^i, \bar{\underline{u}}_k) - (\underline{p}_k^i)^\top \bar{\underline{x}}_k^i + \frac{1}{2} \underline{\boldsymbol{x}}_k^\top \mathbf{Q}_k^i \underline{\boldsymbol{x}}_k - \underline{\boldsymbol{x}}_k^\top \widetilde{\underline{p}}_k - \underline{\boldsymbol{x}}_k^\top \mathbf{Q}_k^i \bar{\underline{x}}_k^i - \underline{\boldsymbol{x}}_k^\top \widetilde{\mathbf{P}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{r}}_k \right.\right.$$
$$+ \underline{\boldsymbol{x}}_k^\top \widetilde{\mathbf{P}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{p}}_k + \left(\frac{\epsilon_k^2}{2} - \epsilon_k \right) \widetilde{\underline{r}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{r}}_k - \frac{1}{2} \widetilde{\underline{p}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{p}}_k + \widetilde{\underline{r}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\underline{p}}_k$$
$$\left.\left. + \mathrm{E}\left\{ \underline{\boldsymbol{x}}_k \right\}^\top \widetilde{\mathbf{P}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\mathbf{P}}_k \mathrm{E}\left\{ \underline{\boldsymbol{x}}_k \right\} \right] \right\} .$$

Please note that the term $\mathrm{E}\left\{ \underline{\boldsymbol{x}}_k \right\}^\top \widetilde{\mathbf{P}}_k^\top \widetilde{\mathbf{R}}_k^{-1} \widetilde{\mathbf{P}}_k \mathrm{E}\left\{ \underline{\boldsymbol{x}}_k \right\}$ is nonlinear. Therefore, it depends on the measurement $\underline{\boldsymbol{y}}_k$, which causes problems when trying to calculate $V_{k-1}(p_{k-1}^x(\underline{x}_{k-1}))$. For this reason, we apply the identity $\mathrm{E}\left\{ \underline{\boldsymbol{x}} \mathbf{A} \underline{\boldsymbol{x}} \right\} = \mathrm{tr}\left[\mathbf{A} \mathrm{E}\left\{ (\underline{\boldsymbol{x}} - \mathrm{E}\left\{ \underline{\boldsymbol{x}} \right\})(\underline{\boldsymbol{x}} - \mathrm{E}\left\{ \underline{\boldsymbol{x}} \right\})^\top \right\} \right] + \mathrm{E}\left\{ \underline{\boldsymbol{x}}_k^\top \right\} \mathbf{A} \mathrm{E}\left\{ \underline{\boldsymbol{x}} \right\}$ for deterministic matrices $\mathbf{A}$, which yields (2.10). The big advantage of the latter step is that the covariance $\mathbf{C}_k = \mathrm{E}\left\{ (\underline{\boldsymbol{x}}_k - \mathrm{E}\left\{ \underline{\boldsymbol{x}}_k \right\})(\underline{\boldsymbol{x}}_k - \mathrm{E}\left\{ \underline{\boldsymbol{x}}_k \right\})^\top \right\}$ is independent of the measurement $\underline{\boldsymbol{y}}_k$, if we assume that the state and the measurement are jointly Gaussian. □

Please note that matrix $\widetilde{\mathbf{R}}_k$ introduced in Theorem 2.5 is invertible because it is a convex combination of matrices $\mathbf{R}_k^i$ that are positive definite, and thus invertible, according to Assumption 2.9. Therefore, the optimization of the approximation of $V_k(p_k^x(\underline{x}_k))$ in (2.15) is well-posed, i.e., it has a unique minimum that is attained for (2.16).

Equation (2.10) in Theorem 2.5 resembles the equation for the costs-to-go in LQG control due to the second-order approximation of the value function. However, in LQG, the covariance of the state estimate is independent of the control inputs. Thus, the term $\mathrm{tr}\left[\mathbf{S}_k \mathbf{C}_k \right]$ can be omitted from optimization, which eliminates the dual effect. In the considered nonlinear case, however, the covariance $\mathbf{C}_k$ depends on the control input $\underline{u}_{k-1}$ and cannot be ignored if the controller has to be closed-loop.

With the results from Theorems 2.4 and 2.5, we are now able to formulate the algorithm for computation of the parameters $\mathbf{L}_k$ and $\underline{d}_k$ for $k = 0, \ldots, K-1$ of the affine policy induced by the statistical second-order Taylor expansion of the value function $V_k(p_k^x(\underline{x}_k))$. The proposed algorithm is given in Algorithm 2.2. This algorithm requires an initial policy that without loss of generality can be $\mathbf{L}_k = \mathbf{0}$ and $\underline{d}_k = \underline{0}$ for $k = 0, \ldots, K-1$. However, it is possible to use rapidly-exploring random belief trees [26] to generate a good initial policy as proposed in [159]. This approach has two advantages. First, the convergence to a local minimum of the value function is faster. And second, the chance to find the global minimum increases.

The backtracking line search in *Step 5* of Algorithm 2.2 ensures that the algorithm converges to a local minimum of the costs $\mathcal{J}$. The basic algorithm for this step is proposed in Algorithm 2.3. In this algorithm, the current policy is considered locally optimal if the costs-to-go associated with this policy are smaller than the costs associated with the policy (2.18) for $\epsilon_k < \alpha$, where α is a small positive constant. If the current policy is locally optimal at the time step k, we do not change it. Otherwise, the policy is updated. Please note that Algorithm 2.3 is only the

Algorithm 2.2 Computation of the controller parameters.

- *Step 1:* Choose an initial policy $\{\mathbf{L}_{0K-1}^{[0]}, \underline{d}_{0:K-1}^{[0]}\}$ and set $\eta = 0$.
- *Step 2:* Use the current policy $\{\mathbf{L}_{0K-1}^{[\eta]}, \underline{d}_{0:K-1}^{[\eta]}\}$ to generate a reference trajectory $\{\bar{p}_{0:K}^{x}(\underline{x}_{0:K}), \bar{\underline{u}}_{0:K-1}\}$ according to Algorithm 2.1.
- *Step 3:* Compute an approximation of the value function $V_K(p_K^x(\underline{x}_K))$ according to Theorem 2.4 and set $k = K - 1$.
- *Step 4:* Compute $\widetilde{\underline{p}}_k$, $\widetilde{\underline{r}}_k$, $\widetilde{\mathbf{P}}_k$, and $\widetilde{\mathbf{R}}_k$ using (2.12).
- *Step 5:* Perform backtracking line search in order to determine the parameter ϵ_k.
- *Step 6:* Store the policy for the current time step, i.e., set $\mathbf{L}_k^{[\eta+1]} = -\widetilde{\mathbf{R}}_k^{-1}\widetilde{\mathbf{P}}_k$ and $\underline{d}_k^{[\eta+1]} = \widetilde{\mathbf{R}}_k^{-1}\widetilde{\underline{p}}_k - \epsilon_k\widetilde{\mathbf{R}}_k^{-1}\widetilde{\underline{r}}_k + \bar{\underline{u}}_k^{[\eta]}$.
- *Step 7:* Compute s_k, $\underline{\mathrm{v}}_k$, $\mathbf{V}_k$, and $\mathbf{S}_k$ using (2.11).
- *Step 8:* If $k = 0$, go to *Step 9*. Otherwise, return to *Step 4*.
- *Step 9:* Stop the algorithm if the costs converged, i.e., if the absolute difference between costs $\mathcal{J}^{[\eta+1]}$ of the current iteration and the costs $\mathcal{J}^{[\eta]}$ of the previous algorithm iteration is smaller than a predefined constant. Otherwise, return to *Step 2*.

most basic form of admissible line search and much more sophisticated line search algorithms can be applied in *Step 5* of Algorithm 2.2.

Algorithm 2.3 Basic line search.

- *Step 1:* Set $\epsilon_k = 1$.
- *Step 2:* For the reference state estimate $\bar{p}_k^x(\underline{x}_k)$, compute $\underline{u}_k = -\widetilde{\mathbf{R}}_k^{-1}\widetilde{\mathbf{P}}_k\mathrm{E}\{\underline{\boldsymbol{x}}_k\} + \widetilde{\mathbf{R}}_k^{-1}\widetilde{\underline{p}}_k - \epsilon_k\widetilde{\mathbf{R}}_k^{-1}\widetilde{\underline{r}}_k + \bar{\underline{u}}_k$.
- *Step 3:* Compute the costs-to-go $V_k(\bar{p}_k^x(\underline{x}_k))$ for $\underline{u}_k = \bar{\underline{u}}_k$ and $V_k^\epsilon(\bar{p}_k^x(\underline{x}_k))$ for $\underline{u}_k$ from *Step 2*.
- *Step 4:* If $V_k^\epsilon(\bar{p}_k^x(\underline{x}_k)) \leq V_k(\bar{p}_k^x(\underline{x}_k))$ stop the algorithm an return ϵ_k. Otherwise, set $\epsilon_k = \epsilon_k/2$ and return to *Step 2*.

We want to conclude this section with a brief remark on how Algorithm 2.2 can be improved. Please observe that in Algorithm 2.2, we compute the policy by optimizing the value function along a single reference trajectory that is generated using a single sample trajectory from the process and the measurement noises, and a single sample from the initial state estimate $p_0^x(\underline{x}_0)$. However, if we generate several sample trajectories, we can construct several trajectories of state estimates for the current policy and then optimize the policy for the entire bundle of reference trajectories. By doing so, the value approximation quality can be increased, which then leads to a better policy.

2.5.3 Numerical Example

In this section, we demonstrate the proposed control approach in a scenario with state-dependent measurement noise that is similar to the one depicted in Fig. 2.1. We consider simple 2D time-invariant dynamics

$$\underline{\boldsymbol{x}}_{k+1} = \underline{\boldsymbol{x}}_k + \underline{u}_k + 0.5\underline{\boldsymbol{w}}_k \ ,$$

where the process noise $\underline{\boldsymbol{w}}_k$ is zero-mean and has identity covariance, and the nonlinear measurement equation

$$\underline{\boldsymbol{y}}_k = \|\underline{\boldsymbol{x}}_k - \underline{\hat{x}}\| + \underline{\boldsymbol{v}}_k \ ,$$

where $\underline{\hat{x}}$ denotes the position of the landmark and $\underline{\boldsymbol{v}}_k \in \mathbb{R}$ is a Gaussian zero-mean noise with state-dependent covariance $0.1^2 \cdot \|\underline{\boldsymbol{x}}_k - \underline{\hat{x}}\|^2$. In the considered scenario, we set the position of the landmark to $\underline{\hat{x}} = [-3 \ \ 0]^\top$.

The initial system state is assumed to be Gaussian with mean $\mathrm{E}\{\underline{\boldsymbol{x}}_0\} = [-5 \ \ 3]^\top$ and covariance $\mathrm{E}\left\{(\underline{\boldsymbol{x}}_0 - \mathrm{E}\{\underline{\boldsymbol{x}}_0\})(\underline{\boldsymbol{x}}_0 - \mathrm{E}\{\underline{\boldsymbol{x}}_0\})^\top\right\} = 0.2^2 \cdot \mathbf{I}_2$, and the target state is $\underline{\widetilde{x}}_T = [0 \ \ 3]^\top$. The cost function is chosen to

$$\mathcal{J} = \mathrm{E}\left\{(\underline{\boldsymbol{x}}_K - \underline{\widetilde{x}}_T)^\top(\underline{\boldsymbol{x}}_K - \underline{\widetilde{x}}_T) + \sum_{k=0}^{K-1}(\underline{\boldsymbol{x}}_k - \underline{\widetilde{x}}_T)^\top(\underline{\boldsymbol{x}}_k - \underline{\widetilde{x}}_T) + 0.1\underline{u}_k^\top\underline{u}_k\right\} \tag{2.19}$$

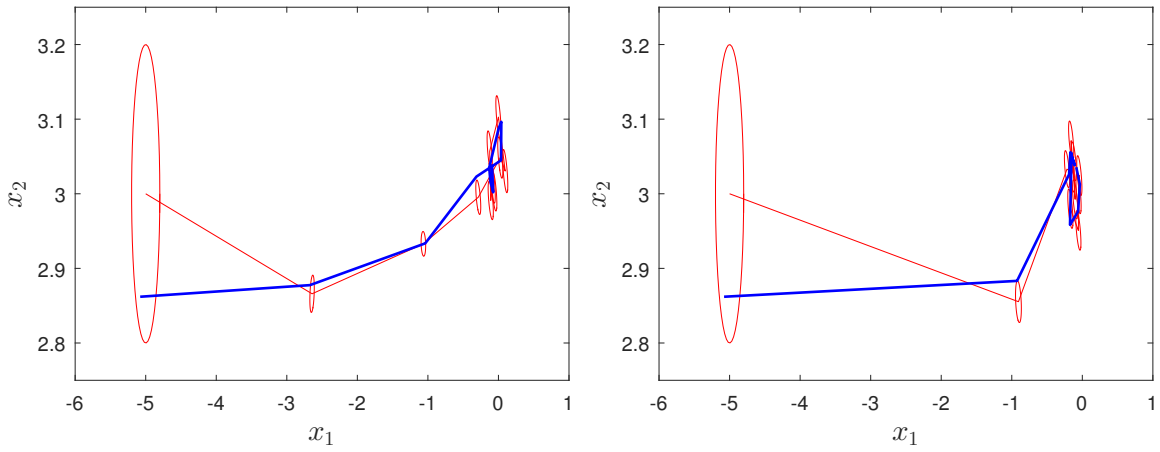
with horizon length $K = 20$. As the estimator, we employ the UKF from the Nonlinear Estimation Toolbox [144]. To compute the derivatives in (2.12), we use the algorithm for adaptive robust numerical differentiation that is available at [46].

The proposed approach is compared with the trajectory optimization approach from [159]. In this approach, the state estimates are expressed with Gaussians that are calculated using an EKF and the control law is assumed affine in the mean and the vectorized square root of the covariance. These assumptions allow a reformulation of the closed-loop system dynamics in terms of the Gaussian state. Finally, the value function is approximated using a second-order Taylor expansion at the Gaussian states of a reference trajectory. Because the approach from [159] requires the cost function to be explicit in the mean and the variance of the Gaussian state, we apply the identity $\underline{x}^\top\mathbf{A}\underline{x} = \mathrm{tr}\left[\mathbf{A}\underline{x}\underline{x}^\top\right]$ and use the property $\mathrm{E}\left\{\underline{x}\underline{x}^\top\right\} = \mathrm{E}\left\{\underline{x}\right\}\mathrm{E}\left\{\underline{x}^\top\right\} + \mathrm{E}\left\{(\underline{x} - \mathrm{E}\left\{\underline{x}\right\})\right\}(\underline{x} - \mathrm{E}\left\{\underline{x}\right\})^\top$ in order to rewrite the cost function (2.19) as

$$\mathcal{J} = \underline{\mu}_K^\top\underline{\mu}_K + \mathbf{C}_K - 2\underline{\widetilde{x}}_T^\top\underline{\mu}_K + K\underline{\widetilde{x}}_T^\top\underline{\widetilde{x}}_T + \sum_{k=0}^{K-1}\underline{\mu}_k^\top\underline{\mu}_k + \mathbf{C}_k - 2\underline{\widetilde{x}}_T^\top\underline{\mu}_k + 0.1\underline{u}_k\underline{u}_k^\top \ , \tag{2.20}$$

where $\underline{\mu}_k$ and $\mathbf{C}_k$ are the mean and the covariance of the Gaussian state estimate $p_k^x(\underline{x}_k)$ at time step k. Please note that in contrast to the approach from [159], the proposed approach does not require the cost function to be formulated in terms of the Gaussian state estimates.

In Fig. 2.3, we can see the trajectories of a simulation run that were generated by the proposed approach and the approach from [159]. In particular, the trajectory of the true system state and



(a) Proposed.

(b) Van den Berg et al [159].

Figure 2.3: Trajectories of the true system state (blue) and the Gaussian state estimate maintained by the controller (red, covariances are depicted as ellipses) of a simulation run.

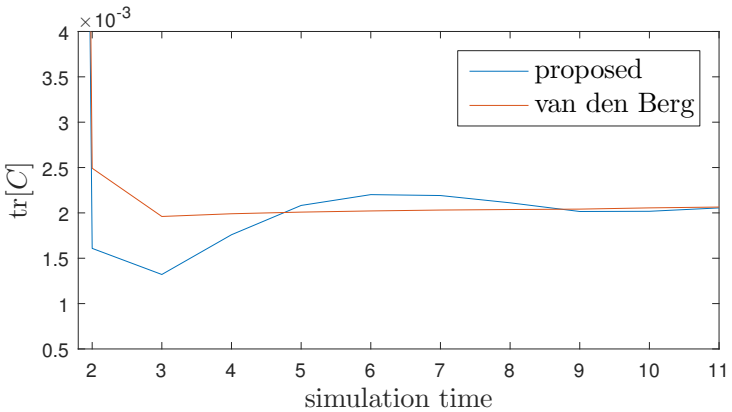


Figure 2.4: Trace of the covariances of the Gaussian state estimate maintained by the proposed controller and the controller from [159] during simulation.

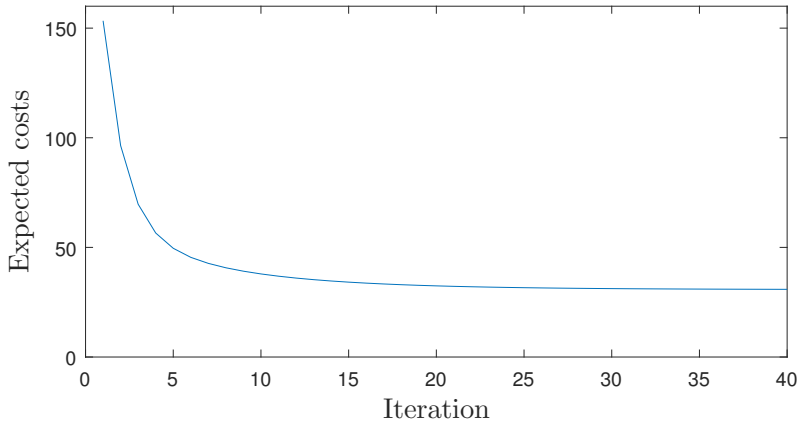


Figure 2.5: Expected costs $\mathcal{J}$ of the proposed controller over the number of iterations.

the Gaussian state estimate that is maintained by the controller are depicted. It can be seen that the trajectories of the both controllers deviate towards the landmark from the straight line that connects the initial mean of the state estimate which serves as the starting point and the goal state $\widetilde{\underline{x}}_T$. This deviation increases the actuation-related costs. However, the costs related to the quality of state estimation are simultaneously decreased due to a better localization near the landmark. Please note that this effect is explicit for the controller from [159], because the costs that are related to the localization quality (the covariance of the state estimate) explicitly appear in the cost function (2.20). On the other hand, the deviation towards the landmark performed by the proposed controller captures the trade-off between the actuation-related and the localization-related costs implicitly. The improvement of the localization quality can be seen in Fig. 2.4 that shows the traces of the covariances of the Gaussian state estimates maintained by the two compared controllers. It can be seen that the proposed approach achieves a smaller covariance in time steps 2–4. The covariance then increases, because the true state moves away from the landmark towards the goal state. Nevertheless, the proposed controller achieves a better localization quality with minimal covariance trace of $1.343 \cdot 10^{-3}$ compared to the cumulated covariance trace of $1.9924 \cdot 10^{-3}$ of the controller from [159].

Finally, Fig. 2.5 depicts the expected costs $\mathcal{J}$ of the proposed controller over the number of iterations of Algorithm 2.2. It can be seen that the cumulative costs converge in the considered scenario after approximately 20 iterations. This convergence is achieved via back stepping line search and indicates that the proposed controller converges to a local minimum. Moreover, the line search guarantees that local convergence is always ensured. However, the quality of the local minimum strongly depends on the initial reference trajectory. Thus, by using global search methods for the initial reference trajectory such as RRT* as proposed in [159] increases the chance of attaining at least a good local minimum.

2.6 Non-parametric Approximation of Functions on Probability Distributions

In the previous section, we presented a stochastic controller with a linear affine policy that is computed using trajectory optimization with local statistical Taylor approximation of the value function along a reference trajectory that is maintained in form of Dirac mixtures. This approximation method induces a linear controller that can be too limiting in some applications. Thus, we propose an alternative method for approximation of the value function along a reference trajectory and approximation of the control law using Gaussian Processes (GPs). Classical GPs are defined for deterministic inputs. However, both the value function and the control law are functions of the state estimate, which is a probability distribution. Thus, we propose a novel framework for GPs defined over general probability distributions in this section and evaluate it in a numerical example.

2.6.1 Gaussian Process Regression with Deterministic Inputs

In regression, the task consists in making predictions of function values at test inputs given a data set $(\mathbf{X}, \underline{y})$ that consists of training inputs $\mathbf{X} = \{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_M\}$, $\underline{x}_i \in \mathbb{R}^{n_x}$, $n_x \in \mathbb{N}$, $i = 1, 2, \ldots, M$, $M \in \mathbb{N}$ and the corresponding outputs $\underline{y} = \{y_1, y_2, \ldots, y_M\}$, $y_i \in \mathbb{R}$, $i = 1, 2, \ldots, M$. To this end, we usually assume a function model and estimate its parameters. Here, we particularly consider the model

$$y = a(\underline{x}) + \nu \ ,$$

where $a : \mathbb{R}^{n_x} \to \mathbb{R}$ is a nonlinear function and $\nu \in \mathbb{R}$ denotes an independent and identically distributed Gaussian measurement noise with zero mean and variance σ_ν^2.

In parametric regression, there is only one parameter set to be estimated. However, in the non-parametric regression, we assume that there are infinitely many parameter sets equipped with a probability distribution and we estimate the parameters of this distribution [124][5]. The parameters of the distribution of model parameters are referred to as the *hyperparameters*.

A special class of non-parametric regression methods is the GP, where we assume that the outputs y_i have a joint Gaussian distribution. More generally, a GP is defined as follows [134].

Definition 2.6. *A Gaussian process is a collection of random variables, where any finite subset has a joint Gaussian distribution.*

Under this definition, a GP on the space of vector-valued inputs $\underline{x} \in \mathbb{R}^{n_x}$ is determined by a mean function m and a covariance function k, whose parameters $\underline{\boldsymbol{w}}$ are distributed according to a Gaussian with hyperparameters $\underline{\theta}$. To emphasize that an output y is generated by a GP with hyperparameters $\underline{\theta}$, we write

$$y \sim \mathcal{GP}_{\underline{\theta}}(m, k) \ .$$

[5] Often, the term 'non-parametric' is also used when the number of parameter sets grows with data.

In spatial statistics, GP regression is referred to as kriging [47].

Given the training data, we can compute the hyperparameters, here the mean and the covariance of the probability distribution of the (random) parameter vector $\underline{\boldsymbol{w}}$, according to

$$p(\underline{w}|\mathbf{X}, \underline{y}) = \frac{p(\underline{y}|\mathbf{X}, \underline{w})p(\underline{w})}{p(\underline{y}|\mathbf{X})} \ ,$$

where $p(\underline{w})$ is an assumed prior of the parameters, which is parametrized with prior hyperparameters, and $p(\underline{y}|\mathbf{X})$ is referred to as the marginal likelihood. The marginal likelihood is independent of the meters $\underline{\boldsymbol{w}}$ and is only required for normalization of $p(\underline{w}|\mathbf{X}, \underline{y})$. If no prior $p(\underline{w})$ is available, we can compute a point estimate of the parameters $\underline{\boldsymbol{w}}$ by maximizing the likelihood $p(\underline{y}|\mathbf{X}, \underline{w})$, or better, its logarithm.

Having obtained the hyperparameters, we can predict the probability distribution $p(y_*|\underline{x}_*, \mathbf{X}, \underline{y})$ of the output $\boldsymbol{y}_*$ at a test input $\underline{x}_*$ according to

$$p(y_*|\underline{x}_*, \mathbf{X}, \underline{y}) = \int\limits_{\mathbb{R}^{n_w}} p(y_*|\underline{x}_*, \underline{w})p(\underline{w}|\mathbf{X}, \underline{y}) \ \mathrm{d}\underline{w} \ .$$

By exploiting the Gaussianity of the output, the parameters of the distribution $p(y_*|\underline{x}_*, \mathbf{X}, \underline{y})$ can be calculated using the next proposition [19]. To this end, for a given set of training inputs $\mathbf{X}$, we introduce the function $\underline{m}(\mathbf{X})$ and the kernel $\boldsymbol{\Sigma}(\mathbf{X})$ with

$$\underline{m}(\mathbf{X}) = \begin{bmatrix} m(\underline{x}_1) \\ m(\underline{x}_2) \\ \vdots \\ m(\underline{x}_M) \end{bmatrix} \ , \qquad \boldsymbol{\Sigma}(\mathbf{X}) = \begin{bmatrix} k(\underline{x}_1, \underline{x}_1) & k(\underline{x}_1, \underline{x}_2) & \cdots & k(\underline{x}_1, \underline{x}_M) \\ k(\underline{x}_2, \underline{x}_1) & k(\underline{x}_2, \underline{x}_2) & \cdots & k(\underline{x}_2, \underline{x}_M) \\ \vdots & \vdots & \ddots & \vdots \\ k(\underline{x}_M, \underline{x}_1) & k(\underline{x}_M, \underline{x}_2) & \cdots & k(\underline{x}_M, \underline{x}_M) \end{bmatrix} \ .$$

Please recall that both $\underline{m}(\mathbf{X})$ and $\boldsymbol{\Sigma}(\mathbf{X})$ are parametrized with $\underline{\boldsymbol{w}}$.

Proposition 2.7. *The Gaussian distribution $p(y_*|\underline{x}_*, \mathbf{X}, \underline{y})$ of the output $\boldsymbol{y}_*$ is determined by its mean μ_* and its variance σ_*^2 with*

$$\mu_* = m(\underline{x}_*) + k(\underline{x}_*, \mathbf{X})^\top \boldsymbol{\Sigma}_{\underline{y}}^{-1}(\underline{y} - \underline{m}(\mathbf{X})) \ ,$$
$$\sigma_*^2 = k(\underline{x}_*, \underline{x}_*) - k(\underline{x}_*, \mathbf{X})^\top \boldsymbol{\Sigma}_{\underline{y}}^{-1} k(\underline{x}_*, \mathbf{X}) \ ,$$

where $\boldsymbol{\Sigma}_{\underline{y}} = \boldsymbol{\Sigma}(\mathbf{X}) + \sigma_n^2 \mathbf{I}_M$ and

$$k(\underline{x}_*, \mathbf{X}) = \begin{bmatrix} k(\underline{x}_*, \underline{x}_1) & k(\underline{x}_*, \underline{x}_2) & \cdots & k(\underline{x}_*, \underline{x}_M) \end{bmatrix}^\top \ .$$

An interesting property of GPs is that they can be seen as a distribution over functions. This can be visualized by sampling from the probability distribution of the parameters and plotting the functions of these parameters. An example is depicted in Fig. 2.6, where we plot 20 functions drawn from a GP with training data $\mathbf{X} = \{-1, -2, 2, 2.4\}$ and $\underline{y} = \{-0.9478, -1.6054, 1.3895, 1.1547\}$. The output data was generated according to $y = 0.25x +$

$\sin(x)+\nu$, where ν is a zero-mean Gaussan noise with variance 0.1^2. To generate the 20 function samples from the GP, we sampled 20 parameter sets from the Gaussian distribution with mean $\underline{\mu} = [1.2401 \quad 5.8161]^\top$ and covariance

$$\mathbf{C} = \begin{bmatrix} 0.4724 & 0.3511 \\ 0.3511 & 0.4686 \end{bmatrix} .$$

It can be seen that the shapes of these functions resemble the true function.

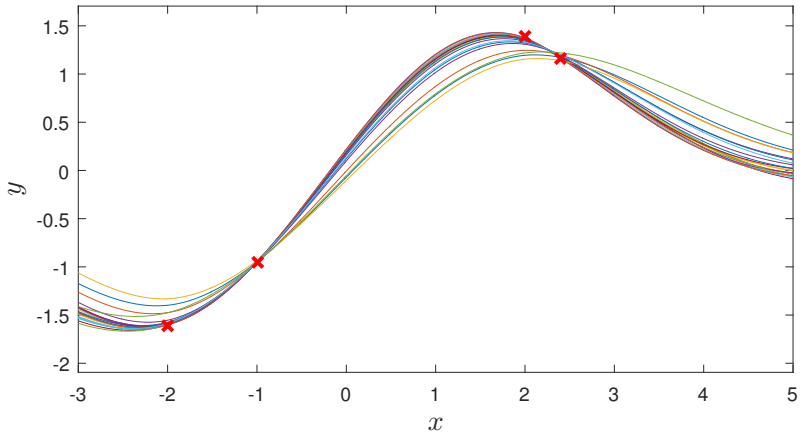


Figure 2.6: Depicted are 20 function samples from a GP. The training data consists of the points indicated with red crosses.

A more detailed introduction to GP regression can be found in the monograph [134].

2.6.2 Related Work

As pointed out above, we propose to use GPs as function approximators for the value function and the policy. These functions map the state estimates to the scalar-valued costs or to the vector-valued control inputs, respectively. Thus, we need a function approximator for inputs provided in form of probability distributions and therefore require GPs defined over the space of general probability distributions.

The amount of work on GPs over probability distributions available in literature is sparse. For example, Girard et al. consider GP regression with deterministic vector-valued training inputs $\mathbf{X}$ and Gaussian test inputs in [73]. In this work, they show that the distribution of $\boldsymbol{y}_*$ given the input distribution $d_* = p_*(\underline{x})$ can be computed according to

$$p(y_*|d_*, \mathbf{X}, \underline{y}) = \int_{\mathbb{R}^{n_x}} p(y_*|\underline{x}_*, \mathbf{X}, \underline{y}) p_*(\underline{x}_*) \, \mathrm{d}\underline{x}_* \, , \tag{2.21}$$

where $p(y_*|\underline{x}_*, \mathbf{X}, \underline{y})$ is determined for a particular deterministic $\underline{x}_*$ following the classical GP framework. Integral (2.21) can be solved analytically if the covariance function is the Squared Exponential (SE) function

$$k(\underline{x}_i, \underline{x}_j) = \alpha^2 \exp\left(-\frac{1}{2}(\underline{x}_i - \underline{x}_j)^\top \mathbf{\Lambda}^{-1}(\underline{x}_i - \underline{x}_j)\right) ,$$

where $\alpha \in \mathbb{R}_+$ and $\mathbf{\Lambda} \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{\Lambda} > 0$ are the parameters of the covariance function [29, 74, 28, 72]. However, the integral (2.21) cannot be solved analytically in general. Furthermore, the distribution of $\boldsymbol{y}_*$ is not necessarily Gaussian. In order to address the first issue, Girard et al. propose in [73] to use Monte-Carlo integration according to

$$p(y_*|d_*, \mathbf{X}, \underline{y}) \approx \sum_{i=1}^{N} \nu_i p(y_*|\underline{x}_*, \underline{x}_i, \mathbf{X}, \underline{y}) ,$$

where $\underline{x}_i$, $i = 1, \ldots, N$, $N \in \mathbb{N}$ are the positions and $\nu_i > 0$, $\sum_{i=1}^{N} \nu_i = 1$ are the corresponding weights of the Dirac density that approximates the probability distribution d_*. As a solution for the second issue, Girard et al. propose to maintain only the first two moments of $p(y_*|d_*, \mathbf{X}, \underline{y})$ [73]. In this way, the probability distribution of $\boldsymbol{y}_*$ is approximated with a Gaussian.

A very related approach to the method that we present in the next section was published by Dallaire et al. [42] and even applied to stochastic control in [43]. The main notion of this approach is to use so-called *mean covariance functions*. For the two inputs given in form of probability distributions $d_i = p_i(\underline{x})$ and $d_j = p_j(\underline{x})$, the mean covariance function is computed according to

$$k(d_i, d_j) = \int_{\mathbb{R}^{n_x}} \int_{\mathbb{R}^{n_x}} k(\underline{x}_i, \underline{x}_j) p_i(\underline{x}_i) p_j(\underline{x}_j) \, \mathrm{d}\underline{x}_i \, \mathrm{d}\underline{x}_j , \tag{2.22}$$

where $k(\underline{x}_i, \underline{x}_j)$ is any standard covariance function defined for deterministic vector-valued inputs. If the input distributions d_i and d_j are Gaussian and $k(\underline{x}_i, \underline{x}_j)$ is the SE covariance function, (2.22) can be expressed analytically as a function of the means and the covariances of d_i and d_j. Unfortunately, the estimation of the hyperparameters of a GP that uses mean covariance functions is not trivial due to local minima and usually requires a good prior $p(\underline{w})$ [42]. For this reason, we propose an alternative approach to GP regression with inputs given in terms of general probability distributions in the next section.

Other works such as [119] or [92] consider a related but substantially different problem of inputs corrupted by noise, i.e., each training input is a vector-valued sample of the corresponding input distribution. Such inputs are still finite, while the approach proposed in the next section is able to deal with inputs provided in form of probability distributions, i.e., infinite-dimensional inputs.

2.6.3 Gaussian Process Regression with Probability Distributions as Inputs

The main notion of our approach to GP regression with inputs provided in terms of general probability distributions consists in using stationary covariance functions[6] defined over *distances between the input distributions*. To illustrate this idea, consider the definition of the isotropic SE covariance function for deterministic vector-valued inputs $\underline{x}_i$ and $\underline{x}_j$

$$k(\underline{x}_i, \underline{x}_j) = \alpha^2 \exp\left(-\frac{1}{2}\frac{(\underline{x}_i - \underline{x}_j)^\top (\underline{x}_i - \underline{x}_j)}{l^2}\right) ,$$

where $l \in \mathbb{R}_+$. In this definition, the term $(\underline{x}_i - \underline{x}_j)^\top (\underline{x}_i - \underline{x}_j)$ corresponds to the squared Euclidean distance between $\underline{x}_i$ and $\underline{x}_j$. Now let the inputs to the GP be given in terms of probability distributions. Then, for two input distributions d_i and d_j, we define

$$k(d_i, d_j) = \alpha^2 \exp\left(-\frac{1}{2}\frac{D(d_i, d_j)^2}{l^2}\right) ,$$

where $D(d_i, d_j)$ denotes an admissible distance measure between the distributions d_i and d_j. The proposed approach can be applied to any stationary covariance function. Table 2.1 provides a small overview of covariance functions from [134] redefined in terms of the distance between the input distributions. A selection of distance measures that can be used within the proposed framework is presented below.

Covariance Function	**Expression**
squared exponential	$\exp\left(-\frac{1}{2}\frac{\Delta^2}{l^2}\right)$
Matérn	$\frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\sqrt{2\nu}}{l}\Delta\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}}{l}\Delta\right)$
exponential	$\exp\left(-\frac{\Delta}{l}\right)$
γ-exponential	$\exp\left(-\left(\frac{\Delta}{l}\right)^\gamma\right)$
rational quadratic	$\left(1 + \frac{\Delta^2}{2\alpha l^2}\right)^{-\alpha}$

Table 2.1: Covariance functions defined in terms of the distance $\Delta = D(d_i, d_j)$ between the input distributions d_i and d_j.

Please note that the choice of stationary covariance functions as the basis for the proposed framework is natural, because the space of probability distributions is, at best, only partially-ordered. Therefore, in contrast to the vector space $\mathbb{R}^{n_x}$, $n_x \in \mathbb{N}$, the notion of absolute position is undefined in this space. Nevertheless, it is possible to define covariance functions for inputs provided in terms of probability distributions that are not only defined over the distance

[6] A covariance function for deterministic vector-valued inputs is called stationary, if its value only depends on the distance between the inputs and not on their absolute positions.

between the inputs but also operate on the means or the modes of these inputs. To this end, we propose to combine stationary covariance functions defined over the distance between the input distributions with covariance functions of the means or the modes of these distributions. For this purpose, it is possible to use standard combination methods for covariance functions such as addition, multiplication, convolution, etc. [134]. For example, we can construct a covariance function from a bilinear function and the SE covariance function for the two input distributions d_i and d_j according to

$$k(d_i, d_j) = \hat{\underline{x}}_i^\top \mathbf{\Sigma} \hat{\underline{x}}_j \exp\left(-\frac{1}{2}\frac{D(d_i, d_j)^2}{l^2}\right) ,$$

where $\mathbf{\Sigma} \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{\Sigma} \geq 0$ and $\hat{\underline{x}}_i$ is the mean of d_i and $\hat{\underline{x}}_j$ is the mean of d_j, respectively.

Please note that in our framework, the parameter estimation remains the same as in the classical GP framework with deterministic vector-valued inputs. Moreover, the distances between the input distributions are independent of the GP hyperparameters and thus only have to be computed once. This constitutes and important advantage over the mean-kernel approach from [42], where the mean kernel has to be recomputed at each step of the parameter estimation algorithm.

Next, we provide a small, not necessarily complete overview of distance measures between probability distributions that can be used in the presented framework. In particular, the following distances are discussed:

- Total variation and L_p distances,
- Hellinger distance,
- Jensen–Shannon divergence,
- Wasserstein distance and OSPA,
- modified Cramér–von Mises distance.

A more thorough discussion of these distances together with other possible distance measures is available, e.g., in [178] and [69]. In our discussion, we not only provide the definitions of these distance measures but also emphasize the classes of the probability distributions that are amenable to them, i.e., whether the distance between two continuous distributions, two Dirac mixtures, or between a continuous and a Dirac distributions can be evaluated using the particular measure.

1. **Total variation and L_P distances:**
   The total variation distance [178] is defined for two continuous probability distributions with densities $p_1(\underline{x})$ and $p_2(\underline{x})$, $\underline{x} \in \Omega$ with respect to the measure μ according to

$$d_q(p_1, p_2) = \left(\int_\Omega (p_1(\underline{x}) - p_2(\underline{x}))^q \, \mathrm{d}\mu(\underline{x})\right)^{\frac{1}{q}}$$

with $q \in \mathbb{R}_{\geq 1}$. By choosing $\mathrm{d}\mu(\underline{x}) = \mathrm{d}\underline{x}$, the total variation becomes the L_q distance

$$d_q(p_1, p_2) = \left(\int_{\Omega} (p_1(\underline{x}) - p_2(\underline{x}))^q \, \mathrm{d}\underline{x} \right)^{\frac{1}{q}} .$$

2. **Hellinger distance:**
The Hellinger distance [178] is also defined for continuous probability distributions. For the two distributions with densities $p_1(\underline{x})$ and $p_2(\underline{x})$, $\underline{x} \in \Omega$, it holds

$$d(p_1, p_2) = \left(1 - \int_{\Omega} \sqrt{p_1(\underline{x}) p_2(\underline{x})} \, \mathrm{d}\underline{x} \right)^2 .$$

Please note that $0 \leq d(p_1, p_2) \leq 1$ holds for any two continuous distributions.

3. **Jensen–Shannon divergence:**
The main disadvantage of the Kullback–Leibler divergence [136] is that it is not symmetric. Therefore, Lin introduced a symmetric generalization thereof referred to as the Jensen–Shannon divergence in [111]. The Jensen–Shannon divergence is defined for two continuous distributions with densities $p_1(\underline{x})$ and $p_2(\underline{x})$, $\underline{x} \in \Omega$ according to

$$d(p_1, p_2) = \frac{1}{2} \int_{\Omega} p_1(\underline{x}) \log \frac{p_1(\underline{x})}{g(\underline{x})} + p_2(\underline{x}) \log \frac{p_2(\underline{x})}{g(\underline{x})} \, \mathrm{d}\underline{x} ,$$

where $g(\underline{x}) = (p_1(\underline{x}) + p_2(\underline{x}))/2$.

4. **Wasserstein distance and OSPA:**
The Wasserstein distance [169] between two continuous probability distributions with densities $p_1(\underline{x})$ and $p_2(\underline{x})$, $\underline{x} \in \Omega$ is defined according to

$$d_q(p_1, p_2) = \inf_h \left(\int_{\Omega} \int_{\Omega} d_e(\underline{x}, \underline{y})^q h(\underline{x}, \underline{y}) \, \mathrm{d}\underline{x} \, \mathrm{d}\underline{y} \right)^{\frac{1}{q}} ,$$

where $q \in \mathbb{R}_{\geq 1}$, $d_e(\underline{x}, \underline{y})$ is the Euclidean distance between the vectors $\underline{x}$ and $\underline{y}$, and $h(\underline{x}, \underline{y})$ is a probability density function with marginals p_1 and p_2, i.e., we have $\int_{\Omega} h(\underline{x}, \underline{y}) \, \mathrm{d}\underline{x} = p_2(\underline{y})$ and $\int_{\Omega} h(\underline{x}, \underline{y}) \, \mathrm{d}\underline{y} = p_1(\underline{x})$. Please observe that an evaluation of the Wasserstein distance is generally intractable because it requires to take the infimum over all possible joint distributions of p_1 and p_2. However, if the two distributions are Gaussian with means $\underline{\mu}_1$ and $\underline{\mu}_2$ and covariances $\mathbf{C}_1$ and $\mathbf{C}_2$, the Wasserstein distance with $q = 2$ evaluates to [75]

$$d_2(p_1, p_2)^2 = \|\underline{\mu}_1 - \underline{\mu}_2\|_2^2 + \mathrm{tr}\left[\mathbf{C}_1\right] + \mathrm{tr}\left[\mathbf{C}_2\right] - 2\mathrm{tr}\left[\sqrt{\sqrt{\mathbf{C}_1} \mathbf{C}_2 \sqrt{\mathbf{C}_1}} \right] .$$

For two Dirac densities with equal numbers of components, the derivation of the Wasserstein distance is presented in [89] and it is usually referred to as the Optimal MAss

Transfer (OMAT) distance [139]. The OMAT distance between two Dirac distributions p_1 and p_2 with $m \in \mathbb{N}$ equally weighted components $\underline{x}_i$ and $\underline{y}_i$, $i = 1, \ldots, m$ can be computed according to

$$d_q(p_1, p_2) = \min_{\pi \in \Pi} \left(\frac{1}{m} \sum_{i=1}^{m} d_e(\underline{x}_i, \underline{y}_{\pi,i})^q \right)^{\frac{1}{q}} , \tag{2.23}$$

where $q \in \mathbb{R}_{\geq 1}$, Π is the set of all possible assignments between the samples of p_1 and p_2, and $\underline{y}_{\pi,i}$ is the ith sample from the assignment π. In order to evaluate (2.23), it is possible to use the linear assignment algorithm from [100]. An extension of the OMAT distance to Dirac densities with different numbers of components is presented in the context of multi-target tracking in [139] under the name Optimal Sub-Pattern Assignment (OSPA) metric.

5. **Modified Cramér–von Mises distance:**
Except for the Wasserstein distance, the distances presented above are either suitable for two continuous distribution or two Dirac distributions and do not allow for computation of the distance between a continuous and a Dirac distributions. On the other hand, evaluation of the Wasserstein distance between two continuous distributions and a continuous and a discrete distribution is usually intractable. Therefore, we present the next distance measure which addresses this issue. This distance, referred to as the modified Cramér–von Mises distance (mCvMd), was proposed in [83] and further analyzed in [82, 70, 71, 69, 81].

Before presenting the mCvMd for two arbitrary (continuous or Dirac) distributions p_1 and p_2, we need to introduce the notion of the Localized Cumulative Distribution (LCD). The LCD $P(\underline{m}, b)$, $\underline{m} \in \mathbb{R}^{n_x}$, $n_x \in \mathbb{N}$, $b \in \mathbb{R}_+$ of a probability distribution $p(\underline{x})$, $\underline{x} \in \mathbb{R}^{n_x}$, $n_x \in \mathbb{N}$ is an integral transform defined according to

$$P(\underline{m}, b) = \int_{\mathbb{R}^{n_x}} p(\underline{x}) k(\underline{x}, \underline{m}, b) \, \mathrm{d}\underline{x} ,$$

where $k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times \mathbb{R}_+ \to \mathbb{R}_+$ is the LCD kernel such as the radial basis function

$$k(\underline{x}, \underline{m}, b) = \exp\left(-\frac{1}{2} \frac{(\underline{x} - \underline{m})^\top (\underline{x} - \underline{m})}{b^2} \right) .$$

Having computed the LCD $P_1(\underline{m}, b)$ of p_1 and the LCD $P_2(\underline{m}, b)$ of p_2, we can evaluate the mCvMd between p_1 and p_2 according to

$$d(p_1, p_2) = \left(\int_0^{b_{\max}} \int_{\mathbb{R}^{n_x}} w(b) (P_1(\underline{m}, b) - P_2(\underline{m}, b))^2 \, \mathrm{d}\underline{m} \, \mathrm{d}b \right)^{\frac{1}{2}} ,$$

where $w(b)$ is the weighting function

$$w(b) = \begin{cases} \frac{1}{b^{n_x - 1}} & \text{for } b \leq b_{\max} , \\ 0 & \text{otherwise} , \end{cases}$$

with a large constant $b_{\max} \in \mathbb{R}_+$.

In Sec. 2.2, we decided to represent the state estimates using Dirac distributions. For this reason, the GP regression also will have to operate on Dirac distributions. Thus, we present the mCvMd for Dirac distributions, whose derivation can be found in [81], for completeness. The mCvMd of the two Dirac distributions p_1 and p_2 with samples at $\underline{x}_i$, $i = 1, \ldots, M$ and $\underline{y}_j$, $j = 1, \ldots, L$ with weights α_i and β_j, respectively, and a large $b_{\max}$ is given by

$$d(p_1, p_2) = \left(\frac{\pi^{\frac{n_x}{2}}}{8} (D_1 - 2D_{12} + D_2 + 2c_b D_m) \right)^{\frac{1}{2}} ,$$

where $c_b = \log(4b_{\max}^2) - \Gamma$ with Γ being the Euler gamma constant and

$$\begin{aligned}
D_1 &= \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i \alpha_j \operatorname{xlog}\left((\underline{x}_i - \underline{x}_j)^\top (\underline{x}_i - \underline{x}_j) \right) , \\
D_2 &= \sum_{i=1}^{L} \sum_{j=1}^{L} \beta_i \beta_j \operatorname{xlog}\left((\underline{y}_i - \underline{y}_j)^\top (\underline{y}_i - \underline{y}_j) \right) , \\
D_{12} &= \sum_{i=1}^{M} \sum_{j=1}^{L} \alpha_i \beta_j \operatorname{xlog}\left((\underline{x}_i - \underline{y}_j)^\top (\underline{x}_i - \underline{y}_j) \right) , \text{ and} \\
D_m &= \left(\sum_{i=1}^{M} \alpha_i \underline{x}_i - \sum_{j=1}^{L} \beta_j \underline{y}_j \right)^\top \left(\sum_{i=1}^{M} \alpha_i \underline{x}_i - \sum_{j=1}^{L} \beta_j \underline{y}_j \right) ,
\end{aligned}$$

with

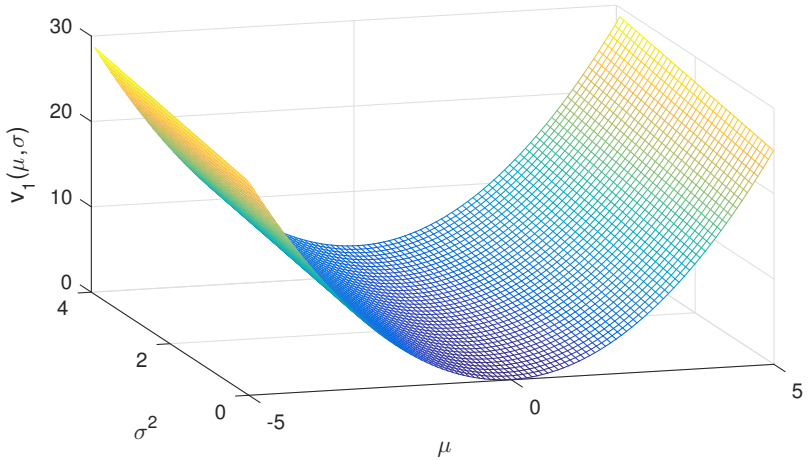
$$\operatorname{xlog}(x) = \begin{cases} 0 & \text{for } x = 0 , \\ x \log(x) & \text{otherwise} . \end{cases}$$

Finally, we briefly discuss, how vector-valued outputs can be approximated using a GP framework. An overview of available methods can be found in [2]. The main idea of these approaches is to define matrix-valued covariance functions as a sum of products of functions over the input space and functions that encode the interactions between the vector-valued outputs. Usually, so-called separable covariance functions are employed with

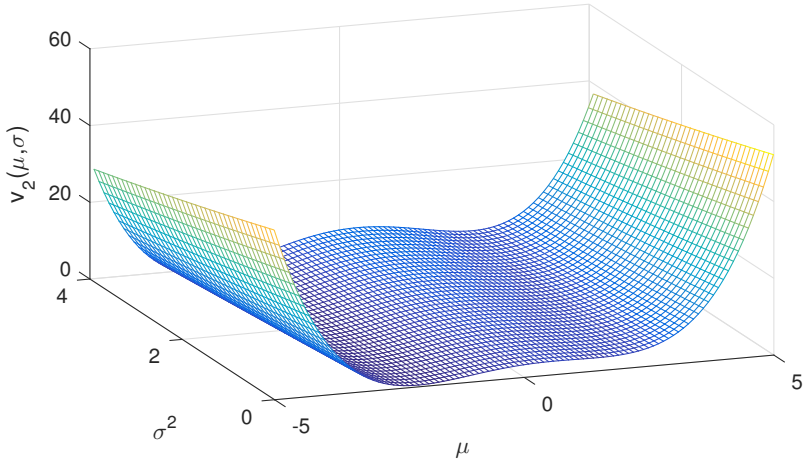
$$k(\underline{x}_i, \underline{x}_j) = \sum_{q=1}^{Q} k_q(\underline{x}_i, \underline{x}_j) \mathbf{B}_q ,$$

where $Q \in \mathbb{N}$ is the dimension of the output, $k_q(\underline{x}_i, \underline{x}_j)$ is the covariance function over the input space for qth component of the output for $q = 1, \ldots, Q$, and $\mathbf{B}_q$ is a symmetric positive definite matrix that encodes the interactions between the components of the output. To adapt separable covariance functions for multivariate outputs to the proposed framework, we need to replace the covariance functions $k_q(\underline{x}_i, \underline{x}_j)$ with functions defined over the distances between the probability distributions provided as inputs. As in GPs with univariate outputs, the proposed modification does not influence the parameter estimation of the GP. Thus, standard approaches can be applied. A review of available methods can be found in [2].

2.6.4 Numerical Evaluation



(a) Quadratic function $v_1(\mu, \sigma)$.



(b) Modified Rosenbrock function $v_2(\mu, \sigma)$.

Figure 2.7: Functions $v_1(\mu, \sigma)$ and $v_2(\mu, \sigma)$ that are used for evaluation.

In this section, we present a brief numerical evaluation of the GP regression that we described in the previous section. For the sake of visualization, our example shows GP regression with Gaussian input distributions and scalar outputs. In order to show the applicability of our approach in the scenario described in Sec. 2.2, the Gaussian inputs are only used to obtain the ground truth, while the GP is trained and evaluated with Dirac distributions sampled from the corresponding Gaussians.

We use two functions for evaluation, a simple quadratic function

$$\begin{aligned}\widetilde{v}_1(\underline{\boldsymbol{x}}) &= \mathrm{E}\left\{\underline{\boldsymbol{x}}^\top \underline{\boldsymbol{x}}\right\} \\ &= \mathrm{E}\left\{\underline{\boldsymbol{x}}\right\}^\top \mathrm{E}\left\{\underline{\boldsymbol{x}}\right\} + \mathrm{E}\left\{(\underline{\boldsymbol{x}} - \mathrm{E}\left\{\underline{\boldsymbol{x}}\right\})^\top (\underline{\boldsymbol{x}} - \mathrm{E}\left\{\underline{\boldsymbol{x}}\right\})\right\} \\ &= \underline{\mu}^\top \underline{\mu} + \mathbf{C} \ ,\end{aligned}$$

where $\underline{\mu}$ is the mean of $\underline{\boldsymbol{x}}$ and $\mathbf{C}$ its covariance, and a Rosenbrock function $\widetilde{v}_2$ slightly modified with an additional parameter d with

$$\widetilde{v}_2(x_1, x_2) = a(x_1 + b)^2 + c(x_1^2 - x_2 + d)^2 \ ,$$

where we set $x_1 = \mu$, $x_2 = \sigma$, $a = c = 0.1$, $b = 0.4$, and $d = -4$. In the considered scenario with univariate Gaussians as inputs, the two functions then are

$$\begin{aligned}v_1(\mu, \sigma) &= \mu^2 + \sigma^2 \ , \\ v_2(\mu, \sigma) &= 0.1(\mu + 4)^2 + 0.1(\mu^2 - \sigma^2 - 4)^2 \ .\end{aligned}$$

We choose the ranges for the mean and the variance to $\mu \in [-5, 5]$ and $\sigma^2 \in [0.1^2, 2^2]$. The functions v_1 and v_2 are depicted in Fig. 2.7.

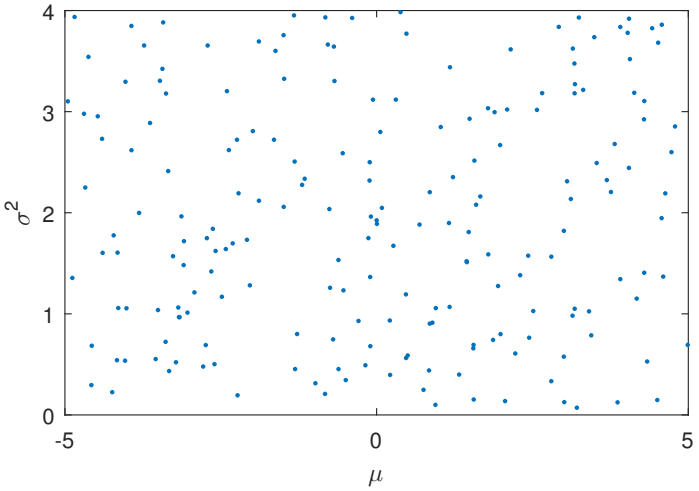
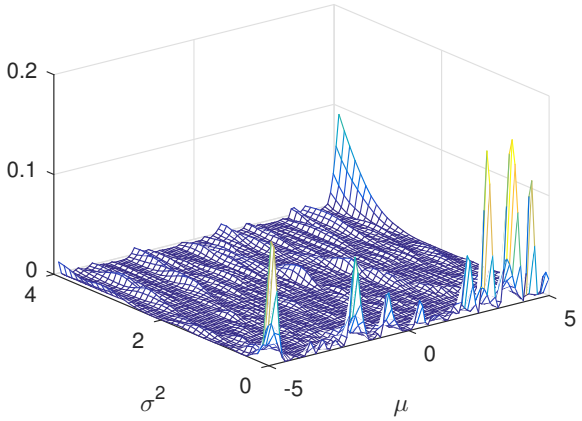


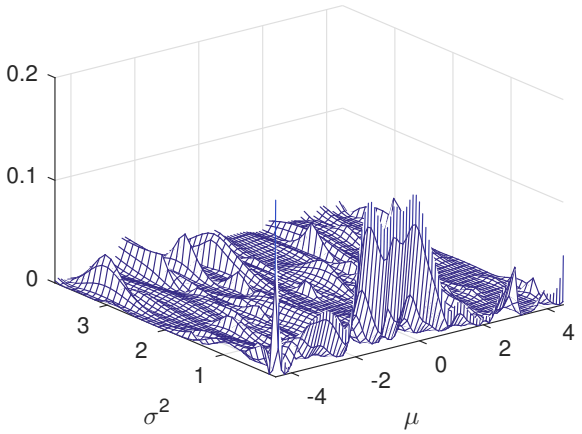
Figure 2.8: The set of 200 Gaussians that were used as training inputs. Each dot represents a Gaussian with corresponding mean μ and variance σ^2.

To generate the evaluation results, we proceed as follows. First, we draw 200 Gaussians from the considered set of Gaussians by drawing uniformly 200 means $\{\mu_i\}$, $i = 1, \ldots, 200$ from $[-5, 5]$ and 200 variances $\{\sigma_i^2\}$ from $[0.1^2, 2^2]$. Each pair (μ_i, σ_i^2) constitutes the parameters of a Gaussian for which we evaluate v_1 and v_2 in order to get the output values. The Gaussians that were generated in the presented numerical examples are depicted in Fig. 2.8, where each dot represents a Gaussian with corresponding mean and variance. Next, we draw 10 samples from each training Gaussian using the deterministic density approximation method from [82]. These samples of each Gaussian and the corresponding output values are used to train the GPs.

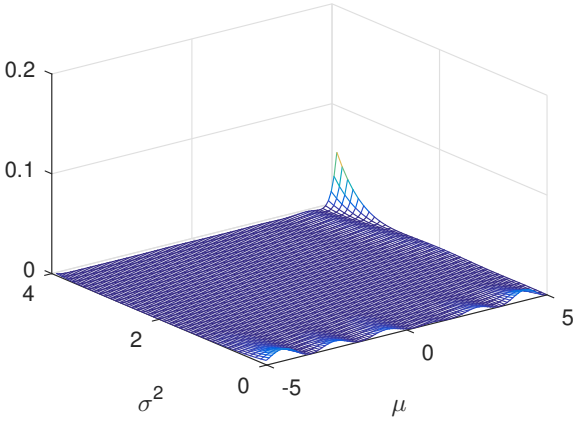
In our evaluation, we compare the following three GPs: (1) a GP designed according to the presented framework with mCvMd, (2) a GP designed according to the presented framework with OMAT distance, and (3) a GP designed according to the mean-kernel approach from [42]. All three GPs use the SE covariance function.



(a) Proposed – mCvMd.



(b) Proposed – OMAT distance (cropped to $\mu \in [-4.5, 4.5]$, $\sigma^2 \in [0.1, 3.9]$).

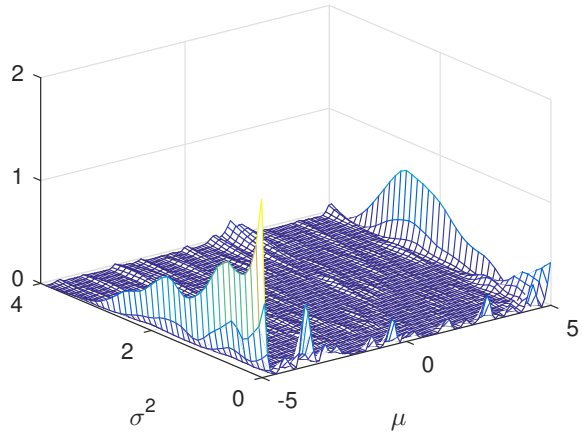


(c) Mean kernel [42].

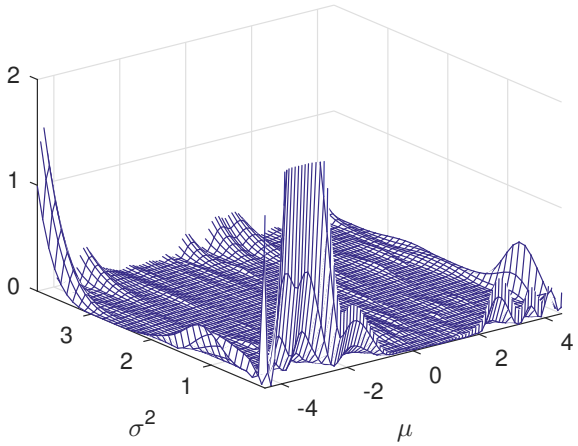
Figure 2.9: Quadratic error of the GP regression of the quadratic function $v_1(\mu, \sigma)$.

The simulation results are depicted in Fig. 2.9 and Fig. 2.10, where the quadratic errors between the true functions v_1 and v_2 and the corresponding approximations can be seen. To generate these plots, we use a uniform grid of test Gaussians with an increment of 0.1 in the mean and the variance. Again, as test inputs, we use Dirac distributions with 10 components that are generated from the test Gaussians using deterministic sampling from [82].

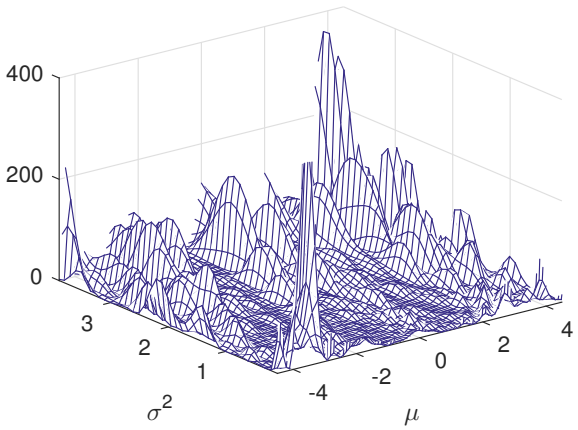
In Fig. 2.9, we can see that all three GPs perform well. The GP based on the mean kernel shows the best results and the GP based on the proposed framework with mCvMd is only slightly worse. The performance of the GP designed according to the presented framework with the OMAT distance is also acceptable. The results of the approximation of the Rosenbrock



(a) Proposed – mCvMd.



(b) Proposed – OMAT distance (cropped to $\mu \in [-4.5, 4.5]$, $\sigma^2 \in [0.1, 3.9]$).



(c) Mean kernel [42] (cropped to $\mu \in [-4.5, 4.5]$, $\sigma^2 \in [0.1, 3.9]$).

Figure 2.10: Quadratic error of the GP regression of the Rosenbrock function $v_2(\mu, \sigma)$.

function are more interesting. Here, we see that the GP designed with the presented framework with mCvMd shows the best performance and the GP based on our framework with the OMAT distance also provides an acceptable approximation quality. However, the GP based on the mean kernel approach from [42] is unable to approximate the Rosenbrock function v_2. As pointed out in [42], this issue can be caused by the nonlinearity of the likelihood and the lack of a good prior for the parameters. For this reason, we conclude that the usage of our GP framework for inputs given in terms of probability distributions is more convenient in practice.

CHAPTER 3

Control of MJLS without Mode Observation

Contents

In this chapter, we consider control of Markov Jump Linear Systems whose mode is not available to the controller. In particular, we address the static output-feedback problem, where the control law is an affine mapping of the measurements of the MJLS state to control inputs, and both the infinite-horizon and the finite-horizon dynamic output-feedback problems, where the controller possesses an internal state that is updated using the measurements of the MJLS state and that is linearly mapped to the control inputs. To this end, we first formally introduce MJLS and discuss related work on control of MJLS with both the observed and the non-observed mode. Before presenting the main results of this chapter in Sec. 3.5 and in Sec. 3.6, we will introduce the notions of stability for MJLS and review the stability conditions available in literature.

3.1 Motivation

Historically, the development of control theory concentrated on systems whose dynamics can be described by the laws of mechanics and thermodynamics, for example combustion motors, turbines, pendulums, etc. Such systems can be modeled by smooth (non)linear continuous-valued difference equations in discrete time or by differential equations in continuous time. However, the expansion of control theory into other areas such as economics, social and biological systems and networks, and interconnection of digital and mechanical systems resulted in system models that are subject to abrupt changes and thus cannot be captured by smooth dynamics. A convenient approach to model such systems is to introduce discrete-valued variables that describe the changes in the continuous-valued system dynamics. These systems are then referred to as hybrid systems because their state comprises both continuous-valued and discrete-valued state variables [138, 31]. There are many possibilities, how the discrete-valued and the continuous-valued states of a hybrid system can interact. For example, a change in the discrete-valued state can be driven by an independent random process, the discrete-valued state can be directly switched by a control mechanism, or change automatically when the continuous-valued state hits a bound in the state space. The variety of possible interactions between the continuous-valued and the discrete-valued dynamics allows for modeling of a broad range of dynamical systems. Examples of hybrid systems include among others interconnected power systems, chemical processes, job scheduling, traffic control, robotic systems, high-level manufacturing systems, automotive power trains, and many many more [138, 160, 154, 17, 90, 129, 4].

In this chapter, we will consider a special class of hybrid systems, namely the Markov Jump Linesr Systems (MJLS). This class of hybrid systems was introduced by Krasovskii and Lidskii in 1961 [104, 105, 106]. In MJLS, the continuous-valued dynamics are linear and the discrete-valued state, also referred to as the *mode*, is assumed to be independent of the continuous-valued dynamics and is modeled as a Markov chain. Although this assumption may seem restrictive, there are plenty of real-world applications that can be modeled as an MJLS. These applications include control of systems with component failures [115, 165, 164, 163], networked control [85, 65, 64], control of manufacturing processes [115], design of policies for economics [51, 57, 55], control of discrete event systems [33], etc. (see [38, 39] for more examples). However, probably the most famous application of MJLS is the experimental solar power plant Solar Two in the Mojave Desert in California, USA (Fig. 3.1) [38]. This plant uses light focused by mirrors on a boiler tower in order to heat molten salt to 565°C. This salt then boils water and the steam drives a turbine. Under optimal conditions, the plant can generate up to 10MWe. For efficient power generation, it is necessary to keep the steam temperature within adequate boundaries by controlling the flow rate of the water in the boiler. Control of the flow rate is necessary because clouds can cause abrupt variations in the amount of light focused on the boiler, which leads to abrupt changes in the dynamics of the boiler temperature. The developers of the Solar Two chose to model these abrupt variations as a Markov chain because modeling the occlusion of the sunlight by clouds as an independent and identically distributed stochastic process is not appropriate [157].

Figure 3.1: Experimental solar power plant Solar Two in the Mojave Desert in California, USA (source: National Renewable Energy Laboratory, USA)

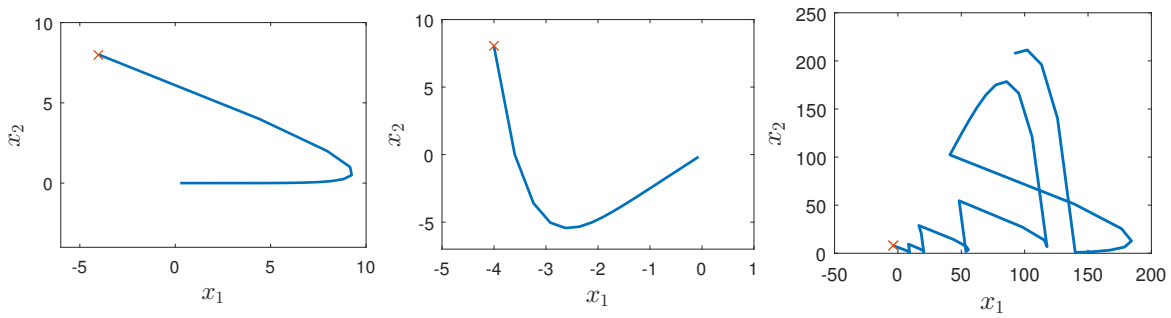
Most of the research devoted to control of MJLS considers settings, where the mode of the system is available to the controller. However, in many problems, e.g., the state of an economic market ("good" or "bad"), the mode often cannot be observed or is only available with such a large delay that it becomes irrelevant for control. For this reason, we address the problem of controlling MJLS, whose discrete-valued state is not observed.

3.2 Markov Jump Linear Systems

As pointed out in the introduction, the term MJLS refers to systems that comprise linear continuous-valued and discrete-valued dynamics. The discrete-valued dynamics, referred to as the *mode* of the system, are assumed to be driven by a Markov chain that is independent from the continuous-valued dynamics. Formally, the simplest discrete-time MJLS can be modeled by the difference equation

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_{\boldsymbol{\theta}_k} \underline{\boldsymbol{x}}_k \ , \tag{3.1}$$

where $k \in \mathbb{N}_0$ is the discrete time index and $\underline{\boldsymbol{x}}_k \in \mathbb{R}^{n_x}, n_x \in \mathbb{N}$, denotes the continuous-valued state of the system. In the remainder of this thesis, we will denote $\underline{\boldsymbol{x}}_k$ simply as the *state* of the system. The linear mapping $\mathbf{A}_{\boldsymbol{\theta}_k}$, referred to as the system matrix, is selected at each time step from a set of matrices $\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_M\}$, $M \in \mathbb{N}$ according to the value of the random variable $\boldsymbol{\theta}_k \in \{1, 2, \ldots, M\}$ that forms a Markov chain $\{\boldsymbol{\theta}_k\}$. If the set $\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_M\}$ is time-invariant, we will call the MJLS time-invariant. Otherwise, the MJLS will be referred to as time-variant MJLS. Throughout this thesis, we assume that the matrix matrix of transition probabilities $\mathbf{T} = (p_{ij})_{M \times M}$, $p_{ij} = \mathrm{P}(\boldsymbol{\theta}_{k+1} = j | \boldsymbol{\theta}_k = i)$, of this Markov chain is known. With a slight abuse of terminology, we will not only refer to $\boldsymbol{\theta}_k$ as the mode of the system but also to the corresponding continuous-valued dynamics $\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_i \underline{\boldsymbol{x}}_k$, $i \in \{1, \ldots, M\}$. Finally, we



(a) State trajectory of $\mathbf{A}_1$ **(b)** State trajectory of $\mathbf{A}_2$ **(c)** State trajectory of the MJLS

Figure 3.2: Influence of the switching on stability of an MJLS.

demand that the Markov chain $\{\boldsymbol{\theta}_k\}$ is time-homogeneous[1] whenever infinite-horizon control of MJLS without mode observation is under consideration.

Next, we want to demonstrate an important issue in MJLS regarding their stability. When we talk about stability of dynamic systems and in particular about Lyapunov stability [7], we refer to the property of the system to react to disturbances in such a way that the state of the system converges to the origin in infinite time, i.e., for the state $\underline{x}_k$ it must hold

$$\lim_{k\to\infty} \underline{x}^\top \underline{x} \to 0 \ .$$

For Linear Time-Invariant (LTI) systems, Lyapunov stability can be verified by calculating the eigenvalues of the system matrix: if the absolute values of all eigenvalues are smaller than one, then the system is stable; if the absolute value of at least one eigenvalue is larger than one, then the system is unstable; otherwise, it is not possible to infer stability or instability because it depends on the magnitude of the initial disturbance $\underline{\boldsymbol{x}}_0$, whether the state converges to the origin or not. In MJLS, stability of each individual mode does not imply stability of the MJLS, i.e., the switching in the dynamics (3.1) driven by the Markov chain $\{\boldsymbol{\theta}_k\}$ can make the MJLS unstable even if the dynamics $\mathbf{A}_i$ of each mode are stable. To demonstrate this issue, consider the dynamics (3.1) with

$$\mathbf{A}_1 = \begin{bmatrix} 0.9 & 1 \\ 0 & 0.5 \end{bmatrix} \ , \ \mathbf{A}_1 = \begin{bmatrix} 0.9 & 0 \\ 1 & 0.5 \end{bmatrix} \ , \text{ and } \mathbf{T} = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \ .$$

The eigenvalues of the modes are $\text{eig}\,(\mathbf{A}_1) = \{0.9, 0.5\}$ and $\text{eig}\,(\mathbf{A}_2) = \{0.5, 0.9\}$. From the stability theory of linear systems, we see that the individual modes are stable because the eigenvalues lie within the complex unit circle. This can also be visualized in the phase plane. Fig. 3.2 depicts an example trajectory of both modes with $\underline{\boldsymbol{x}}_0 = [-4 \ \ 8]^\top$. It can be seen that the trajectories converge to the origin. However, if we plot the trajectory of the MJLS for the same $\underline{\boldsymbol{x}}_0$, we see that the MJLS is unstable, which implies that we need a suitable notion of stability for MJLS and corresponding conditions to check it. Conditions available in the literature are summarized in the next section.

[1] A Markov chain is denoted as time-homogeneous if its state transition probabilities are time-invariant [79].

The above example also illustrates that controller design for MJLS has to consider the switching and its statistical properties in order to be able to stabilize the system. Otherwise even if the mode $\boldsymbol{\theta}_k$ is observed and we design a controller independently for each mode, the switching still could destabilize the system [38]. In the above example, $\mathbf{A}_{\boldsymbol{\theta}_k}$ would be the closed-loop system matrix of each mode. For this reason, research concentrates on control methods that explicitly consider the switching. Before addressing the control of MJLS and discussing the issues of control of MJLS without mode observation, we summarize the results on stochastic stability of MJLS available in literature in the next section.

3.3 Stochastic Stability of MJLS

For LTI systems, there is a unified definition of stability. However, for stochastic systems several notions of stability exist [108, 103, 66]. Particularly for MJLS, Ji et al. [99] proposed to consider the following three notions:

1. Stochastic stability
$$\mathrm{E}\left\{\sum_{k=0}^{\infty} \underline{\boldsymbol{x}}_k^\top \underline{\boldsymbol{x}}_k\right\} < \infty \ ,$$
2. Mean square stability
$$\lim_{k\to\infty} \mathrm{E}\left\{\underline{\boldsymbol{x}}_k^\top \underline{\boldsymbol{x}}_k\right\} = 0 \ ,$$
3. Exponential mean square stability
$$\mathrm{E}\left\{\underline{\boldsymbol{x}}_k^\top \underline{\boldsymbol{x}}_k\right\} < \alpha\beta^k \underline{x}_0^\top \underline{x}_0 \ ,$$
where $\alpha > 0$ and $0 < \beta < 1$,

for every initial condition $(\underline{\boldsymbol{x}}_0, \boldsymbol{\theta}_0)$ and showed that they are equivalent. In the remainder of this thesis, we will refer to the mean square stability only due to its analogy to Lyapunov stability of deterministic systems.

Derivation of stability criteria for MJLS has been addressed in a plethora of publications, e.g., in [18, 122, 123, 115, 107, 61, 36, 59, 21, 22, 77, 60, 151, 112]. The main results on necessary and sufficient conditions for mean square stability are given in the following theorem [122, 61, 60].

Theorem 3.1. *An MJLS with dynamics* (3.1) *is mean square stable, if there exist positive definite matrices* $\{\mathbf{S}_1, \ldots, \mathbf{S}_M\}$ *and* $\{\mathbf{Q}_1, \ldots, \mathbf{Q}_M\}, M \in \mathbb{N}$ *such that*

$$\sum_{j=1}^{M} p_{ij} \mathbf{A}_i^\top \mathbf{Q}_j \mathbf{A}_i - \mathbf{Q}_i = -\mathbf{S}_i \ ,$$

or

$$\sum_{i=1}^{M} p_{ij} \mathbf{A}_i \mathbf{Q}_i \mathbf{A}_i^\top - \mathbf{Q}_j = -\mathbf{S}_j \ .$$

The authors of [60] also present a very convenient condition for mean square stability of MJLS given in the following corollary.

Corollary 3.2. *The MJLS* (3.1) *is mean square stable, if the matrix*

$$\mathbf{M} = \operatorname{diag}\left[\mathbf{A}_1 \otimes \mathbf{A}_1,\ \mathbf{A}_2 \otimes \mathbf{A}_2,\ \ldots,\ \mathbf{A}_M \otimes \mathbf{A}_M\right](\mathbf{T}^\top \otimes \mathbf{I}_{n_x{}^2}) \ ,$$

where $\otimes$ *denotes the Kronecker product [14], is Schur*[2]*.*

In [112], it was shown that this condition is not only sufficient but also necessary.

Finally, we review the concept of mean square stabilizability and detectability for MJLS [93, 95, 35, 38].

Definition 3.3. *The MJLS* $\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \mathbf{B}_{\boldsymbol{\theta}_k}\underline{u}_k$, $\boldsymbol{\theta}_k \in \{1, 2, \ldots, M\}, M \in \mathbb{N}$ *is mean square stabilizable via state feedback*

i) with mode observation, if there exist matrices $\{\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_M\}$ *such that the closed-loop system* $\underline{\boldsymbol{x}}_{k+1} = (\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{B}_{\boldsymbol{\theta}_k}\mathbf{L}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{x}}_k$ *is mean square stable;*

ii) without mode observation, if there exists a matrix $\mathbf{L}$ *such that the closed-loop system* $\underline{\boldsymbol{x}}_{k+1} = (\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{B}_{\boldsymbol{\theta}_k}\mathbf{L})\underline{\boldsymbol{x}}_k$ *is mean square stable.*

Definition 3.4. *The MJLS with dynamics*

$$\begin{aligned} \underline{\boldsymbol{x}}_{k+1} &= \mathbf{A}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k \ , \\ \underline{\boldsymbol{y}}_k &= \mathbf{C}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k \ , \end{aligned}$$

$\boldsymbol{\theta}_k \in \{1, 2, \ldots, M\}, M \in \mathbb{N}$, *is mean square detectable*

i) with mode observation, if there exist matrices $\{\mathbf{K}_1, \mathbf{K}_2, \ldots, \mathbf{K}_M\}$ *such that the system* $\underline{\boldsymbol{z}}_{k+1} = (\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{K}_{\boldsymbol{\theta}_k}\mathbf{C}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{z}}_k$ *is mean square stable;*

ii) without mode observation, if there exists a matrix $\mathbf{K}$ *such that the system* $\underline{\boldsymbol{z}}_{k+1} = (\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{K}\mathbf{C}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{z}}_k$ *is mean square stable.*

3.4 Related Work and Intractability of the Optimal Control Law

The first control methods for MJLS with mode feedback can be traced back to 1971 when Wonham has solved the infinite-horizon continuous-time optimal control problem with state feedback using dynamic programming [174]. The optimal control of MJLS has then had its highest popularity in the late 80s and early 90s. At that time, continuous-time optimal control

[2] A matrix is said to be Schur, if all its eigenvalues lie within the open unit circle or alternatively if its spectral radius, i.e., the square of its largest eigenvalue, is smaller than 1.

with observed mode was considered, e.g., in [114, 114, 148] and in [117, 96, 97, 49, 50, 48], where the authors addressed the static output-feedback problem, i.e., the case where instead of the state the controller only uses an output thereof as a feedback. Furthermore, the authors of [97] presented controllability and observability conditions for MJLS with mode feedback. The dynamic output-feedback problem, where the controller maintains an estimate of the system state, was solved in [98]. In parallel, research on discrete-time control of MJLS with mode observation was reported, e.g., in [34, 93, 32, 94], where both the state and the dynamic output-feedback problems were solved. The corresponding controllability conditions were published in [95]. A more recent research on control of MJLS with observed mode concentrates on H_2 and H_∞ problems [37, 76, 68]. We refer to [39, 38] for a more thorough discussion of the publications mentioned so far.

One of the main insights in control of MJLS with mode observation is the fact that the *separation principle* holds. This implies that in the state and the static output feedback, the optimal control law consists of a (switched) linear regulator, while in the dynamic output-feedback case, the estimation of the continuous-valued state is independent of the control and vice versa [15]. Thus, the optimal controller consists of an optimal linear estimator and an optimal linear regulator with mode-dependent parameters. This is only possible because the choice of control inputs does not affect the uncertainty of the state estimates (see discussion in the Introduction). However, if the mode is not available to the controller, the separation principle does not hold and therefore there is a dual effect that makes the optimal control law nonlinear and intractable [78, 30, 11]. To demonstrate this issue, let us consider the MJLS with dynamics

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \mathbf{B}_{\boldsymbol{\theta}_k}\underline{u}_k + \underline{\boldsymbol{w}}_k \ , \tag{3.2}$$

where as in the previous chapter $\underline{\boldsymbol{x}}_k \in \mathbb{R}^{n_x}$ denotes the state, $\underline{u}_k \in \mathbb{R}^{n_u}, n_u \in \mathbb{N}$, the control input, and $\underline{\boldsymbol{w}}_k$ represents disturbances of the state dynamics. We assume that these disturbances $\underline{\boldsymbol{w}}_k$ are modeled as a stationary independent and identically distributed Gaussian stochastic process with zero mean and covariance $\mathbf{W}$. For simplicity, let us assume that we have state feedback, i.e., the state $\underline{\boldsymbol{x}}_k$ is available to the controller. However, only the initial mode $\boldsymbol{\theta}_0$ is given to the controller and the subsequent modes $\boldsymbol{\theta}_k, k \geq 1$ are not observed. For system (3.2), we seek to find a closed-loop control policy $\mu_k(\underline{\boldsymbol{x}}_{0:k}, \boldsymbol{\theta}_0)$ that yields the control inputs $\underline{u}_k$ based on the available state feedback $\underline{\boldsymbol{x}}_{0:k}$ and initial mode $\boldsymbol{\theta}_0$. The control inputs determined by the policy $\mu_k(\underline{\boldsymbol{x}}_{0:k}, \boldsymbol{\theta}_0)$ shall minimize the finite-horizon quadratic cost function

$$\mathcal{J} = \mathrm{E}\left\{ \underline{\boldsymbol{x}}_K^\top \mathbf{Q}_{\boldsymbol{\theta}_K}\underline{\boldsymbol{x}}_K + \sum_{k=0}^{K-1} \underline{\boldsymbol{x}}_k^\top \mathbf{Q}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \underline{u}_k^\top \mathbf{R}_{\boldsymbol{\theta}_k}\underline{u}_k \right\} \ , \tag{3.3}$$

where $K \in \mathbb{N}$ denotes the horizon length, the matrices $\mathbf{Q}_{\boldsymbol{\theta}_k}$ are positive semidefinite and $\mathbf{R}_{\boldsymbol{\theta}_k}$ are positive definite, and the expectation is taken with respect to $\underline{\boldsymbol{w}}_k$ and $\boldsymbol{\theta}_k$. In (3.3), $\mathbf{Q}_{\boldsymbol{\theta}_k}$ punishes the deviation of the state $\underline{\boldsymbol{x}}_k$ from the origin and $\mathbf{R}_k$ punishes the actuation. Note that minimizing the finite-horizon cost function (3.3) is easier than solving the infinite-horizon problems considered in Sec. 3.5 and Sec. 3.6.

In order to calculate the policies that yield optimal control inputs $\underline{u}_k$, we apply the DP algorithm [12, 15]. According to this algorithm, the optimal minimal costs (3.3) can be obtained by solving the recursion

$$\begin{aligned} \mathcal{J}_K^* &= \mathrm{E}\left\{ \mathcal{C}_K(\underline{\boldsymbol{x}}_K, \boldsymbol{\theta}_K) \middle| \mathcal{I}_K \right\} \ , \\ \mathcal{J}_k^* &= \inf_{\underline{u}_k} \mathrm{E}\left\{ \mathcal{C}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k, \boldsymbol{\theta}_k) + \mathcal{J}_{k+1}^* \middle| \mathcal{I}_k \right\} \ , \end{aligned} \tag{3.4}$$

where the stage costs $\mathcal{C}_K(\cdot)$ and $\mathcal{C}_k(\cdot)$ are defined according to

$$\begin{aligned} \mathcal{C}_K(\underline{\boldsymbol{x}}_K, \boldsymbol{\theta}_K) &= \underline{\boldsymbol{x}}_K^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \underline{\boldsymbol{x}}_K \ , \\ \mathcal{C}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k, \boldsymbol{\theta}_k) &= \underline{\boldsymbol{x}}_k^\top \mathbf{Q}_{\boldsymbol{\theta}_k} \underline{\boldsymbol{x}}_k + \underline{u}_k^\top \mathbf{R}_{\boldsymbol{\theta}_k} \underline{u}_k \ , \end{aligned}$$

and $\mathcal{I}_k$ denotes the information set that will be available to the controller at time step k. In our scenario, we have

$$\mathcal{I}_K = \{\underline{\boldsymbol{x}}_0, \boldsymbol{\theta}_0\} \text{ and } \mathcal{I}_k = \{\underline{\boldsymbol{x}}_k, \underline{u}_{k-1}, \mathcal{I}_{k-1}\} \ .$$

In order to show the intractability of the optimal control law that minimizes (3.3), consider the costs at time step $K-2$ of the recursion (3.4). These costs are given in the following proposition.

Proposition 3.5. *For the optimal costs at time step $K-2$, it holds*

$$\mathcal{J}_{K-2}^* = \inf_{\underline{u}_{K-2}} \mathrm{E}\left\{ \underline{\boldsymbol{x}}_{K-2}^\top \mathbf{Q}_{\boldsymbol{\theta}_{K-2}} \underline{\boldsymbol{x}}_{K-2} + \underline{u}_{K-2}^\top \mathbf{R}_{\boldsymbol{\theta}_{K-2}} \underline{u}_{K-2} + \underline{\boldsymbol{x}}_{K-1}^\top \boldsymbol{\Psi}(\mathcal{I}_{K-1}) \underline{\boldsymbol{x}}_{K-1} \middle| \mathcal{I}_{K-2} \right\} + \Delta \ , \tag{3.5}$$

where Δ is a positive constant and

$$\begin{aligned} \boldsymbol{\Psi}(\mathcal{I}_{K-1}) = \mathrm{E}&\left\{ \mathbf{Q}_{\boldsymbol{\theta}_{K-1}} + \mathbf{A}_{\boldsymbol{\theta}_{K-1}}^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \mathbf{A}_{\boldsymbol{\theta}_{K-1}} \middle| \mathcal{I}_{K-1} \right\} - \mathrm{E}\left\{ \mathbf{A}_{\boldsymbol{\theta}_{K-1}}^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \mathbf{B}_{\boldsymbol{\theta}_{K-1}} \middle| \mathcal{I}_{K-1} \right\} \\ &\times \left(\mathrm{E}\left\{ \mathbf{R}_{\boldsymbol{\theta}_{K-1}} + \mathbf{B}_{\boldsymbol{\theta}_{K-1}}^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \mathbf{B}_{\boldsymbol{\theta}_{K-1}} \middle| \mathcal{I}_{K-1} \right\} \right)^{-1} \mathrm{E}\left\{ \mathbf{B}_{\boldsymbol{\theta}_{K-1}}^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \mathbf{A}_{\boldsymbol{\theta}_{K-1}} \middle| \mathcal{I}_{K-1} \right\} \ . \end{aligned}$$

Proof. To show the result from Proposition 3.5, we evaluate the recursion (3.4). Time step K is trivial. Thus, beginning at time step $K-1$, it holds

$$\begin{aligned} \mathcal{J}_{K-1}^* &= \inf_{\underline{u}_{K-1}} \mathrm{E}\left\{ \underline{\boldsymbol{x}}_{K-1}^\top \mathbf{Q}_{\boldsymbol{\theta}_{K-1}} \underline{\boldsymbol{x}}_{K-1} + \underline{u}_{K-1}^\top \mathbf{R}_{\boldsymbol{\theta}_{K-1}} \underline{u}_{K-1} + \underline{\boldsymbol{x}}_K^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \underline{\boldsymbol{x}}_K \middle| \mathcal{I}_{K-1} \right\} \\ &= \inf_{\underline{u}_{K-1}} \underline{\boldsymbol{x}}_{K-1}^\top \mathrm{E}\left\{ \mathbf{Q}_{\boldsymbol{\theta}_{K-1}} + \mathbf{A}_{\boldsymbol{\theta}_{K-1}}^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \mathbf{A}_{\boldsymbol{\theta}_{K-1}} \middle| \mathcal{I}_{K-1} \right\} \underline{\boldsymbol{x}}_{K-1} \\ &\quad + 2\underline{u}_{K-1}^\top \mathrm{E}\left\{ \mathbf{B}_{\boldsymbol{\theta}_{K-1}}^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \mathbf{A}_{\boldsymbol{\theta}_{K-1}} \middle| \mathcal{I}_{K-1} \right\} \underline{\boldsymbol{x}}_{K-1} \\ &\quad + \underline{u}_{K-1}^\top \mathrm{E}\left\{ \mathbf{R}_{\boldsymbol{\theta}_{K-1}} + \mathbf{B}_{\boldsymbol{\theta}_{K-1}}^\top \mathbf{Q}_{\boldsymbol{\theta}_K} \mathbf{B}_{\boldsymbol{\theta}_{K-1}} \middle| \mathcal{I}_{K-1} \right\} \underline{u}_{K-1} + \mathrm{tr}\left[\mathrm{E}\left\{ \mathbf{Q}_{\boldsymbol{\theta}_K} \underline{\boldsymbol{w}}_{K-1} \underline{\boldsymbol{w}}_{K-1}^\top \middle| \mathcal{I}_{K-1} \right\} \right] \ , \end{aligned}$$

where we used (3.2) in order to express $\underline{\boldsymbol{x}}_K$ in terms of $\underline{\boldsymbol{x}}_{K-1}$ and $\underline{u}_{K-1}$, and exploited the fact that at time step $K-1$ the state $\underline{\boldsymbol{x}}_{K-1}$ is available to the controller. Evaluation of the

necessary optimality conditions with respect to $\underline{u}_{K-1}$ yields the optimal control input at time step $K-1$ that can be calculated according to

$$\underline{u}_{K-1} = -\left(\mathrm{E}\left\{\mathbf{R}_{\boldsymbol{\theta}_{K-1}} + \mathbf{B}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{B}_{\boldsymbol{\theta}_{K-1}}\middle|\mathcal{I}_{K-1}\right\}\right)^{-1}\mathrm{E}\left\{\mathbf{B}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{A}_{\boldsymbol{\theta}_{K-1}}\middle|\mathcal{I}_{K-1}\right\}\underline{\boldsymbol{x}}_{K-1} \ .$$

Thus, for the optimal costs at time step $K-1$, it holds

$$\begin{aligned}\mathcal{J}_{K-1}^{*} = \underline{\boldsymbol{x}}_{K-1}^{\top}\Big[&\mathrm{E}\left\{\mathbf{Q}_{\boldsymbol{\theta}_{K-1}} + \mathbf{A}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{A}_{\boldsymbol{\theta}_{K-1}}\middle|\mathcal{I}_{K-1}\right\} - \mathrm{E}\left\{\mathbf{A}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{B}_{\boldsymbol{\theta}_{K-1}}\middle|\mathcal{I}_{K-1}\right\}\\ &\times\left(\mathrm{E}\left\{\mathbf{R}_{\boldsymbol{\theta}_{K-1}} + \mathbf{B}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{B}_{\boldsymbol{\theta}_{K-1}}\middle|\mathcal{I}_{K-1}\right\}\right)^{-1}\mathrm{E}\left\{\mathbf{B}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{A}_{\boldsymbol{\theta}_{K-1}}\middle|\mathcal{I}_{K-1}\right\}\Big]\underline{\boldsymbol{x}}_{K-1} + \Delta \ .\end{aligned}$$

Using this result at time step $K-2$ concludes the proof. □

Note that at time step $K-1$, the optimal control law and the optimal costs are tractable although it is not linear in the state feedback $\underline{\boldsymbol{x}}_{K-1}$, because the expected values are functions of the mode distribution at time step $K-1$ that is a function of $\underline{\boldsymbol{x}}_{K-1}$. However, at time step $K-2$ the optimal control law is no longer tractable. Because $\mathrm{E}\left\{\underline{\boldsymbol{x}}_{K-1}^{\top}\boldsymbol{\Psi}(\mathcal{I}_{K-1})\underline{\boldsymbol{x}}_{K-2}\middle|\mathcal{I}_{K-2}\right\}$ contains a nested product of expected values, we cannot exploit the linearity of conditional expectations. In this way, the dynamic programming recursion (3.4) becomes intractable. To see that, we will evaluate the last term in (3.5). It holds

$$\begin{aligned}&\mathrm{E}\left\{\underline{\boldsymbol{x}}_{K-1}^{\top}\boldsymbol{\Psi}(\mathcal{I}_{K-1})\underline{\boldsymbol{x}}_{K-2}\middle|\mathcal{I}_{K-2}\right\}\\ &\quad= \int_{\mathbb{R}^{n_x}}\sum_{q=1}^{M}\underline{\boldsymbol{x}}_{K-1}^{\top}\Bigg[\left(\sum_{i=1}^{M}\sum_{j=1}^{M}\boldsymbol{\Psi}_1(i,j)p(\boldsymbol{\theta}_K=i,\boldsymbol{\theta}_{K-1}=j|\mathcal{I}_{K-1})\right)\\ &\qquad\qquad-\left(\sum_{i=1}^{M}\sum_{j=1}^{M}(\boldsymbol{\Psi}_2(i,j))^{\top}p(\boldsymbol{\theta}_K=i,\boldsymbol{\theta}_{K-1}=j|\mathcal{I}_{K-1})\right)\\ &\qquad\qquad\times\left(\sum_{i=1}^{M}\sum_{j=1}^{M}\boldsymbol{\Psi}_3(i,j)p(\boldsymbol{\theta}_K=i,\boldsymbol{\theta}_{K-1}=j|\mathcal{I}_{K-1})\right)^{-1}\\ &\qquad\qquad\times\left(\sum_{i=1}^{M}\sum_{j=1}^{M}\boldsymbol{\Psi}_2(i,j)p(\boldsymbol{\theta}_K=i,\boldsymbol{\theta}_{K-1}=j|\mathcal{I}_{K-1})\right)\Bigg]\\ &\qquad\times\underline{\boldsymbol{x}}_{K-1}p(\underline{\boldsymbol{x}}_{K-1},\boldsymbol{\theta}_{K-2}=q|\mathcal{I}_{K-2})\mathrm{d}\underline{\boldsymbol{x}}_{K-1} \ ,\end{aligned} \tag{3.6}$$

where we used the abbreviations

$$\begin{aligned}\boldsymbol{\Psi}_1(\boldsymbol{\theta}_K,\boldsymbol{\theta}_{K-1}) &= \mathbf{Q}_{\boldsymbol{\theta}_{K-1}} + \mathbf{A}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{A}_{\boldsymbol{\theta}_{K-1}} \ ,\\ \boldsymbol{\Psi}_2(\boldsymbol{\theta}_K,\boldsymbol{\theta}_{K-1}) &= \mathbf{B}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{A}_{\boldsymbol{\theta}_{K-1}} \ ,\\ \boldsymbol{\Psi}_3(\boldsymbol{\theta}_K,\boldsymbol{\theta}_{K-1}) &= \mathbf{R}_{\boldsymbol{\theta}_{K-1}} + \mathbf{B}_{\boldsymbol{\theta}_{K-1}}^{\top}\mathbf{Q}_{\boldsymbol{\theta}_K}\mathbf{B}_{\boldsymbol{\theta}_{K-1}} \ .\end{aligned}$$

The probability distributions of mode transitions in (3.6) are given by

$$p(\boldsymbol{\theta}_K=i,\boldsymbol{\theta}_{K-1}=j|\mathcal{I}_{K-1}) = p(\boldsymbol{\theta}_K=i,\boldsymbol{\theta}_{K-1}=j|\underline{\boldsymbol{x}}_{K-1:0},\underline{u}_{K-2:0},\boldsymbol{\theta}_0)$$

$$
\begin{aligned}
&= \sum_{t=1}^{M} p(\boldsymbol{\theta}_K = i, \boldsymbol{\theta}_{K-1} = j, \boldsymbol{\theta}_{K-2} = t | \underline{\boldsymbol{x}}_{K-1:0}, \underline{u}_{K-2:0}, \boldsymbol{\theta}_0) \\
&= \sum_{t=1}^{M} p(\boldsymbol{\theta}_K = i | \boldsymbol{\theta}_{K-1} = j) p(\boldsymbol{\theta}_{K-1} = j | \boldsymbol{\theta}_{K-2} = t) p(\boldsymbol{\theta}_{K-2} = t | \underline{\boldsymbol{x}}_{K-1:0}, \underline{u}_{K-2:0}, \boldsymbol{\theta}_0) \\
&= \sum_{t=1}^{M} p_{ji} p_{tj} \frac{p(\underline{\boldsymbol{x}}_{K-1} | \boldsymbol{\theta}_{K-2} = t, \underline{\boldsymbol{x}}_{K-2}, \underline{u}_{K-2})}{\sum_{r=1}^{M} p(\underline{\boldsymbol{x}}_{K-1} | \boldsymbol{\theta}_{K-2} = r, \underline{\boldsymbol{x}}_{K-2}, \underline{u}_{K-2})} p(\boldsymbol{\theta}_{K-2} = t | \underline{\boldsymbol{x}}_{K-2:0}, \underline{u}_{K-3:0}, \boldsymbol{\theta}_0) \ ,
\end{aligned}
$$

where $\underline{\boldsymbol{x}}_{K-1:0}$ denotes the sequence $\{\underline{\boldsymbol{x}}_{K-1}, \underline{\boldsymbol{x}}_{K-2}, \dots, \underline{\boldsymbol{x}}_1, \underline{\boldsymbol{x}}_0\}$ and where we can identify

$$
\frac{p(\underline{\boldsymbol{x}}_{K-1} | \boldsymbol{\theta}_{K-2} = t, \underline{\boldsymbol{x}}_{K-2}, \underline{u}_{K-2})}{\sum_{r=1}^{M} p(\underline{\boldsymbol{x}}_{K-1} | \boldsymbol{\theta}_{K-2} = r, \underline{\boldsymbol{x}}_{K-2}, \underline{u}_{K-2})} p(\boldsymbol{\theta}_{K-2} = t | \underline{\boldsymbol{x}}_{K-2:0}, \underline{u}_{K-3:0}, \boldsymbol{\theta}_0)
$$

as the estimate of a Wonham filter [173] for being in mode t at time step $K-2$ given the information set $\mathcal{I}_{K-2}$, and

$$
p(\underline{\boldsymbol{x}}_{K-1}, \boldsymbol{\theta}_{K-2} = q | \mathcal{I}_{K-2}) = p(\underline{\boldsymbol{x}}_{K-1} | \boldsymbol{\theta}_{K-2} = q, \underline{\boldsymbol{x}}_{K-2}, \underline{u}_{K-2}, \boldsymbol{\theta}_0) p(\boldsymbol{\theta}_{K-2} = q | \underline{\boldsymbol{x}}_{K-2:0}, \underline{u}_{K-3:0}, \boldsymbol{\theta}_0) \ .
$$

For the transition distribution $p(\underline{\boldsymbol{x}}_{K-1} | \boldsymbol{\theta}_{K-2} = r, \underline{\boldsymbol{x}}_{K-2}, \underline{u}_{K-2})$, it holds

$$
p(\underline{\boldsymbol{x}}_{K-1} | \boldsymbol{\theta}_{K-2} = r, \underline{\boldsymbol{x}}_{K-2}, \underline{u}_{K-2}) = \mathcal{N}(\underline{\boldsymbol{x}}_{k-1} | \mathbf{A}_r \underline{\boldsymbol{x}}_{k-2} + \mathbf{B}_r \underline{u}_{k-1}, \mathbf{W}) \ ,
$$

where $\mathcal{N}(\underline{x} | \underline{\mu}, \mathbf{C})$ denotes a Gaussian with mean $\underline{\mu}$ and covariance $\mathbf{C}$ evaluated at $\underline{x}$.

Finally, we can demonstrate the intractability of (3.6) in a one-dimensional example. To this end, assume the dynamics

$$
\underline{\boldsymbol{x}}_{k+1} = \begin{cases} \underline{\boldsymbol{x}}_k + \underline{u}_k + 0.2 \underline{\boldsymbol{w}}_k \ , & \text{for } \boldsymbol{\theta}_k = 1 \ , \\ 0.3 \underline{\boldsymbol{x}}_k + 0.8 \underline{u}_k + 0.2 \underline{\boldsymbol{w}}_k \ , & \text{for } \boldsymbol{\theta}_k = 2 \ , \end{cases}
$$

with transition matrix

$$
\mathbf{T} = \begin{bmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{bmatrix} \ ,
$$

and set $\mathbf{Q}_i = \mathbf{R}_i = 1$ for $i = \{1, 2\}$. Let the filtered probability distribution of $\boldsymbol{\theta}_{K-2}$ conditioned on the information set $\mathcal{I}_{K-2}$ be $p(\boldsymbol{\theta}_{K-2} | \underline{\boldsymbol{x}}_{K-2:0}, \underline{u}_{K-3:0}, \boldsymbol{\theta}_0) = [0.4 \ , \ 0.6]$. Then, for $\underline{\boldsymbol{x}}_{K-2} \in [-5, 5]$ and $\underline{u}_{K-2} \in [-5, 5]$, we obtain the expected costs depicted in Fig. 3.3. In this figure, we see that the costs are no longer quadratic as it is the case for LTI systems and MJLS with mode observation. Even worse, the costs are no longer convex. In fact, they involve sums and quotients of exponential functions, which makes an analytical computation of the optimal costs as a function of $\underline{\boldsymbol{x}}_{K-2}$ impossible and thus, the optimal control input $\underline{u}_{K-2}$ can only be determined using numerical global optimization algorithms. However, in this way, already computing the optimal control input at time step $K-3$ becomes intractable.

As we have seen, optimal control for MJLS without mode observation is intractable even in the state-feedback case. Thus, research concentrates on approximate but tractable control

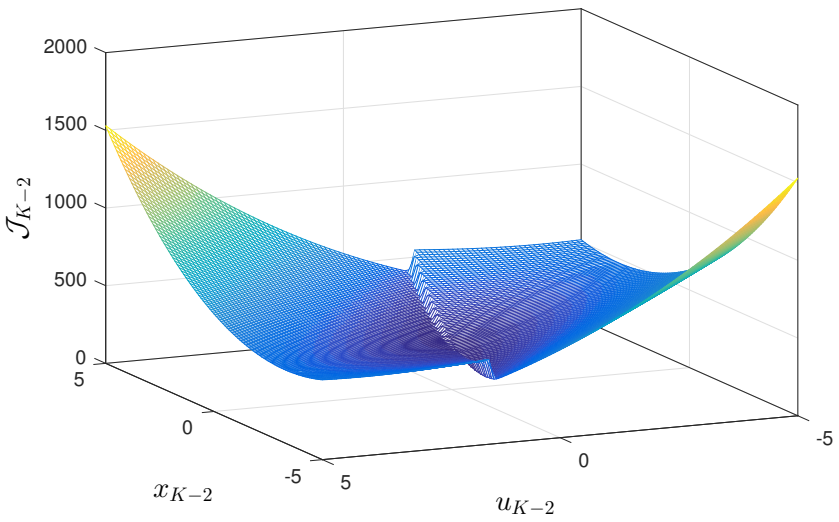


Figure 3.3: Costs $\mathcal{J}_{K-2}$ for different values of $\underline{\boldsymbol{x}}_{K-2}$ and $\underline{u}_{K-2}$

laws. A popular approach is to assume a linear controller. We will discuss the implications of this assumption in Sec. 3.5. Finite-horizon state-feedback control of MJLS without mode observation and under the assumption of a time-variant linear control law was, e.g., considered in [51, 165] and later extended to static output feedback in [163]. Time-invariant control was addressed in [52, 127, 164]. In [52], the authors considered H_2 state-feedback control with clustered mode observation, i.e., some of the mode values are observed and the others are not. The case of no mode observation is recovered if the cluster of observed modes is empty. In this work, the solution to the problem was formulated as an optimization problem with Linear Matrix Inequalities (LMI). An LMI-based formulation with a better bound on the H_2 norm was derived in [127] and an approach based on minimax optimization was presented in [67]. In contrast to the infinite-horizon control addressed in [52, 127], [164] considered the finite-horizon control problem. The time-invariant regulator gain is computed using nonlinear optimization algorithms. Finally, [63] considers dynamic output-feedback control of MJLS with clustered observations. However, the system parameters of the modes in the non-observed modes can only have jumps in the parameters that are related to noise. Other parameters must be deterministic. This assumption is very restrictive if we do not have mode feedback. While the control laws in [51, 52, 165, 127, 164, 163] are mode-independent, the authors of [40] assumed a feedback of a mode estimate. By doing so, they are able address the H_2 control problem. An approach to optimal control with a similar a priori separation of control and estimation was considered in [27, 80], where the authors employed a multiple-model filter in order to jointly estimate the state and the mode of the MJLS. In [27], the filter outputs the mean of the state and the filtered mode distribution to a linear mode-dependent regulator, while in [80] the most likely mode value is given to the regulator. However, by a priori assuming a separation of control and estimation, it becomes very hard to prove stability. An overview of relevant related work is given in Table 3.1.

	horizon	**feedback**	**time-**	**noise**
do Val et al. [51]	finite	state	variant	no
Vargas et al. [165]	finite	state	variant	yes
Vargas et al. [164]	finite	state	invariant	no
Vargas et al. [163]	finite	static output	variant	yes
do Val et al. [52]	infinite	state	invariant	yes
Oliveira et al. [127]	infinite	state	invariant	yes
Fioravanti et al. [63][3]	infinite	dynamic output	invariant	yes

Table 3.1: Overview of relevant related work.

The publications reviewed so far assume that the transition matrix of the MJLS is perfectly known. However, this may not always be the case. Thus, besides research on control of MJLS without mode observation, there is an important branch of research on control methods for MJLS whose transition matrix is uncertain (see e.g., [120, 45, 126, 176, 140]). Furthermore, we did not cover open-loop control methods for MJLS, such as Model Predictive Control (MPC) considered, e.g., in [20, 166, 143, 55]. A review of these methods is out of scope of this thesis.

As we have seen, control of MJLS with mode observation and state-feedback control of MJLS without mode observation can be considered as solved. On the other hand, the output-feedback control of MJLS without mode observation is still an active research topic [68] and will be addressed in the following three sections, where we consider time-invariant static output feedback, time-invariant dynamic output feedback, and time-variant dynamic output feedback.

3.5 Static Output Feedback

3.5.1 Problem Formulation

In this section, we address static output-feedback control of MJLS whose mode is not observed. The presented results were previously published in [184] (own publication) and extend our results from [183]. In particular, we consider an MJLS with state dynamics (3.2), where we add the measured output $\underline{\boldsymbol{y}}_k \in \mathbb{R}^{n_y}, n_y \in \mathbb{N}$ such that

$$\begin{aligned} \underline{\boldsymbol{x}}_{k+1} &= \mathbf{A}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \mathbf{B}_{\boldsymbol{\theta}_k}\underline{u}_k + \mathbf{H}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{w}}_k \ , \\ \underline{\boldsymbol{y}}_k &= \mathbf{C}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k \ . \end{aligned} \tag{3.7}$$

Furthermore, $\underline{\boldsymbol{w}}_k \in \mathbb{R}^{n_w}, n_w \in \mathbb{N}$ no longer necessarily has to have the same dimension as the state, but it is still assumed to have zero mean and identity covariance. This assumption can

[3] Considers only a very restrictive scenario.

be made without loss of generality because an appropriate choice of the matrices $\mathbf{H}_{\boldsymbol{\theta}_k}$ allows to model different covariances. Finally, we assume $\mathbf{C}_{\boldsymbol{\theta}_k}$ to have full row rank. This assumption is not limiting, since we can ignore outputs that are linear combinations of other outputs because they do not carry any additional information[4].

For the dynamics (3.7), we seek to determine the parameters of the static linear mode-independent output-feedback control law

$$\underline{u}_k = \mathbf{L}\underline{\boldsymbol{y}}_k \tag{3.8}$$

with a constant regulator gain $\mathbf{L}$ that is independent of the initial condition $(\underline{\boldsymbol{x}}_0, \boldsymbol{\theta}_0)$ and the modes at subsequent time steps, i.e., $\boldsymbol{\theta}_k, k \geq 1$. The regulator gain shall be chosen such that it minimizes the infinite-horizon quadratic cost function

$$\mathcal{J} = \lim_{K\to\infty} \frac{1}{K} \mathrm{E}\left\{ \sum_{k=0}^{K} \underline{\boldsymbol{x}}_k^\top \mathbf{Q}_{\boldsymbol{\theta}_k} \underline{\boldsymbol{x}}_k + \underline{u}_k^\top \mathbf{R}_{\boldsymbol{\theta}_k} \underline{u}_k \right\} , \tag{3.9}$$

where, as in (3.3), the matrices $\mathbf{Q}_{\boldsymbol{\theta}_k}$ are positive semidefinite, $\mathbf{R}_{\boldsymbol{\theta}_k}$ are positive definite, and the expectation is taken with respect to $\underline{\boldsymbol{w}}_k$ and $\boldsymbol{\theta}_k$.

Combining (3.7) and (3.8), we can construct the closed-loop dynamics

$$\underline{\boldsymbol{x}}_{k+1} = (\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{B}_{\boldsymbol{\theta}_k}\mathbf{L}\mathbf{C}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{x}}_k + \mathbf{H}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{w}}_k \ . \tag{3.10}$$

With this representation, we can summarize the considered control problem as follows.

Problem 3.6. *The optimal linear static output-feedback regulator gain for the MJLS* (3.7) *that minimizes* (3.9) *without mode observation is determined by the optimization problem*

$$\begin{aligned} \inf_{\mathbf{L}} \quad & \lim_{K\to\infty} \frac{1}{K} \mathrm{E}\left\{ \sum_{k=0}^{K} \underline{\boldsymbol{x}}_k^\top (\mathbf{Q}_{\boldsymbol{\theta}_k} + \mathbf{L}^\top \mathbf{R}_{\boldsymbol{\theta}_k}\mathbf{L}) \underline{\boldsymbol{x}}_k \right\} \\ \text{s.t.} \quad & \underline{\boldsymbol{x}}_{k+1} = (\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{B}_{\boldsymbol{\theta}_k}\mathbf{L}\mathbf{C}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{x}}_k + \mathbf{H}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{w}}_k \ . \end{aligned} \tag{3.11}$$

For this problem, we derive a feasibility condition and present an iterative algorithm for computation of the regulator gain $\mathbf{L}$ in the next two sections.

3.5.2 Feasibility Condition

In this section, we derive a sufficient feasibility condition for static output-feedback control of MJLS without mode observation (3.11) in terms of LMIs. To this end, we will convert the optimization problem (3.11) into an H_2 control problem [175] and express the latter as a Semi Definite Program (SDP) [87] that can be efficiently solved using, e.g., an interior-point

[4] This property only holds if the measurement equation is deterministic. If the measurement equation is perturbed by noise, each measurement entry contributes to inference.

method [23, 125]. The presented results mainly build upon the work of Crusius and Trofino [41] and do Val et al. [52]. Crusius and Trofino [41] derive LMI-based sufficient conditions for stabilizability of deterministic linear systems via static output feedback and do Val et al. [52] consider state-feedback control for MJLS with clustered observations, i.e., the modes are divided into two clusters, a cluster of observed and a cluster of non-observed modes. If the cluster of observed modes is empty, the problem corresponds to control of MJLS without mode observation. As in the scenario with deterministic parameters [41], the presented conditions are sufficient. Furthermore, they depend on the particular state space representation.

To recast Problem 3.6 as an H_2 problem, let $\underline{\boldsymbol{z}}_k \in \mathbb{R}^{(n_x+n_u)}$ denote the performance output of (3.7) with

$$\underline{\boldsymbol{z}}_k = \mathbf{D}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \mathbf{E}_{\boldsymbol{\theta}_k}\underline{u}_k \ , \tag{3.12}$$

where

$$\mathbf{D}_{\boldsymbol{\theta}_k} = \begin{bmatrix} \mathbf{Q}_{\boldsymbol{\theta}_k}^{\frac{1}{2}} \\ \mathbf{0}_{n_u \times n_x} \end{bmatrix} \text{ and } \mathbf{E}_{\boldsymbol{\theta}_k} = \begin{bmatrix} \mathbf{0}_{n_x \times n_u} \\ \mathbf{R}_{\boldsymbol{\theta}_k}^{\frac{1}{2}} \end{bmatrix} \ .$$

The matrices $\mathbf{Q}_{\boldsymbol{\theta}_k}^{\frac{1}{2}}$ and $\mathbf{R}_{\boldsymbol{\theta}_k}^{\frac{1}{2}}$ can be computed using, e.g., the Cholesky decomposition of $\mathbf{Q}_{\boldsymbol{\theta}_k}$ and $\mathbf{R}_{\boldsymbol{\theta}_k}$, respectively. For the closed-loop system (3.10), the performance output is given by

$$\underline{\boldsymbol{z}}_k = (\mathbf{D}_{\boldsymbol{\theta}_k} + \mathbf{E}_{\boldsymbol{\theta}_k}\mathbf{L}\mathbf{C}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{x}}_k \ . \tag{3.13}$$

With these prerequisites, we can state the following result.

Proposition 3.7. *The H_2 norm of a mean square stable MJLS* (3.10) *with performance output* (3.13) *is given by*

$$\|\mathcal{G}\|_2^2 = \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)\mathbf{X}_\infty^i(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)^\top\right] \ ,$$

where for $i \in \{1, \ldots, M\}$, $\mathbf{X}_\infty^i$ is the controllability Grammian that satisfies

$$\mathbf{X}_\infty^j = \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{X}_\infty^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top + \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top\right] \ ,$$

with μ_∞^i being the limit probability[5] $\lim_{k\to\infty} \mathrm{P}(\boldsymbol{\theta}_k = i)$ *[79].*

Proof. This result follows directly from [38], where it is shown that for the H_2 norm of a mean square stable MJLS

$$\begin{aligned} \underline{\boldsymbol{x}}_{k+1} &= \widetilde{\mathbf{A}}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \mathbf{H}_k\underline{\boldsymbol{w}}_k \ , \\ \underline{\boldsymbol{z}}_k &= \widetilde{\mathbf{D}}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k \end{aligned}$$

[5] Please recall that we assumed the Markov chain $\{\boldsymbol{\theta}_k\}$ to be time-homogeneous in infinite-horizon control. This implies that the limit distribution μ_∞ exists.

with performance output $\underline{\boldsymbol{z}}_k$, it holds

$$\begin{aligned}\|\mathcal{G}\|_2^2 &= \sum_{i=1}^{M} \operatorname{tr}\left[\widetilde{\mathbf{D}}_i \mathbf{X}_i \widetilde{\mathbf{D}}_i^\top\right] \ , \\ \mathbf{X}_j &= \sum_{i=1}^{M} p_{ij}\left[\widetilde{\mathbf{A}}_i \mathbf{X}_i \widetilde{\mathbf{A}}_i^\top + \mu_\infty^i \mathbf{H}_i \mathbf{H}_i^\top\right] \ , \qquad\qquad j \in \{1,\dots,M\} \ .\end{aligned}$$

Using the substitutions $\widetilde{\mathbf{A}}_i = \mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i$ and $\widetilde{\mathbf{D}}_i = \mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i$ yields the result of Proposition 3.7. □

By comparing the cost function in (3.11) and the squared H_2 norm $\|\mathcal{G}\|_2^2$, we see that the regulator gain $\mathbf{L}$ that yields the infimum of $\|\mathcal{G}\|_2^2$ also minimizes (3.11) because the two problems are equivalent. Now, consider the optimization problem

$$\begin{aligned}\inf_{\mathbf{L},\mathbf{X}_\infty^i} \quad & \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)\mathbf{P}_\infty^i(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)^\top\right] \\ \text{s.t.} \quad & \mathbf{P}_\infty^j > \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{P}_\infty^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top + \mu_\infty^i \mathbf{H}_i\mathbf{H}_i^\top\right] \ , \quad j \in \{1,\dots,M\} \ .\end{aligned} \tag{3.14}$$

According to [52], the solution of this optimization problem bounds the H_2 norm in Proposition 3.7 from above, because we have the partial ordering in the positive semidefinite sense $\mathbf{P}_\infty^i > \mathbf{X}_\infty^i \geq 0$. The result of the next lemma allows us to express (3.14) as an optimization problem that involves a *Bilinear Matrix Inequality* (BMI) [162].

Lemma 3.8. *Consider symmetric positive definite matrices* $\mathbf{G}$ *and* $\{\mathbf{Y}_1,\dots,\mathbf{Y}_M\}$*. The solution of the optimization problem* (3.14) *is bounded from above by the solution of*

$$\begin{aligned}\inf \quad & \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)^\top\right] \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i \mathbf{H}_i\mathbf{H}_i^\top & (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{G} \\ \mathbf{G}^\top(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top & \mathbf{G} + \mathbf{G}^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\ & \qquad\qquad\qquad\qquad i \in \{1,\dots,M\} \ ,\end{aligned} \tag{3.15}$$

where

$$\mathcal{D}_i(\mathbf{Y}) = \sum_{j=1}^{M} p_{ji}\mathbf{Y}_j \ .$$

Proof. Let $\{\mathbf{S}_1,\dots,\mathbf{S}_M\}$ be a set of symmetric positive definite auxiliary matrices. Then, we can write the condition in (3.14) for $j \in \{1,\dots,M\}$ as

$$\mathbf{P}_\infty^j = \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{P}_\infty^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top + \mu_\infty^i \mathbf{H}_i\mathbf{H}_i^\top + \mathbf{S}_i\right] \ .$$

Introducing

$$\mathbf{Y}_i = (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{P}_\infty^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top + \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top + \mathbf{S}_i$$

yields $\mathbf{P}_\infty^i = \mathcal{D}_i(\mathbf{Y})$. Applied to the condition in (3.14), this identity provides that

$$\mathbf{Y}_i - (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{P}_\infty^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top > 0\ . \tag{3.16}$$

Now, substituting $\mathbf{P}_\infty^i$ in the cost function in (3.14) and in (3.16) with $\mathbf{P}_\infty^i = \mathcal{D}_i(\mathbf{Y})$ gives

$$\begin{aligned} \inf\ & \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)^\top\right] \\ \text{s.t.}\ & \mathbf{Y}_i - (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top > 0 \\ & \qquad\qquad i \in \{1, \dots, M\}\ . \end{aligned} \tag{3.17}$$

Next, taking the Schur complement [113] of the inequality in (3.17) in $\mathcal{D}_i(\mathbf{Y})$ yields

$$\begin{aligned} &\begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top & (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i) \\ (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top & \mathcal{D}_i(\mathbf{Y})^{-1} \end{bmatrix} > 0 \\ &\qquad\qquad i \in \{1, \dots, M\}\ . \end{aligned} \tag{3.18}$$

Premultiplying this inequality with

$$\begin{bmatrix} \mathbf{I}_{n_x} & \mathbf{0}_{n_x} \\ \mathbf{0}_{n_x} & \mathbf{G} \end{bmatrix}\ , \tag{3.19}$$

postmultiplying with its transpose, and applying the inequality [68]

$$\mathbf{A}\mathbf{B}^{-1}\mathbf{A}^\top \geq \mathbf{A} + \mathbf{A}^\top - \mathbf{B} \tag{3.20}$$

concludes the proof. Note, that multiplication of (3.18) with (3.19) and postmultiplication with its transpose preserves the partial order of positive definite matrices because $\mathbf{G}$ is symmetric and positive definite. □

The optimization problem in Lemma 3.8 is a BMI because $(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{G}$ involves products of the decision variables $\mathbf{L}$ and $\mathbf{G}$. This issue renders the problem impractical because BMIs have been shown to be NP-hard [153]. For this reason, the main result of this section is the conversion of the problem from Lemma 3.8 into an optimization problem that only contains LMIs. By doing so, the problem becomes convex [23] and can be efficiently solved in polynomial time using, e.g., an interior-point method [125].

From the reformulated optimization problem (3.15), we obtain the sufficient feasibility conditions for the initial control problem. This result is given in the next theorem.

Theorem 3.9. *The optimization problem* (3.11) *is feasible if there exist a matrix* $\mathbf{F}$ *and symmetric positive definite matrices* $\mathbf{G}$, $\mathbf{M}$, $\{\mathbf{W}_1, \dots, \mathbf{W}_M\}$, $\{\mathbf{Y}_1, \dots, \mathbf{Y}_i\}$ *such that the LMIs*

$$\begin{bmatrix} \mathbf{W}_i & \mathbf{D}_i\mathbf{G} + \mathbf{E}_i\mathbf{F}\mathbf{C}_i \\ \mathbf{G}^\top\mathbf{D}_i^\top + \mathbf{C}_i^\top\mathbf{F}^\top\mathbf{E}_i^\top & \mathbf{G} + \mathbf{G}^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0$$
$$\begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top & \mathbf{A}_i\mathbf{G} + \mathbf{B}_i\mathbf{F}\mathbf{C}_i \\ \mathbf{G}^\top\mathbf{A}_i^\top + \mathbf{C}_i^\top\mathbf{F}^\top\mathbf{B}_i^\top & \mathbf{G} + \mathbf{G}^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0$$
$$\mathbf{M}\mathbf{C}_i = \mathbf{C}_i\mathbf{G}$$
$$i \in \{1, \dots, M\}$$

are feasible. The mean square stabilizing regulator gain is then given by $\mathbf{L} = \mathbf{F}\mathbf{M}^{-1}$.

Proof. We begin the proof by selecting matrices $\{\mathbf{W}_1, \dots, \mathbf{W}_M\}$ such that for $i \in \{1, \dots, M\}$

$$\mathbf{W}_i > (\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)^\top \ . \tag{3.21}$$

This implies

$$\sum_{i=1}^{M} \operatorname{tr}[\mathbf{W}_i] > \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)^\top\right] \ .$$

Thus, the solution of the optimization problem in Lemma 3.8 is bounded from above by the solution of

$$\begin{aligned} \inf \quad & \sum_{i=1}^{M} \operatorname{tr}[\mathbf{W}_i] \\ \text{s.t.} \quad & \mathbf{W}_i - (\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)^\top > 0 \\ & \begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top & \mathbf{A}_i\mathbf{G} + \mathbf{B}_i\mathbf{L}\mathbf{C}_i\mathbf{G} \\ \mathbf{G}^\top\mathbf{A}_i^\top + \mathbf{G}^\top\mathbf{C}_i^\top\mathbf{L}^\top\mathbf{B}_i^\top & \mathbf{G} + \mathbf{G}^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\ & i \in \{1, \dots, M\} \ . \end{aligned}$$

Taking the Schur complement of the first inequality constraint in the above equation, premultiplying it with (3.19), postmultiplying with its transpose, and applying the inequality (3.20) as in the proof of Lemma 3.8 yields

$$\begin{aligned} \inf \quad & \sum_{i=1}^{M} \operatorname{tr}[\mathbf{W}_i] \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{W}_i & \mathbf{D}_i\mathbf{G} + \mathbf{E}_i\mathbf{L}\mathbf{C}_i\mathbf{G} \\ \mathbf{G}^\top\mathbf{D}_i^\top + \mathbf{G}^\top\mathbf{C}_i^\top\mathbf{L}^\top\mathbf{E}_i^\top & \mathbf{G} + \mathbf{G}^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\ & \begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top & \mathbf{A}_i\mathbf{G} + \mathbf{B}_i\mathbf{L}\mathbf{C}_i\mathbf{G} \\ \mathbf{G}^\top\mathbf{A}_i^\top + \mathbf{G}^\top\mathbf{C}_i^\top\mathbf{L}^\top\mathbf{B}_i^\top & \mathbf{G} + \mathbf{G}^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\ & i \in \{1, \dots, M\} \ . \end{aligned} \tag{3.22}$$

Next, using the approach from [41], we introduce M additional constraints

$$\mathbf{M}\mathbf{C}_i = \mathbf{C}_i\mathbf{G}$$

for $i \in \{1, \ldots, M\}$. By doing so, it holds

$$\mathbf{C}_i = \mathbf{M}^{-1}\mathbf{C}_i\mathbf{G} \ .$$

Using this identity in (3.22) and introducing $\mathbf{F} = \mathbf{L}\mathbf{M}$, we obtain the LMI from Theorem 3.9. Finally, note that $\mathbf{M}$ is invertible since we assumed $\mathbf{C}_i$ to have full row rank. Thus, we have $\mathbf{L} = \mathbf{F}\mathbf{M}^{-1}$, which concludes the proof. □

Please note that the condition provided by Theorem 3.9 is only sufficient and not necessary for controllability of MJLS with non-observed mode via static output feedback, because we minimize an upper bound of the true H_2 norm and introduce additional constraints $\mathbf{M}\mathbf{C}_i = \mathbf{C}_i\mathbf{G}$, i.e., the derived optimization problem is a conservative approximation of the initial problem (3.30). In this context, the term conservative means that a feasible solution of the optimization problem

$$\begin{aligned}
\inf \quad & \sum_{i=1}^{M} \operatorname{tr}\left[\mathbf{W}_i\right] \\
\text{s.t.} \quad & \begin{bmatrix} \mathbf{W}_i & \mathbf{D}_i\mathbf{G} + \mathbf{E}_i\mathbf{F}\mathbf{C}_i \\ \mathbf{G}^\top\mathbf{D}_i^\top + \mathbf{C}_i^\top\mathbf{F}^\top\mathbf{E}_i^\top & \mathbf{G} + \mathbf{G}^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\
& \begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top & \mathbf{A}_i\mathbf{G} + \mathbf{B}_i\mathbf{F}\mathbf{C}_i \\ \mathbf{G}^\top\mathbf{A}_i^\top + \mathbf{C}_i^\top\mathbf{F}^\top\mathbf{B}_i^\top & \mathbf{G} + \mathbf{G}^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\
& \mathbf{M}\mathbf{C}_i = \mathbf{C}_i\mathbf{G} \\
& \qquad\qquad\qquad\qquad i \in \{1, \ldots, M\}
\end{aligned} \tag{3.23}$$

is also a feasible solution of the initial problem (3.11). Furthermore, the presented feasibility condition depends on the particular choice of the state space representation as pointed out in [41], i.e., the condition may be feasible in one representation and infeasible in an another. In case of deterministic parameters, there exist similarity transforms that ensure that the LMI becomes feasible if the initial system is stabilizable via static output feedback. If such transforms also exist for MJLS is an open question that needs to be investigated. Thus, we propose an alternative method in the next section, where we first compute the gain $\mathbf{L}$ and then use standard methods in order to check whether the closed-loop system is stable.

To conclude this section, we want to remark that the presented stabilizability condition also can be derived for MJLS with mode feedback as follows. If the mode is available, the regulator (3.8) becomes mode-dependent, i.e., we have

$$\underline{u}_k = \mathbf{L}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{y}}_k \ , \tag{3.24}$$

where the regulator gains $\mathbf{L}_{\boldsymbol{\theta}_k}$ are selected from the time-invariant set of gains $\{\mathbf{L}_1, \ldots, \mathbf{L}_M\}$ according to the mode $\boldsymbol{\theta}_k$ at time step k. Consequently, the H_2 norm of a mean square stable MJLS can be formulated as follows.

Proposition 3.10. *The H_2 norm of a mean square stable MJLS* (3.10) *with mode feedback and performance output* (3.13) *is given by*

$$\|\mathcal{G}\|_2^2 = \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)\mathbf{X}_\infty^i(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)^\top\right] \ ,$$

where $\mathbf{X}_\infty^i$ is the controllability Grammian that satisfies

$$\mathbf{X}_\infty^j = \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_i\mathbf{C}_i)\mathbf{X}_\infty^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_i\mathbf{C}_i)^\top + \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top\right] \ , \qquad j \in \{1,\dots,M\} \ .$$

Proof. The proof follows the proof of Proposition 3.7. □

As argued above, we now consider the problem

$$\begin{aligned} \inf_{\mathbf{L}_i,\mathbf{X}_\infty^i} \quad & \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)\mathbf{P}_\infty^i(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)^\top\right] \\ \text{s.t.} \quad & \mathbf{P}_\infty^j > \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_i\mathbf{C}_i)\mathbf{P}_\infty^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_i\mathbf{C}_i)^\top + \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top\right] \\ & \qquad\qquad j \in \{1,\dots,M\} \ . \end{aligned} \tag{3.25}$$

whose solution constitutes an upper bound of the H_2 norm from Proposition 3.7, because by construction we have $\mathbf{P}_\infty^i > \mathbf{X}_\infty^i \geq 0$. Next, we present an equivalent to Lemma 3.8 for MJLS with mode observation, where we reformulate (3.25) in terms of a BMI.

Lemma 3.11. *Consider symmetric positive definite matrices $\{\mathbf{G}_1,\dots,\mathbf{G}_M\}$ and $\{\mathbf{Y}_1,\dots,\mathbf{Y}_M\}$. The solution of the optimization problem* (3.25) *is bounded from above by the solution of*

$$\begin{aligned} \inf \quad & \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}\mathbf{C}_i)^\top\right] \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top & (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_i\mathbf{C}_i)\mathbf{G}_i \\ \mathbf{G}_i^\top(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_i\mathbf{C}_i)^\top & \mathbf{G}_i + \mathbf{G}_i^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\ & \qquad\qquad i \in \{1,\dots,M\} \ , \end{aligned}$$

where

$$\mathcal{D}_i(\mathbf{Y}) = \sum_{j=1}^{M} p_{ji}\mathbf{Y}_j \ .$$

Proof. The proof is omitted here, because it works along the lines of proof of Lemma 3.8. □

Finally, the next theorem presents sufficient conditions for stabilization of MJLS with mode observation via static output feedback. Please note that these conditions also depend on the state space representation as argued in [41].

Theorem 3.12. *The MJLS* (3.7) *is stabilizable via static output feedback* (3.24) *if there exist matrices* $\{\mathbf{F}_1, \ldots, \mathbf{F}_M\}$ *and symmetric positive definite matrices* $\{\mathbf{M}_1, \ldots, \mathbf{M}_M\}$, $\{\mathbf{G}_1, \ldots, \mathbf{G}_M\}$, $\{\mathbf{W}_1, \ldots, \mathbf{W}_M\}$, $\{\mathbf{Y}_1, \ldots, \mathbf{Y}_i\}$ *such that the LMIs*

$$\begin{aligned}
&\begin{bmatrix} \mathbf{W}_i & \mathbf{D}_i\mathbf{G}_i + \mathbf{E}_i\mathbf{F}_i\mathbf{C}_i \\ \mathbf{G}_i^\top\mathbf{D}_i^\top + \mathbf{C}_i^\top\mathbf{F}_i^\top\mathbf{E}_i^\top & \mathbf{G}_i + \mathbf{G}_i^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\
&\begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top & \mathbf{A}_i\mathbf{G}_i + \mathbf{B}_i\mathbf{F}_i\mathbf{C}_i \\ \mathbf{G}_i^\top\mathbf{A}_i^\top + \mathbf{C}_i^\top\mathbf{F}_i^\top\mathbf{B}_i^\top & \mathbf{G}_i + \mathbf{G}_i^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\
&\mathbf{M}_i\mathbf{C}_i = \mathbf{C}_i\mathbf{G}_i \\
&\qquad\qquad\qquad\qquad\qquad\qquad i \in \{1, \ldots, M\}
\end{aligned}$$

are feasible. The mean square stabilizing regulator gain is then given by $\mathbf{L}_i = \mathbf{F}_i\mathbf{M}_i^{-1}$.

Proof. We begin the proof by choosing matrices $\{\mathbf{W}_1, \ldots, \mathbf{W}_M\}$ such that for $i \in \{1, \ldots, M\}$, it holds

$$\mathbf{W}_i > (\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)^\top \ . \tag{3.26}$$

This choice of matrices $\{\mathbf{W}_1, \ldots, \mathbf{W}_M\}$ implies

$$\sum_{i=1}^{M} \operatorname{tr}[\mathbf{W}_i] > \sum_{i=1}^{M} \operatorname{tr}\left[(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)^\top\right] \ .$$

Thus, the solution of the optimization problem in Lemma 3.11 is bounded from above by the solution of

$$\begin{aligned}
\inf \quad & \sum_{i=1}^{M} \operatorname{tr}[\mathbf{W}_i] \\
\text{s. t.} \quad & \mathbf{W}_i - (\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)\mathcal{D}_i(\mathbf{Y})(\mathbf{D}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i)^\top > 0 \\
& \begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top & \mathbf{A}_i\mathbf{G}_i + \mathbf{B}_i\mathbf{L}_i\mathbf{C}_i\mathbf{G}_i \\ \mathbf{G}_i^\top\mathbf{A}_i^\top + \mathbf{G}_i^\top\mathbf{C}_i^\top\mathbf{L}_i^\top\mathbf{B}_i^\top & \mathcal{D}_i(\mathbf{Y})^{-1} \end{bmatrix} > 0 \\
& \qquad\qquad\qquad\qquad\qquad\qquad i \in \{1, \ldots, M\} \ .
\end{aligned}$$

Taking the Schur complement of the first inequality constraint in the above equation for each $i \in \{1, \ldots, M\}$, premultiplying it with

$$\begin{bmatrix} \mathbf{I}_{n_x} & \mathbf{0}_{n_x} \\ \mathbf{0}_{n_x} & \mathbf{G}_i \end{bmatrix}$$

and postmultiplying with its transpose, and applying the inequality (3.20) as in the proof of Lemma 3.8 yields

$$\begin{aligned} \inf \quad & \sum_{i=1}^{M} \operatorname{tr}\left[\mathbf{W}_i\right] \\ \text{s. t.} \quad & \begin{bmatrix} \mathbf{W}_i & \mathbf{D}_i\mathbf{G}_i + \mathbf{E}_i\mathbf{L}_i\mathbf{C}_i\mathbf{G}_i \\ \mathbf{G}_i^\top\mathbf{D}_i^\top + \mathbf{G}_i^\top\mathbf{C}_i^\top\mathbf{L}_i^\top\mathbf{E}_i^\top & \mathbf{G}_i + \mathbf{G}_i^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\ & \begin{bmatrix} \mathbf{Y}_i - \mu_\infty^i \mathbf{H}_i\mathbf{H}_i^\top & \mathbf{A}_i\mathbf{G}_i + \mathbf{B}_i\mathbf{L}_i\mathbf{C}_i\mathbf{G}_i \\ \mathbf{G}_i^\top\mathbf{A}_i^\top + \mathbf{G}_i^\top\mathbf{C}_i^\top\mathbf{L}_i^\top\mathbf{B}_i^\top & \mathbf{G}_i + \mathbf{G}_i^\top - \mathcal{D}_i(\mathbf{Y}) \end{bmatrix} > 0 \\ & \qquad\qquad\qquad\qquad\qquad\qquad i \in \{1, \dots, M\} \ . \end{aligned} \tag{3.27}$$

Next, using the approach from [41], we introduce M additional constraints

$$\mathbf{M}_i\mathbf{C}_i = \mathbf{C}_i\mathbf{G}_i$$

for $i \in \{1, \dots, M\}$. By doing so, it holds

$$\mathbf{C}_i = \mathbf{M}_i^{-1}\mathbf{C}_i\mathbf{G}_i \ .$$

Using this identity in (3.27) and introducing $\mathbf{F}_i = \mathbf{L}_i\mathbf{M}_i$, we obtain the LMI from Theorem 3.12. Note that $\mathbf{M}_i$ for $i \in \{1, \dots, M\}$ are invertible since we assumed $\mathbf{C}_i$ to have full row rank. Thus, we can calculate $\mathbf{L}_i = \mathbf{F}_i\mathbf{M}_i^{-1}$. □

3.5.3 Iterative Computation of the Regulator Gain

As argued in the previous section, the bound on the optimal value of (3.11) provided by the LMI in Theorem 3.9 is not necessarily tight and thus, the corresponding regulator gain can yield poor performance. For this reason, we present an alternative algorithm for computation of the regulator gain $\mathbf{L}$. To this end, we assume that the MJLS (3.7) is mean square stabilizable via static output feedback without mode observation according to the following definition.

Definition 3.13. *The MJLS* (3.7) *is stabilizable via static output feedback* (3.8)*, if there exists a matrix* $\mathbf{L}$ *such that the closed-loop system* (3.10) *is mean square stable.*

In our derivation of the iterative algorithm, we will express the initial optimization problem (3.11) in terms of the second moment of the state $\underline{\boldsymbol{x}}_k$. Then, from the necessary optimality conditions, we will obtain a set of coupled nonlinear equations that determine the optimal regulator gain $\mathbf{L}$. Because solving these equations is not trivial, we will provide a recursion that converges to the solution. Finally, we prove that if the initial problem is feasible, i.e., if the MJLS (3.7) is stabilizable via static mode-independent output feedback, then the algorithm converges to the optimal solution of the coupled equations. Note that the converse does not hold, i.e., the convergence of the iterative algorithm is not sufficient for stability of the closed-loop system. We will elaborate on this issue later in this section and in the numerical example in Sec. 3.5.4.

We begin the derivation of the algorithm for computation of the regulator gain $\mathbf{L}$ by defining the second moment of the state $\underline{\boldsymbol{x}}_k$ as

$$\mathbf{X}_k^i = \mathrm{E}\left\{\underline{\boldsymbol{x}}_k \underline{\boldsymbol{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \ , \tag{3.28}$$

where $\mathbb{1}$ denotes the indicator function with $\mathbb{1}_{\boldsymbol{\theta}_k=i} = 1$ if $\boldsymbol{\theta}_k = i$ and 0 otherwise. The next result gives the dynamics of the second moment $\mathbf{X}_k^i$.

Proposition 3.14. *The closed-loop dynamics of the second moment of* (3.7) *are given by*

$$\mathbf{X}_{k+1}^j = \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{X}_k^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^\top + \mu_k^i \mathbf{H}_i\mathbf{H}_i^\top\right] \ , \tag{3.29}$$

where $\mu_k^i = \mathrm{P}(\boldsymbol{\theta}_k = i)$ is the probability of being in mode $\boldsymbol{\theta}_k = i$ at time step k.

Proof. Let $p(A)$ denote the probability density function of A. Then, we have

$$\begin{aligned}
\mathbf{X}_{k+1}^j &= \mathrm{E}\left\{\underline{\boldsymbol{x}}_{k+1}\underline{\boldsymbol{x}}_{k+1}^\top \mathbb{1}_{\boldsymbol{\theta}_{k+1}}\right\} \\
&= \int_{\mathbb{R}^{n_x}} \sum_{j=1}^{M} \underline{\boldsymbol{x}}_{k+1}\underline{\boldsymbol{x}}_{k+1}^\top \mathbb{1}_{\boldsymbol{\theta}_{k+1}=j} p(\underline{\boldsymbol{x}}_{k+1}, \boldsymbol{\theta}_{k+1}) \ \mathrm{d}\underline{\boldsymbol{x}}_{k+1} \\
&= \int_{\mathbb{R}^{n_w}}\int_{\mathbb{R}^{n_x}}\int_{\mathbb{R}^{n_x}} \sum_{j=1}^{M} \underline{\boldsymbol{x}}_{k+1}\underline{\boldsymbol{x}}_{k+1}^\top \mathbb{1}_{\boldsymbol{\theta}_{k+1}=j} p(\underline{\boldsymbol{x}}_{k+1}, \underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_{k+1}, \boldsymbol{\theta}_k = i, \underline{\boldsymbol{w}}_k) \ \mathrm{d}\underline{\boldsymbol{x}}_{k+1} \ \mathrm{d}\underline{\boldsymbol{x}}_k \ \mathrm{d}\underline{\boldsymbol{w}}_k \\
&= \int_{\mathbb{R}^{n_w}}\int_{\mathbb{R}^{n_x}}\int_{\mathbb{R}^{n_x}} \sum_{j=1}^{M}\sum_{i=1}^{M} \underline{\boldsymbol{x}}_{k+1}\underline{\boldsymbol{x}}_{k+1}^\top \mathbb{1}_{\boldsymbol{\theta}_{k+1}=j}\mathbb{1}_{\boldsymbol{\theta}_k=i} p(\boldsymbol{\theta}_{k+1}|\boldsymbol{\theta}_k) p(\underline{\boldsymbol{x}}_{k+1}|\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k, \underline{\boldsymbol{w}}_k) \\
&\qquad\qquad \times p(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k) p(\underline{\boldsymbol{w}}_k) \ \mathrm{d}\underline{\boldsymbol{x}}_{k+1} \ \mathrm{d}\underline{\boldsymbol{x}}_k \ \mathrm{d}\underline{\boldsymbol{w}}_k \ .
\end{aligned}$$

It holds $p(\underline{\boldsymbol{x}}_{k+1}|\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k, \underline{\boldsymbol{w}}_k) = \delta(\underline{\boldsymbol{x}}_{k+1} - (\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{B}_{\boldsymbol{\theta}_k}\mathbf{L}\mathbf{C}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{x}}_k - \mathbf{H}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{w}}_k)$, where $\delta(\cdot)$ is the Dirac delta. Thus, evaluating the integral with respect to $\underline{\boldsymbol{x}}_{k+1}$ and taking the summation over $\boldsymbol{\theta}_{k+1}$, we obtain

$$\begin{aligned}
\mathbf{X}_{k+1}^j = \int_{\mathbb{R}^{n_w}}\int_{\mathbb{R}^{n_x}} \sum_{i=1}^{M} & p(\boldsymbol{\theta}_{k+1} = j|\boldsymbol{\theta}_k)[(\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{B}_{\boldsymbol{\theta}_k}\mathbf{L}\mathbf{C}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{x}}_k - \mathbf{H}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{w}}_k] \\
&\times [(\mathbf{A}_{\boldsymbol{\theta}_k} + \mathbf{B}_{\boldsymbol{\theta}_k}\mathbf{L}\mathbf{C}_{\boldsymbol{\theta}_k})\underline{\boldsymbol{x}}_k - \mathbf{H}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{w}}_k]^\top p(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k) p(\underline{\boldsymbol{w}}_k) \ \mathrm{d}\underline{\boldsymbol{x}}_k \ \mathrm{d}\underline{\boldsymbol{w}}_k \ .
\end{aligned}$$

Finally, integrating out $\underline{\boldsymbol{w}}_k$ yields (3.29). □

Note that the definition of the closed-loop second moment (3.28) with dynamics (3.29) can be interpreted as a representation of the closed-loop system state with a Gaussian mixture probability distribution that has M components. To see that, we can construct the dynamics

of the mean of the closed-loop system state and the central closed-loop second moment. For the means of the components of the Gaussian mixture, it holds (without proof)

$$\begin{aligned}\widetilde{\underline{x}}_{k+1}^{j} &= \mathrm{E}\left\{\underline{\boldsymbol{x}}_{k+1}\mathbb{1}_{\boldsymbol{\theta}_{k+1}=j}\right\} \\ &= \sum_{i=1}^{M} p_{ij}(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\widetilde{\underline{x}}_{k}^{i} \ ,\end{aligned}$$

and for the covariances

$$\begin{aligned}\widetilde{\mathbf{X}}_{k+1}^{j} &= \mathrm{E}\left\{(\underline{\boldsymbol{x}}_{k+1} - \widetilde{\underline{x}}_{k+1}^{j})(\underline{\boldsymbol{x}}_{k+1} - \widetilde{\underline{x}}_{k+1}^{j})^{\top}\mathbb{1}_{\boldsymbol{\theta}_{k+1}=j}\right\} \\ &= \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\widetilde{\mathbf{X}}_{k}^{i}(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^{\top} + \mu_k^i\mathbf{H}_i\mathbf{H}_i^{\top}\right] \ ,\end{aligned}$$

where $\widetilde{\mathbf{X}}_0^i = 0$. Due to linearity of the control law, the number of components of the Gaussian mixture remains constant and we obtain a finite-dimensional sufficient statistic for the hybrid state of the MJLS. If the control law is nonlinear, the number of the Gaussian mixture components grows exponentially with time, which yields an intractable control law [78].

Now, consider the cost function (3.9). Using that for the expectation $\mathrm{E}\left\{\underline{\boldsymbol{x}}_k^{\top}(\mathbf{Q}_{\boldsymbol{\theta}_k} + \mathbf{L}^{\top}\mathbf{R}_{\boldsymbol{\theta}_k}\mathbf{L})\underline{\boldsymbol{x}}_k\right\}$ with respect to $\boldsymbol{\theta}_k$ and $\underline{\boldsymbol{w}}_k$, it holds

$$\begin{aligned}\mathrm{E}\left\{\underline{\boldsymbol{x}}_k^{\top}(\mathbf{Q}_{\boldsymbol{\theta}_k} + \mathbf{L}^{\top}\mathbf{R}_{\boldsymbol{\theta}_k}\mathbf{L})\underline{\boldsymbol{x}}_k\right\} &= \sum_{i=1}^{M}\mathrm{E}\left\{\underline{\boldsymbol{x}}_k^{\top}(\mathbf{Q}_{\boldsymbol{\theta}_k} + \mathbf{L}^{\top}\mathbf{R}_{\boldsymbol{\theta}_k}\mathbf{L})\underline{\boldsymbol{x}}_k\mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \\ &= \sum_{i=1}^{M}\mathrm{E}\left\{\mathrm{tr}\left[(\mathbf{Q}_{\boldsymbol{\theta}_k} + \mathbf{L}^{\top}\mathbf{R}_{\boldsymbol{\theta}_k}\mathbf{L})\underline{\boldsymbol{x}}_k\underline{\boldsymbol{x}}_k^{\top}\mathbb{1}_{\boldsymbol{\theta}_k=i}\right]\right\} \\ &= \sum_{i=1}^{M}\mathrm{tr}\left[(\mathbf{Q}_{\boldsymbol{\theta}_k} + \mathbf{L}^{\top}\mathbf{R}_{\boldsymbol{\theta}_k}\mathbf{L})\mathbf{X}_k^i\right]\end{aligned}$$

and taking the limit $K \to \infty$, we obtain the following result.

Proposition 3.15. *The optimization problem* (3.11) *is equivalent to*

$$\begin{aligned}\inf_{\mathbf{L}} \quad & \sum_{i=1}^{M}\mathrm{tr}\left[(\mathbf{Q}_i + \mathbf{L}^{\top}\mathbf{R}_i\mathbf{L})\mathbf{X}_{\infty}^{i}\right] \\ \text{s. t.} \quad & \mathbf{X}_{\infty}^{j} = \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)\mathbf{X}_{\infty}^{i}(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}\mathbf{C}_i)^{\top} + \mu_{\infty}^{i}\mathbf{H}_i\mathbf{H}_i^{\top}\right] \ ,\end{aligned} \tag{3.30}$$

where $\lim_{k\to\infty}\mathrm{P}(\boldsymbol{\theta}_k = i)$ *denotes the limit probability of being in mode* i *[79].*

Next, we introduce the symmetric positive definite Lagrange multipliers $\{\mathbf{P}_{\infty}^1, \ldots, \mathbf{P}_{\infty}^M\}$. By doing so, we obtain the result of Proposition 3.16.

Proposition 3.16. *The Hamiltonian of* (3.30) *is given by*

$$\mathcal{H} = \sum_{i=1}^{M} \operatorname{tr}\Big[(\mathbf{Q}_i + \mathbf{C}_i^\top \mathbf{L}^\top \mathbf{R}_i \mathbf{L} \mathbf{C}_i)\mathbf{X}_\infty^i - \mathbf{P}_\infty^i \mathbf{X}_\infty^i + \mathcal{E}_i(\mathbf{P}_\infty)\Big[(\mathbf{A}_i + \mathbf{B}_i \mathbf{L} \mathbf{C}_i)\mathbf{X}_\infty^i (\mathbf{A}_i + \mathbf{B}_i \mathbf{L} \mathbf{C}_i)^\top + \mu_\infty^i \mathbf{H}_i \mathbf{H}_i^\top\Big]\Big] \ ,$$

where

$$\mathcal{E}_i(\mathbf{Y}) = \sum_{j=1}^{M} p_{ij} \mathbf{Y}_j \ .$$

Proof. Multiplying the constraint in (3.30) with the Lagrange multiplier yields

$$\mathcal{H} = \sum_{i=1}^{M} \operatorname{tr}\Big[(\mathbf{Q}_i + \mathbf{L}^\top \mathbf{R}_i \mathbf{L})\mathbf{X}_\infty^i\Big] - \sum_{j=1}^{M} \operatorname{tr}\Big[\mathbf{P}_\infty^j \mathbf{X}_\infty^j\Big] + \sum_{j=1}^{M}\sum_{i=1}^{M} p_{ij} \operatorname{tr}\Big[\mathbf{P}_\infty^j \Big[(\mathbf{A}_i + \mathbf{B}_i \mathbf{L} \mathbf{C}_i)\mathbf{X}_\infty^i (\mathbf{A}_i + \mathbf{B}_i \mathbf{L} \mathbf{C}_i)^\top + \mu_\infty^i \mathbf{H}_i \mathbf{H}_i^\top\Big]\Big] \ .$$

A change of indexes in the second summand and summation over j in the third summand concludes the proof. □

The optimal regulator gain that minimizes (3.30) and thus also (3.11) is determined by the next theorem.

Theorem 3.17. *The optimal static output-feedback control law* (3.8) *for the MJLS* (3.7) *that minimizes the infinite-horizon cost function* (3.9) *is determined by the following set of coupled nonlinear equations*

$$\sum_{i=1}^{M} (\mathbf{R}_i + \mathbf{B}_i^\top \mathcal{E}_i(\mathbf{P}_\infty)\mathbf{B}_i)\mathbf{L}\mathbf{C}_i \mathbf{X}_\infty^i \mathbf{C}_i^\top + \mathbf{B}_i^\top \mathcal{E}_i(\mathbf{P}_\infty)\mathbf{A}_i \mathbf{X}_\infty^i \mathbf{C}_i^\top = \mathbf{0}_{n_u \times n_y} \ , \tag{3.31}$$

$$\sum_{i=1}^{M} p_{ij}\Big[(\mathbf{A}_i + \mathbf{B}_i \mathbf{L} \mathbf{C}_i)\mathbf{X}_\infty^i (\mathbf{A}_i + \mathbf{B}_i \mathbf{L} \mathbf{C}_i)^\top + \mu_\infty^i \mathbf{H}_i \mathbf{H}_i^\top\Big] - \mathbf{X}_\infty^j = \mathbf{0}_{n_x} \ , \tag{3.32}$$

$$(\mathbf{A}_i + \mathbf{B}_i \mathbf{L} \mathbf{C}_i)^\top \mathcal{E}_i(\mathbf{P}_\infty)(\mathbf{A}_i + \mathbf{B}_i \mathbf{L} \mathbf{C}_i) + \mathbf{Q}_i + \mathbf{C}_i^\top \mathbf{L}^\top \mathbf{R}_i \mathbf{L} \mathbf{C}_i - \mathbf{P}_\infty^i = \mathbf{0}_{n_x} \ , \quad i, j \in \{1, \ldots, M\} \ . \tag{3.33}$$

Proof. The optimal solution of (3.30) and thus also of (3.11) can be found by solving

$$\inf_{\mathbf{L}, \mathbf{X}_\infty, \mathbf{P}_\infty} \ \mathcal{H} \ ,$$

where $\mathcal{H}$ is given in Proposition 3.16. For this problem, the necessary optimality conditions are given by

$$\frac{\partial \mathcal{H}}{\partial \mathbf{L}} = \mathbf{0}_{n_u \times n_y} \ , \ \frac{\partial \mathcal{H}}{\partial \mathbf{X}_\infty^i} = \mathbf{0}_{n_x} \ , \ \frac{\partial \mathcal{H}}{\partial \mathbf{P}_\infty^i} = \mathbf{0}_{n_x} \ .$$

Calculating the derivatives concludes the proof. □

Please note, that (3.31) can be analytically solved for $\mathbf{L}$ using the following proposition.

Proposition 3.18. *For the regulator gain in* (3.31) *from Theorem 3.17, it holds*

$$\mathbf{L} = \mathrm{vec}^{-1}\left[\mathrm{vec}\left[\mathbf{L}\right]\right] \ ,$$

where $\mathrm{vec}\left[\cdot\right]$ *denotes the vectorization operator and* $\mathrm{vec}^{-1}\left[\cdot\right]$ *the inverse vectorization operator [14], and* $\mathrm{vec}\left[\mathbf{L}\right]$ *is given by*

$$\mathrm{vec}\left[\mathbf{L}\right] = -\left(\sum_{i=1}^{M}\left[\mathbf{C}_i\mathbf{X}_\infty^i\mathbf{C}_i^\top \otimes (\mathbf{R}_i + \mathbf{B}_i^\top \mathcal{E}_i(\mathbf{P}_\infty)\mathbf{B}_i)\right]\right)^{-1} \mathrm{vec}\left[\sum_{i=1}^{M}\mathbf{B}_i^\top \mathcal{E}_i(\mathbf{P}_\infty)\mathbf{A}_i\mathbf{X}_\infty^i\mathbf{C}_i^\top\right] \ .$$

Alternatively, according to [23], $\mathbf{L}$ can be obtained via the SDP

$$\begin{aligned}
\min \ & \lambda \\
\text{s.t.} \ & \begin{bmatrix} \lambda\mathbf{I}_{n_u} & \mathbf{L} \\ \mathbf{L}^\top & \mathbf{I}_{n_y} \end{bmatrix} > 0 \\
& \sum_{i=1}^{M}(\mathbf{R}_i + \mathbf{B}_i^\top \mathcal{E}_i(\mathbf{P}_\infty)\mathbf{B}_i)\mathbf{L}\mathbf{C}_i\mathbf{X}_\infty^i\mathbf{C}_i^\top + \mathbf{B}_i^\top \mathcal{E}_i(\mathbf{P}_\infty)\mathbf{A}_i\mathbf{X}_\infty^i\mathbf{C}_i^\top = \mathbf{0}_{n_u \times n_y} \ .
\end{aligned}$$

This latter method to compute $\mathbf{L}$ has the advantage that it does not require the inversion of a $n_y n_u \times n_y n_u$ matrix. Therefore, it is numerically more robust.

As mentioned before, finding a solution to the coupled equations (3.31), (3.32), and (3.33) is not trivial. For this reason, following the discussion in [44], we propose the iterative procedure given in Algorithm 3.1.

In the simulations, we achieved a much faster convergence when $\mathbf{X}_{[0]}^i$ and $\mathbf{P}_{[0]}^i$ were initialized with $1\mathrm{e}{-4} \cdot \mathbf{I}_{n_x}$ rather than with random symmetric positive definite matrices.

At last, let us take a look at the convergence of the algorithm provided above. The result is given in the next theorem.

Theorem 3.19. *If the MJLS* (3.7) *is stabilizable via static output-feedback control* (3.8)*, then recursion* (3.35)–(3.34) *converges to its unique solution* $(\mathbf{X}_\infty^i, \mathbf{P}_\infty^i)$.

Algorithm 3.1 Computation of the regulator gain for static output feedback.

- *Step 1:* Set the counter $\eta = 0$ and initialize $\mathbf{X}^i_{[\eta]}$ and $\mathbf{P}^i_{[\eta]}$ with random symmetric positive definite $n_x \times n_x$ matrices.
- *Step 2:* Compute $\mathbf{L}_{[\eta]}$ using

$$\operatorname{vec}\left[\mathbf{L}_{[\eta]}\right] = -\left(\sum_{i=1}^{M}\left[\mathbf{C}_i\mathbf{X}^i_{[\eta]}\mathbf{C}_i^\top \otimes (\mathbf{R}_i + \mathbf{B}_i^\top \mathcal{E}_i(\mathbf{P}_{[\eta]})\mathbf{B}_i)\right]\right)^{-1} \operatorname{vec}\left[\sum_{i=1}^{M}\mathbf{B}_i^\top \mathcal{E}_i(\mathbf{P}_{[\eta]})\mathbf{A}_i\mathbf{X}^i_{[\eta]}\mathbf{C}_i^\top\right] . \tag{3.34}$$

- *Step 3:* Compute

$$\begin{aligned}\mathbf{X}^j_{[\eta+1]} &= \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[\eta]}\mathbf{C}_i)\mathbf{X}^i_{[\eta]}(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[\eta]}\mathbf{C}_i)^\top + \mu^i_\infty \mathbf{H}_i\mathbf{H}_i^\top\right] ,\\ \mathbf{P}^i_{[\eta+1]} &= (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[\eta]}\mathbf{C}_i)^\top \mathcal{E}_i(\mathbf{P}_{[\eta]})(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[\eta]}\mathbf{C}_i) + \mathbf{Q}_i + \mathbf{C}_i^\top \mathbf{L}_{[\eta]}^\top \mathbf{R}_i\mathbf{L}_{[\eta]}\mathbf{C}_i .\end{aligned} \tag{3.35}$$

- *Step 4:* Stop the algorithm if $\mathbf{X}^i_{[\eta]}$ and $\mathbf{P}^i_{[\eta]}$ converged, i.e., if the distances $\|\mathbf{X}^i_{[\eta]} - \mathbf{X}^i_{[\eta+1]}\|$ and $\|\mathbf{P}^i_{[\eta]} - \mathbf{P}^i_{[\eta+1]}\|$ are sufficiently small, where $\|\cdot\|$ is an admissible matrix norm. Otherwise, set $\eta = \eta + 1$ and return to *Step 2*.

Proof. The proof follows argumentation of the proof of Theorem 3 in [44]. First, we show that any solution of (3.32)–(3.33) yields a stabilizing regulator gain. Since we assume that the considered MJLS is mean square stabilizable via static output feedback, there exists a solution to the coupled equations of Theorem 3.17. Assume that $(\overline{\mathbf{X}^i_L}, \overline{\mathbf{P}^i_L})$ is such a solution for some fixed regulator gain $\mathbf{L}$. Then, the regulator gain $\mathbf{L}$ determined by this solution stabilizes the MJLS in the mean square sense because in order for $(\overline{\mathbf{X}^i_L}, \overline{\mathbf{P}^i_L})$ to exist, the closed-loop system must be mean square stable. This implies that all solutions of the equations (3.31)–(3.33) yield stabilizing regulator gains. Next, we have to show that the equations in Theorem 3.17 have only one solution and that the recursion (3.35) with (3.34) converges to this solution for $\eta \to \infty$. To this end, we define the homotopy

$$\begin{aligned}\mathbf{X}^j_{[\alpha,\eta+1]} &= \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}^i_{[\alpha,\eta]}\mathbf{C}_i)\mathbf{X}^i_{[\eta]}(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}^i_{[\alpha,\eta]}\mathbf{C}_i)^\top + \mu^i_\infty \mathbf{H}_i\mathbf{H}_i^\top\right] ,\\ \mathbf{P}^i_{[\alpha,\eta+1]} &= (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}^i_{[\alpha,\eta]}\mathbf{C}_i)^\top \mathcal{E}_i(\mathbf{P}_{[\alpha,\eta]})(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}^i_{[\alpha,\eta]}\mathbf{C}_i) + \mathbf{Q}_i + \mathbf{C}_i^\top (\mathbf{L}^i_{[\alpha,\eta]})^\top \mathbf{R}_i\mathbf{L}^i_{[\alpha,\eta]}\mathbf{C}_i ,\end{aligned}$$

where

$$\mathbf{L}^i_{[\alpha,\eta]} = (1-\alpha)\mathbf{L}^i_{[\eta]} + \alpha\mathbf{L}_{[\eta]}$$

with $\mathbf{L}^i_{[\eta]}$ being the regulator gains for the case that the mode $\boldsymbol{\theta}_k$ is observed [140] and $\mathbf{L}_{[\eta]}$ being determined by (3.34). For $\alpha = 0$, we have the static output-feedback control problem with observed mode and for $\alpha = 1$, we recover the control problem considered in this section. As α goes from 0 to 1, we follow the solution path $(\mathbf{X}^{i,[\alpha]}_\infty, \mathbf{P}^{i,[\alpha]}_\infty)$ from the

solution $(\mathbf{X}_\infty^{i,[0]}, \mathbf{P}_\infty^{i,[0]})$ to the desired solution $(\mathbf{X}_\infty^i, \mathbf{P}_\infty^i)$. Note that for $\alpha = 0$, the solution $\lim_{\eta\to\infty}(\mathbf{X}_{[0,\eta]}^i, \mathbf{P}_{[0,\eta]}^i) = (\mathbf{X}_\infty^{i,[0]}, \mathbf{P}_\infty^{i,[0]})$ is unique, i.e., the recursion

$$\mathbf{X}_{[0,\eta+1]}^j = \sum_{i=1}^{M} p_{ij}\left[(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[0,\eta]}^i\mathbf{C}_i)\mathbf{X}_{[0,\eta]}^i(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[0,\eta]}^i\mathbf{C}_i)^\top + \mu_\infty^i\mathbf{H}_i\mathbf{H}_i^\top\right] ,$$
$$\mathbf{P}_{[0,\eta+1]}^i = (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[0,\eta]}^i\mathbf{C}_i)^\top\mathcal{E}_i(\mathbf{P}_{[0,\eta]})(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[0,\eta]}^i\mathbf{C}_i) + \mathbf{Q}_i + \mathbf{C}_i^\top(\mathbf{L}_{[0,\eta]}^i)^\top\mathbf{R}_i\mathbf{L}_{[0,\eta]}^i\mathbf{C}_i$$

with $\mathbf{L}_{[0,\eta]}^i$ being the solutions of the SDP

$$\begin{aligned}
&\min \ \lambda \\
&\text{s.t.} \quad \begin{bmatrix} \lambda\mathbf{I}_{n_u} & \mathbf{L}_{[0,\eta]}^i \\ (\mathbf{L}_{[0,\eta]}^i)^\top & \mathbf{I}_{n_y} \end{bmatrix} > \mathbf{0}_{n_u+n_y} \\
&\quad -(\mathbf{R}_i + \mathbf{B}_i^\top\mathcal{E}_i(\mathbf{P}_{[0,\eta]})\mathbf{B}_i)\mathbf{L}_{[0,\eta]}^i\mathbf{C}_i\mathbf{X}_{[0,\eta]}^i\mathbf{C}_i^\top + \mathbf{B}_i^\top\mathcal{E}_i(\mathbf{P}_{[0,\eta]})\mathbf{A}_i\mathbf{X}_{[0,\eta]}^i\mathbf{C}_i^\top = \mathbf{0}_{n_u\times n_y}
\end{aligned}$$

converges to $(\mathbf{X}_\infty^{i,[0]}, \mathbf{P}_\infty^{i,[0]})$, if the MJLS with observed mode is stabilizable via static output feedback. Then, it follows from the topological degree theory that the number of solutions along the homotopy path for α going from 0 to 1 remains constant [116, 135, 1]. This implies that $\lim_{\eta\to\infty}(\mathbf{X}_{[1,\eta]}^i, \mathbf{P}_{[1,\eta]}^i) = \lim_{\eta\to\infty}(\mathbf{X}_{[\eta]}^i, \mathbf{P}_{[\eta]}^i)$ has a unique solution $(\mathbf{X}_\infty^i, \mathbf{P}_\infty^i)$ to which the recursion (3.35), (3.34) converges. □

In the next section, we provide a numerical example, in which we demonstrate the proposed algorithms.

3.5.4 Numerical Examples

To demonstrate the performance of the control law (3.8) with the regulator gain computed using the algorithm from Section 3.5.3, we choose the parameters of (3.7) to

$$\mathbf{A}_1 = \begin{bmatrix} 1.2 & 1.2 \\ 0 & 1 \end{bmatrix} , \ \mathbf{A}_2 = \begin{bmatrix} 1 & 0.8 \\ 0 & 1 \end{bmatrix} , \ \mathbf{B}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} , \ \mathbf{B}_2 = \begin{bmatrix} 0 \\ 0.2 \end{bmatrix} , \ \mathbf{H}_1 = \mathbf{H}_2 = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.1 \end{bmatrix} ,$$

and the cost matrices to $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{I}_{n_x}$ and $\mathbf{R}_1 = \mathbf{R}_2 = \mathbf{I}_{n_x}$. We compare the presented method with the optimal time-invariant state-feedback control law from [38] that requires mode feedback, the state-feedback approach with clustered observations from [52], and the time-variant static output-feedback controller from [163]. A reference implementation of the proposed computation method for the regulator gain is available on GitHub [53].

We conduct a Monte Carlo simulation with 1e5 runs à 100 time steps each and investigate the influence of the transition matrix by considering two different transition matrices

$$\mathbf{T}_{[1]} = \begin{bmatrix} 0.7 & 0.3 \\ 0.6 & 0.4 \end{bmatrix} \text{ and } \mathbf{T}_{[2]} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} .$$

		$\mathbf{T}_1$, $\underline{\boldsymbol{x}}_0 = \begin{bmatrix}0\\0\end{bmatrix}$	$\mathbf{T}_1$, $\underline{\boldsymbol{x}}_0 = \begin{bmatrix}3\\0\end{bmatrix}$	$\mathbf{T}_2$, $\underline{\boldsymbol{x}}_0 = \begin{bmatrix}0\\0\end{bmatrix}$	$\mathbf{T}_2$, $\underline{\boldsymbol{x}}_0 = \begin{bmatrix}3\\0\end{bmatrix}$
state feedback	time-invariant optimal with observed mode [38]	0.9718	24.1879	0.8280	24.5886
	time-invariant with non-observed mode [52]	0.9788	24.4780	1.0088	26.0436
	time-variant with non-observed mode [165]	0.9743	24.4031	0.8786	24.7490
	proposed	0.9730	24.4808	0.8855	25.4265
output feedback	time-variant with non-observed observed mode [163]	1.5611	26.5108	1.0351	30.6060
	proposed	1.5623	28.3532	1.0539	30.7562

Table 3.2: Median costs of the Monte Carlo simulation.

Furthermore, we investigate the influence of the initial state. To this end, we choose two different initial states

$$\underline{\boldsymbol{x}}_0^{[1]} = \begin{bmatrix}0\\0\end{bmatrix} \text{ and } \underline{\boldsymbol{x}}_0^{[2]} = \begin{bmatrix}3\\0\end{bmatrix} .$$

Note that using Corollary 3.2, we obtain the spectral radius $\rho(\mathbf{M}(\mathbf{T}_{[1]}) = 1.2970$ of the MJLS with $\mathbf{T}_{[1]}$ and $\rho(\mathbf{M}(\mathbf{T}_{[2]}) = 1.3295$ for the MJLS with $\mathbf{T}_{[2]}$, respectively. Thus, the uncontrolled MJLS is unstable in the mean square sense in both cases.

First, we choose the observation matrices to

$$\mathbf{C}_1 = \mathbf{C}_2 = \begin{bmatrix}1 & 0\end{bmatrix} .$$

With these output matrices, the LMI from Theorem 3.9 is infeasible for both transition matrices $\mathbf{T}_{[1]}$ and $\mathbf{T}_{[2]}$. On the other hand, the iterative algorithm from Sec. 3.5.3 converges with the regulator gains $\mathbf{L}_{[1]} = -0.0089$ for the transition matrix $\mathbf{T}_{[1]}$ and with the gain $\mathbf{L}_{[2]} = -0.0122$ for $\mathbf{T}_{[2]}$. However, if we evaluate the spectral radii of the controlled MJLS according to Corollary 3.2, we obtain $\rho(\mathbf{M}(\mathbf{L}_{[1]}, \mathbf{T}_{[1]})) = 1.1463$ and $\rho(\mathbf{M}(\mathbf{L}_{[2]}, \mathbf{T}_{[2]})) = 1.1548$, respectively. This means that the MJLS with the transition matrix $\mathbf{T}_{[1]}$ cannot be stabilized using the regulator gain $\mathbf{L}_{[1]}$ and also the MJLS with $\mathbf{T}_{[2]}$ is not stabilizable using $\mathbf{L}_{[2]}$. This observation demonstrates that convergence of the proposed iterative algorithm is not sufficient for stability of the closed-loop system.

Now, we set the observation matrices to

$$\mathbf{C}_1 = \begin{bmatrix}1 & 2\end{bmatrix} \text{ and } \mathbf{C}_2 = \begin{bmatrix}2 & 1\end{bmatrix} .$$

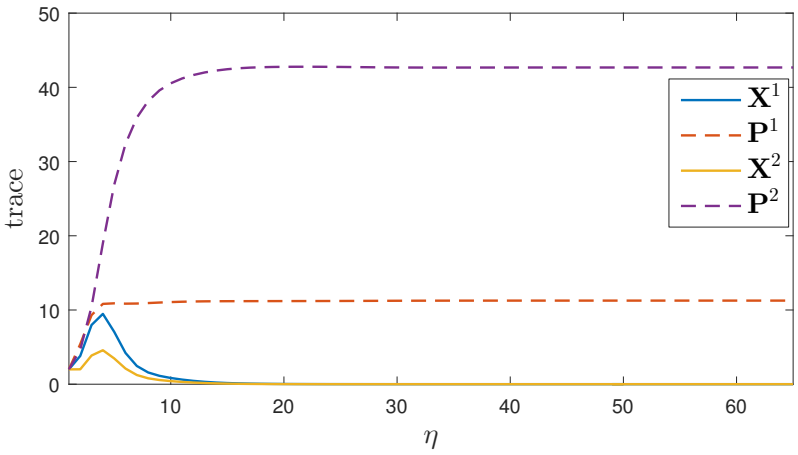


Figure 3.4: Depicted are the traces of $\mathbf{X}_{[\eta]}^i$ and $\mathbf{P}_{[\eta]}^i$ for $i = \{1, 2\}$, whose convergence indicates convergence of Algorithm 3.1.

According to Theorem 3.9, these output matrices yield feasible LMIs for MJLS with both transition matrices $\mathbf{T}_{[1]}$ and $\mathbf{T}_{[2]}$. Thus, the systems can be stabilized via mode-independent static output feedback. Using the proposed iterative algorithm, we obtain the regulator gains $\mathbf{L}_{[1]} = -0.6350$ for $\mathbf{T}_{[1]}$ and $\mathbf{L}_{[2]} = -0.4404$ for $\mathbf{T}_{[2]}$. The regulator gain $\mathbf{L}_{[1]}$ yields the spectral radius $\rho(\mathbf{M}(\mathbf{L}_{[1]}, \mathbf{T}_{[1]})) = 0.8178$ of the closed-loop MJLS and $\mathbf{L}_{[2]}$ yields $\rho(\mathbf{M}(\mathbf{L}_{[2]}, \mathbf{T}_{[2]})) = 0.9871$. Fig. 3.4 depicts the convergence of $\mathbf{X}_{[\eta]}^i$ and $\mathbf{P}_{[\eta]}^i$ for transition matrix $\mathbf{T}_{[1]}$. It can be seen that in this scenario, convergence is achieved after approximately 15 iterations.

The spectral radii of the MJLS controlled with the regulator gains computed using the proposed algorithms are smaller than one, which means that the closed-loop systems are stable. The results of the Monte-Carlo simulation are given in Table 3.2, where we compare the proposed approach in a state-feedback scenario, i.e., $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{I}_2$, and a scenario with output feedback via $\mathbf{C}_1 = [1 \;\; 2]$ and $\mathbf{C}_2 = [2 \;\; 1]$ with the approach from [163]. The state-feedback approaches from [38] and [52] are included for reference. According to Table 3.2, it can be seen that, as expected, state feedback yields lower median costs than output feedback in the considered scenario. The performance of the proposed controller is almost equal to the performance of the time-variant controller from [163] in the steady-state scenario, i.e., for $\underline{\boldsymbol{x}}_0 = [0 \;\; 0]^\top$, and is only slightly worse in the transient scenario where the state has to be driven from $\underline{\boldsymbol{x}}_0 = [3 \;\; 0]^\top$ to the origin. Moreover, we can conclude from Table 3.2 that the choice of the transition matrix has an impact on the costs of all controllers. In output-feedback scenarios, this influence seems to be more severe than for state feedback.

Besides the performance in terms of the costs, it is important to compare the computation time of the regulator gains of the proposed iterative algorithm and the method from [163]. The regulator gain computation time of the proposed algorithm was only 0.025s for the output-feedback scenario with transition matrix $\mathbf{T}_{[1]}$ and 0.03s for $\mathbf{T}_{[2]}$ on a PC with Intel Core i5-3320M and 8 GB RAM running Matlab 2013b. Furthermore, the computation time is independent of the simulation time. On the other hand, the method from [163] took 194.12s for $\mathbf{T}_{[1]}$ and

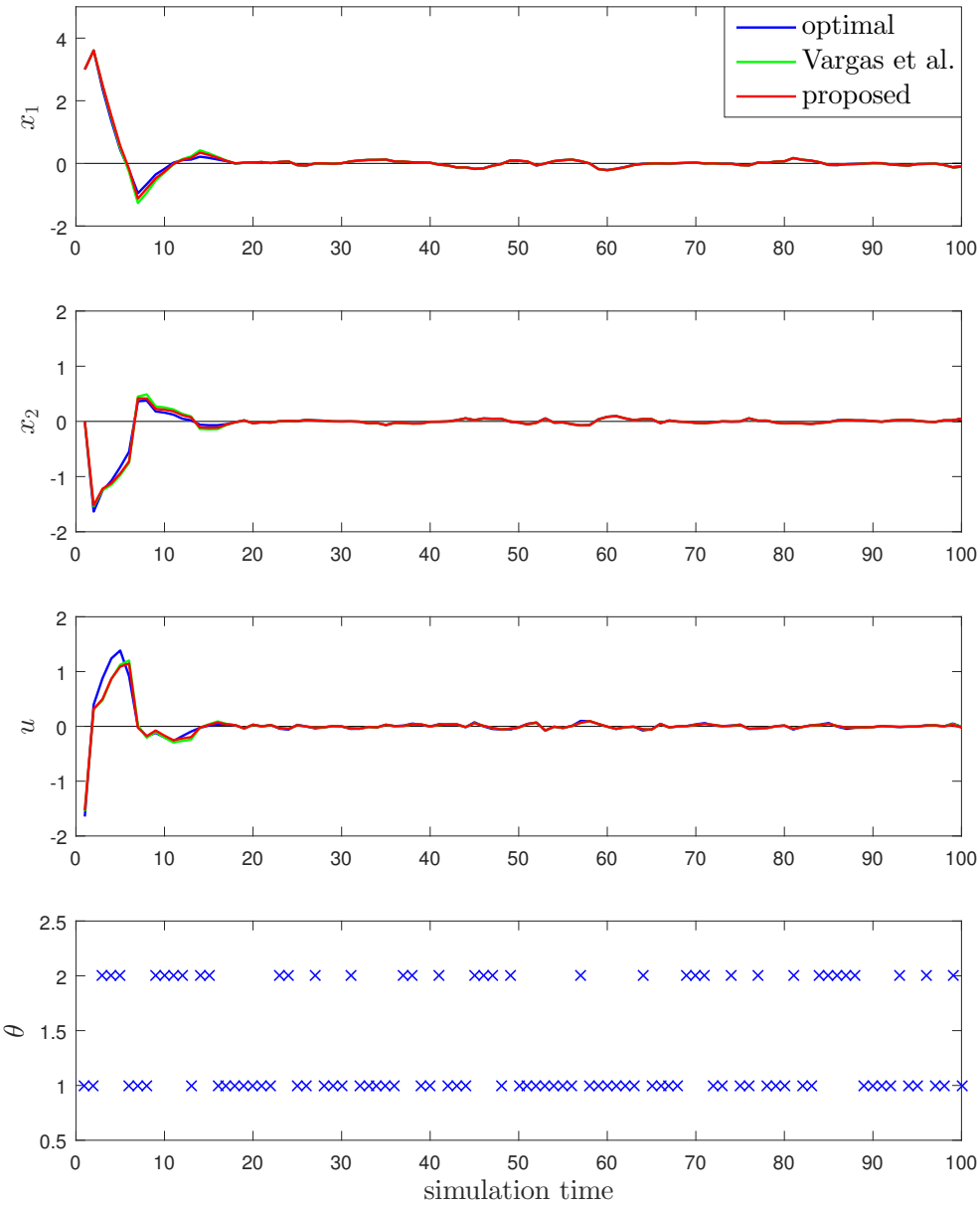


Figure 3.5: State, input, and mode trajectories of an example run with $\mathbf{T}_{[1]}$ and $\underline{\boldsymbol{x}}_0^{[2]}$.

277.68s for $\mathbf{T}_{[2]}$. In order to decrease the computation time of the control approach from [163], we suggest to use the regulator gain computed using the proposed iterative algorithm as the initial gain for the controller from [163].

3.6 Infinite-horizon Dynamic Output Feedback

This section addresses infinite-horizon dynamic output-feedback control of MJLS with non-observed mode. In particular, we propose an algorithm for computation of parameters of a linear time-invariant controller that is robust to the mode. The presented results were partially published in [179] (own publication).

3.6.1 Problem Formulation and Preliminaries

In this section, we consider an MJLS whose dynamics are given by

$$\begin{aligned} \underline{\boldsymbol{x}}_{k+1} &= \mathbf{A}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \mathbf{B}_{\boldsymbol{\theta}_k}\underline{u}_k + \mathbf{H}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{w}}_k \ , \\ \underline{\boldsymbol{y}}_k &= \mathbf{C}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \mathbf{J}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{v}}_k \ , \end{aligned} \tag{3.36}$$

where $\underline{\boldsymbol{x}}_k \in \mathbb{R}^{n_x}$ is the state, $\underline{u}_k \in \mathbb{R}^{n_u}$ the control input, and $\underline{\boldsymbol{y}}_k \in \mathbb{R}^{n_y}$ the measurement that is fed back to the controller. Process disturbances are modeled as a stationary second-order random process $\underline{\boldsymbol{w}}_k \in \mathbb{R}^{n_w}$, while measurement disturbances are modeled with another stationary second-order random process $\underline{\boldsymbol{v}}_k \in \mathbb{R}^{n_v}$. Both random processes are assumed to have zero mean and identity covariance. Further, they are independent from each other and other quantities in (3.36). Please note that the assumption of identity covariance for both processes can be made without loss of generality, because different covariances can be modeled by choosing the matrices $\mathbf{H}_{\boldsymbol{\theta}_k}$ and $\mathbf{J}_{\boldsymbol{\theta}_k}$ appropriately. Finally, we assume that the mode $\boldsymbol{\theta}_k$ is not fed back to the controller.

The goal of this section is to compute the parameters of a time-invariant dynamic output-feedback control law for (3.36) of the form

$$\begin{aligned} \underline{\hat{x}}_{k+1} &= \mathbf{F}\underline{\hat{x}}_k + \mathbf{K}\underline{\boldsymbol{y}}_k \ , \\ \underline{u}_k &= \mathbf{L}\underline{\hat{x}}_k \ , \end{aligned} \tag{3.37}$$

such that it minimizes the infinite-horizon quadratic cost function

$$\mathcal{J} = \lim_{K\to\infty} \frac{1}{K}\mathrm{E}\left\{\sum_{k=0}^{K} \underline{\boldsymbol{x}}_k^\top \mathbf{Q}_{\boldsymbol{\theta}_k}\underline{\boldsymbol{x}}_k + \underline{u}_k^\top \mathbf{R}_{\boldsymbol{\theta}_k}\underline{u}_k\right\} \ , \tag{3.38}$$

where the matrices $\mathbf{Q}_{\boldsymbol{\theta}_k}$ are positive semidefinite and $\mathbf{R}_{\boldsymbol{\theta}_k}$ are positive definite. The expectation is taken with respect to $\underline{\boldsymbol{x}}_k$, $\underline{\boldsymbol{w}}_k$, $\underline{\boldsymbol{v}}_k$, and $\boldsymbol{\theta}_k$. Furthermore, the matrices $\mathbf{F}$, $\mathbf{K}$, and $\mathbf{L}$ shall be independent not only of the mode $\boldsymbol{\theta}_k$ but also of the initial condition $(\underline{\boldsymbol{x}}_0, \boldsymbol{\theta}_0)$.

Similar to the course of Sec. 3.5, we first construct the closed-loop system according to

$$\underline{\widetilde{\boldsymbol{x}}}_{k+1} = \widetilde{\mathbf{A}}_{\boldsymbol{\theta}_k}\underline{\widetilde{\boldsymbol{x}}}_k + \widetilde{\mathbf{H}}_{\boldsymbol{\theta}_k}\underline{\widetilde{\boldsymbol{w}}}_k \tag{3.39}$$

with

$$\widetilde{\underline{\boldsymbol{x}}}_k = \begin{bmatrix} \underline{\boldsymbol{x}}_k \\ \underline{\hat{x}}_k \end{bmatrix} \ , \ \widetilde{\mathbf{A}}_{\boldsymbol{\theta}_k} = \begin{bmatrix} \mathbf{A}_{\boldsymbol{\theta}_k} & \mathbf{B}_{\boldsymbol{\theta}_k}\mathbf{L} \\ \mathbf{K}\mathbf{C}_{\boldsymbol{\theta}_k} & \mathbf{F} \end{bmatrix} \ , \ \widetilde{\underline{\boldsymbol{w}}}_k = \begin{bmatrix} \underline{\boldsymbol{w}}_k \\ \underline{\boldsymbol{v}}_k \end{bmatrix} \ , \ \widetilde{\mathbf{H}}_{\boldsymbol{\theta}_k,k} = \begin{bmatrix} \mathbf{H}_{\boldsymbol{\theta}_k} & \mathbf{0}_{n_w \times n_v} \\ \mathbf{0}_{n_v \times n_w} & \mathbf{K}\mathbf{J}_{\boldsymbol{\theta}_k} \end{bmatrix} \ . \quad (3.40)$$

The next result follows from plugging (3.37) into (3.38) and using the measurement equation from (3.36).

Proposition 3.20. *For the cost function* (3.38) *written in terms of the closed-loop state, it holds*

$$\mathcal{J} = \lim_{K\to\infty} \frac{1}{K} \mathrm{E}\left\{ \sum_{k=0}^{K} \widetilde{\underline{\boldsymbol{x}}}_k^\top \widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_k} \widetilde{\underline{\boldsymbol{x}}}_k \right\} \ ,$$

where

$$\widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_k} = \begin{bmatrix} \mathbf{Q}_{\boldsymbol{\theta}_k} & \mathbf{0}_{n_x} \\ \mathbf{0}_{n_x} & \mathbf{L}^\top \mathbf{R}_{\boldsymbol{\theta}_k} \mathbf{L} \end{bmatrix} \ .$$

Finally, we can summarize the considered problem as follows.

Problem 3.21. *The optimal dynamic output-feedback control law* (3.37) *that minimizes the quadratic cost function* (3.38) *for the MJLS* (3.36) *with non-observed mode is determined by the optimization problem*

$$\begin{aligned} \inf_{\mathbf{F},\mathbf{K},\mathbf{L}} \quad & \lim_{K\to\infty} \frac{1}{K} \mathrm{E}\left\{ \sum_{k=0}^{K} \widetilde{\underline{\boldsymbol{x}}}_k^\top \widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_k} \widetilde{\underline{\boldsymbol{x}}}_k \right\} \\ \text{s. t.} \quad & \widetilde{\underline{\boldsymbol{x}}}_{k+1} = \widetilde{\mathbf{A}}_{\boldsymbol{\theta}_k} \widetilde{\underline{\boldsymbol{x}}}_k + \widetilde{\mathbf{H}}_{\boldsymbol{\theta}_k} \widetilde{\underline{\boldsymbol{w}}}_k \ . \end{aligned} \quad (3.41)$$

The proposed algorithm for this problem is presented in the next section.

3.6.2 Computation of the Control Law

The derivation of the algorithm for computation of the parameters $\mathbf{F}$, $\mathbf{K}$, and $\mathbf{L}$ of the linear dynamic control law (3.37) is similar to the approach from the previous section. First, we reformulate the optimization problem (3.41) in terms of the closed-loop second moment of the MJLS and construct a Hamiltonian using symmetric positive definite Lagrange multipliers. Unfortunately, a naïve evaluation of the necessary optimality conditions for the Hamiltonian does not yield a convergent algorithm. Thus, we propose to upper bound the second moment and the Lagrange multiplier to ensure that they remain positive definite during the proposed iterative computation of the controller parameters. By doing so, we are able to derive an iterative algorithm for computation of the parameters of the assumed control law. Convergence properties of the proposed algorithm constitute a topic for future research.

In what follows, we assume that the MJLS (3.36) is mean square stabilizable via dynamic output feedback without mode observation according to the following definition.

Definition 3.22. *The MJLS* (3.36) *is stabilizable via dynamic output feedback* (3.37)*, if there exist matrices* $\mathbf{F}$*,* $\mathbf{K}$*, and* $\mathbf{L}$ *such that the closed-loop system* (3.39) *is mean square stable.*

The derivation of the algorithm for minimization of (3.41) follows the lines of Sec. 3.5.3. First, we define the closed-loop second moment of (3.39) according to

$$\widetilde{\mathbf{X}}_k^i = \mathrm{E}\left\{\widetilde{\underline{\boldsymbol{x}}}_k\widetilde{\underline{\boldsymbol{x}}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} = \begin{bmatrix} \mathrm{E}\left\{\underline{\boldsymbol{x}}_k\underline{\boldsymbol{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} & \mathrm{E}\left\{\underline{\boldsymbol{x}}_k\underline{\hat{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \\ \mathrm{E}\left\{\underline{\hat{x}}_k\underline{\boldsymbol{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} & \mathrm{E}\left\{\underline{\hat{x}}_k\underline{\hat{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{1,k}^i & \mathbf{X}_{12,k}^i \\ (\mathbf{X}_{12,k}^i)^\top & \mathbf{X}_{2,k}^i \end{bmatrix} .$$

The dynamics of the closed-loop second moment are given in the next proposition.

Proposition 3.23. *For the dynamics of the second moment of* (3.39)*, it holds*

$$\widetilde{\mathbf{X}}_{k+1}^j = \sum_{i=1}^M p_{ij}\left[\widetilde{\mathbf{A}}_i\widetilde{\mathbf{X}}_k^i\widetilde{\mathbf{A}}_i^\top + \mu_k^i\widetilde{\mathbf{H}}_i\widetilde{\mathbf{H}}_i^\top\right] ,$$

where

$$\widetilde{\mathbf{X}}_{k+1}^i = \begin{bmatrix} \mathbf{X}_{1,k+1}^i & \mathbf{X}_{12,k+1}^i \\ (\mathbf{X}_{12,k+1}^i)^\top & \mathbf{X}_{2,k+1}^i \end{bmatrix}$$

with

$$\begin{aligned}
\mathbf{X}_{1,k+1}^j &= \sum_{i=1}^M p_{ij}\Big[\mathbf{A}_i\mathbf{X}_{1,k}^i\mathbf{A}_i^\top + \mathbf{A}_i\mathbf{X}_{12,k}^i\mathbf{L}^\top\mathbf{B}_i^\top + \mathbf{B}_i\mathbf{L}(\mathbf{X}_{12,k}^i)^\top\mathbf{A}_i^\top + \mathbf{B}_i\mathbf{L}\mathbf{X}_{2,k}^i\mathbf{L}^\top\mathbf{B}_i^\top \\
&\qquad +\mu_k^i\mathbf{H}_i\mathbf{H}_i^\top\Big] , \\
\mathbf{X}_{12,k+1}^j &= \sum_{i=1}^M p_{ij}\left[\mathbf{A}_i\mathbf{X}_{1,k}^i\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{A}_i\mathbf{X}_{12,k}^i\mathbf{F}^\top + \mathbf{B}_i\mathbf{L}(\mathbf{X}_{12,k}^i)^\top\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{B}_i\mathbf{L}\mathbf{X}_{2,k}^i\mathbf{F}^\top\right] , \\
\mathbf{X}_{2,k+1}^j &= \sum_{i=1}^M p_{ij}\Big[\mathbf{K}\mathbf{C}_i\mathbf{X}_{1,k}^i\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{K}\mathbf{C}_i\mathbf{X}_{12,k}^i\mathbf{F}^\top + \mathbf{F}(\mathbf{X}_{12,k}^i)^\top\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{F}\mathbf{X}_{2,k}^i\mathbf{F}^\top \\
&\qquad +\mu_k^i\mathbf{K}\mathbf{J}_i\mathbf{J}_i^\top\mathbf{K}^\top\Big] .
\end{aligned}$$

Proof. The proof follows the lines of the proof of Proposition 3.14. □

Using this result, we can also express the cost function (3.38) in terms of the closed-loop second moment. The result is given in the next proposition.

Proposition 3.24. *Under the assumption of a linear time-invariant controller* (3.37)*, the optimization problem* (3.41) *is equivalent to*

$$\begin{aligned}
&\inf_{\mathbf{F},\mathbf{K},\mathbf{L}} \quad \sum_{i=1}^M \mathrm{tr}\left[\widetilde{\mathbf{Q}}_i\widetilde{\mathbf{X}}_\infty^i\right] \\
&\text{s.t.} \quad \widetilde{\mathbf{X}}_\infty^j = \sum_{i=1}^M p_{ij}\left[\widetilde{\mathbf{A}}_i\widetilde{\mathbf{X}}_\infty^i\widetilde{\mathbf{A}}_i^\top + \mu_\infty^i\widetilde{\mathbf{H}}_i\widetilde{\mathbf{H}}_i^\top\right] .
\end{aligned} \tag{3.42}$$

Proof. This result can be obtained by using the identity $\underline{x}^\top \mathbf{A} \underline{x} = \operatorname{tr}\left[\mathbf{A} \underline{x}\underline{x}^\top\right]$, taking the expectation with respect to $\boldsymbol{\theta}_k$, $\underline{\boldsymbol{w}}_k$, and $\underline{\boldsymbol{v}}_k$, and taking the limit $K \to \infty$ in (3.41). □

Now, let $\{\widetilde{\mathbf{P}}^1_\infty, \ldots, \widetilde{\mathbf{P}}^i_\infty\}$ denote a set of Lagrange multipliers that are symmetric positive definite matrices. Then, we can construct the Hamiltonian following the approach in [8]. The result is presented in the following proposition.

Proposition 3.25. *The Hamiltonian of* (3.42) *is given by*

$$\mathcal{H} = \sum_{i=1}^{M} \operatorname{tr}\left[\widetilde{\mathbf{Q}}_i \widetilde{\mathbf{X}}^i_\infty + \mathcal{E}_i(\widetilde{\mathbf{P}}_\infty)\left[\widetilde{\mathbf{A}}_i \widetilde{\mathbf{X}}^i_\infty \widetilde{\mathbf{A}}_i^\top + \mu^i_\infty \widetilde{\mathbf{H}}_i \widetilde{\mathbf{H}}_i^\top\right] - \widetilde{\mathbf{P}}^i_\infty \widetilde{\mathbf{X}}^i_\infty\right]$$

with $\mathcal{E}_i(\mathbf{A})$ as defined in Proposition 3.16.

Proof. The result of Proposition 3.25 can be obtained according to

$$\begin{aligned}
\mathcal{H} &= \sum_{i=1}^{M} \operatorname{tr}\left[\widetilde{\mathbf{Q}}_i \widetilde{\mathbf{X}}^i_\infty\right] - \sum_{i=1}^{M} \operatorname{tr}\left[\widetilde{\mathbf{P}}^i_\infty \mathbf{X}^i_\infty\right] + \sum_{j=1}^{M} \operatorname{tr}\left[\widetilde{\mathbf{P}}^j_\infty \sum_{i=1}^{M} p_{ij}\left[\widetilde{\mathbf{A}}_i \widetilde{\mathbf{X}}^i_\infty \widetilde{\mathbf{A}}_i^\top + \mu^i_\infty \widetilde{\mathbf{H}}_i \widetilde{\mathbf{H}}_i^\top\right]\right] \\
&= \sum_{i=1}^{M} \operatorname{tr}\left[\widetilde{\mathbf{Q}}_i \widetilde{\mathbf{X}}^i_\infty - \widetilde{\mathbf{P}}^i_\infty \mathbf{X}^i_\infty + \sum_{j=1}^{M} p_{ij} \widetilde{\mathbf{P}}^j_\infty \left[\widetilde{\mathbf{A}}_i \widetilde{\mathbf{X}}^i_\infty \widetilde{\mathbf{A}}_i^\top + \mu^i_\infty \widetilde{\mathbf{H}}_i \widetilde{\mathbf{H}}_i^\top\right]\right] \\
&= \sum_{i=1}^{M} \operatorname{tr}\left[\widetilde{\mathbf{Q}}_i \widetilde{\mathbf{X}}^i_\infty - \widetilde{\mathbf{P}}^i_\infty \mathbf{X}^i_\infty + \mathcal{E}_i(\widetilde{\mathbf{P}}_\infty)\left[\widetilde{\mathbf{A}}_i \widetilde{\mathbf{X}}^i_\infty \widetilde{\mathbf{A}}_i^\top + \mu^i_\infty \widetilde{\mathbf{H}}_i \widetilde{\mathbf{H}}_i^\top\right]\right] \ .
\end{aligned}$$

□

Theoretically, we could evaluate the necessary optimality conditions for $\mathcal{H}$ at this point and use these results in order to construct an algorithm for computation of the controller parameters $\mathbf{F}$, $\mathbf{K}$, and $\mathbf{L}$. However, our empirical evaluations have shown that such an algorithm does not converge, because the second moments $\widetilde{\mathbf{X}}^i_\infty$ and the Lagrange multiplier $\widetilde{\mathbf{P}}^i_\infty$ leave the positive definite cone. For this reason, we propose to bound the second moments and the Lagrange multiplier from above according to Proposition 3.26. By doing so, we ensure that they remain within the admissible space of positive definite matrices.

Proposition 3.26. *The second moment $\widetilde{\mathbf{X}}^i_\infty$ and the Lagrange multiplier $\widetilde{\mathbf{P}}^i_\infty$ can be bounded from above according to*

$$\widetilde{\mathbf{X}}^i_\infty \leq \begin{bmatrix} \overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i & \underline{\mathbf{X}}_i \\ \underline{\mathbf{X}}_i & \underline{\mathbf{X}}_i \end{bmatrix} \qquad \textit{and} \qquad \widetilde{\mathbf{P}}^i_\infty \leq \begin{bmatrix} \overline{\mathbf{P}}_i + \underline{\mathbf{P}}_i & -\underline{\mathbf{P}}_i \\ -\underline{\mathbf{P}}_i & \underline{\mathbf{P}}_i \end{bmatrix} \ ,$$

where

$$\begin{aligned}
&\overline{\mathbf{X}}_i = \mathbf{X}^i_{1,\infty} - \mathbf{X}^i_{12,\infty} - (\mathbf{X}^i_{12,\infty})^\top + \mathbf{X}^i_{2,\infty} \ , \\
&\underline{\mathbf{X}}_i = \mathbf{X}^i_{2,\infty} \ ,
\end{aligned} \qquad \textit{and} \qquad \begin{aligned}
&\overline{\mathbf{P}}_i = \mathbf{P}^i_{1,\infty} - \mathbf{P}^i_{12,\infty} - (\mathbf{P}^i_{12,\infty})^\top + \mathbf{P}^i_{2,\infty} \ , \\
&\underline{\mathbf{P}}_i = \mathbf{P}^i_{2,\infty} \ ,
\end{aligned} \tag{3.43}$$

Proof. To show the claim of Proposition 3.26 for $\widetilde{\mathbf{X}}^i_\infty$, we first take the Schur complement of $\widetilde{\mathbf{X}}^i_\infty$ in $\mathbf{X}^i_{2,\infty}$. Then, we use the inequality $\mathbf{X}\mathbf{Y}^{-1}\mathbf{X}^\top \geq \mathbf{X} + \mathbf{X}^\top - \mathbf{Y}$ for positive definite matrices $\mathbf{Y}$ [68] and obtain

$$\mathbf{X}^i_{1,\infty} - \mathbf{X}^i_{12,\infty}(\mathbf{X}^i_{2,\infty})^{-1}(\mathbf{X}^i_{12,\infty})^\top \leq \mathbf{X}^i_{1,\infty} - \mathbf{X}^i_{12,\infty} - (\mathbf{X}^i_{12,\infty})^\top + \mathbf{X}^i_{2,\infty} \ .$$

Next, the usage of substitutions (3.43) yields

$$\mathbf{X}^i_{1,\infty} - \mathbf{X}^i_{12,\infty} - (\mathbf{X}^i_{12,\infty})^\top + \mathbf{X}^i_{2,\infty} = \overline{\mathbf{X}}_i \ .$$

Finally, we get

$$\begin{aligned}\overline{\mathbf{X}}_i &= \overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i - \underline{\mathbf{X}}_i \\ &= \overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i - \underline{\mathbf{X}}_i\underline{\mathbf{X}}_i^{-1}\underline{\mathbf{X}}_i^\top\end{aligned}$$

and reverse the Schur complement. Since the Schur complement preserves the partial order of matrices in the positive semidefinite sense, we obtain the result of Proposition 3.26 for $\widetilde{\mathbf{X}}^i_\infty$. The proof for $\widetilde{\mathbf{P}}^i_\infty$ works analogously. □

Please observe that in classical stochastic optimal control of linear systems, i.e., for $M = 1$, and also in control of systems with white parameters, optimality implies $\mathbf{X}_{12,\infty} = (\mathbf{X}_{12,\infty})^\top = \mathbf{X}_{2,\infty}$ [13, 44, 170, 161]. This convergence can be interpreted as the controller state $\underline{\hat{x}}_k$ being an unbiased estimate of the system state $\underline{\boldsymbol{x}}_k$. Also, the Lagrange multiplier $\widetilde{\mathbf{P}}$ in classical linear stochastic optimal control corresponds to the second moment of the costate. Consequently, optimality implies $-\mathbf{P}_{12,\infty} = -(\mathbf{P}_{12,\infty})^\top = \mathbf{P}_{2,\infty}$ [68], i.e., the controller costate becomes an unbiased estimate of the costate of $\underline{\boldsymbol{x}}_k$. From these observations, we can also conclude

$$\underline{\mathbf{X}}_i = \lim_{k\to\infty} \mathrm{E}\left\{\underline{\hat{x}}_k\underline{\hat{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \text{ and } \overline{\mathbf{X}}_i = \lim_{k\to\infty} \mathrm{E}\left\{(\underline{\boldsymbol{x}}_k - \underline{\hat{x}}_k)(\underline{\boldsymbol{x}}_k - \underline{\hat{x}}_k)^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \ . \tag{3.44}$$

Thus, it follows from (3.43) and (3.44) that the second moment $\mathbf{X}^i_{1,\infty}$ of the system state $\underline{\boldsymbol{x}}_k$ is estimated using the second moment $\underline{\mathbf{X}}_i$ of the controller and the estimation covariance $\overline{\mathbf{X}}_i$. The same property can be observed for $\mathbf{P}^i_{1,\infty}$.

Finally, using substitutions (3.43), we can state the main result of this section in the next theorem.

Theorem 3.27. *The linear dynamic output-feedback control law* (3.37) *for the MJLS* (3.36) *with non-observed mode is determined by the coupled nonlinear equations*

$$\begin{aligned}\overline{\mathbf{X}}_j = \sum_{i=1}^M p_{ij} \Big[&\mu^i_\infty \mathbf{H}_i\mathbf{H}_i^\top + \mu^i_\infty \mathbf{K}\mathbf{J}_i\mathbf{J}_i^\top\mathbf{K}^\top + (\mathbf{A}_i - \mathbf{K}\mathbf{C}_i)\overline{\mathbf{X}}_i(\mathbf{A}_i - \mathbf{K}\mathbf{C}_i)^\top \\ &+(\mathbf{A}_i - \mathbf{F} + \mathbf{B}_i\mathbf{L} - \mathbf{K}\mathbf{C}_i)\underline{\mathbf{X}}_i(\mathbf{A}_i - \mathbf{F} + \mathbf{B}_i\mathbf{L} - \mathbf{K}\mathbf{C}_i)^\top\Big] \ ,\end{aligned} \tag{3.45}$$

$$\underline{\mathbf{X}}_j = \sum_{i=1}^M p_{ij} \left[\mu^i_\infty \mathbf{K}\mathbf{J}_i\mathbf{J}_i^\top\mathbf{K}^\top + \mathbf{K}\mathbf{C}_i\overline{\mathbf{X}}_i\mathbf{C}_i^\top\mathbf{K}^\top + (\mathbf{F} + \mathbf{K}\mathbf{C}_i)\underline{\mathbf{X}}_i(\mathbf{F} + \mathbf{K}\mathbf{C}_i)^\top\right] \ , \tag{3.46}$$

$$\overline{\mathbf{P}}_i = \mathbf{Q}_i + \mathbf{L}^\top \mathbf{R}_i \mathbf{L} + (\mathbf{A}_i + \mathbf{B}_i\mathbf{L})^\top \mathcal{E}_i(\overline{\mathbf{P}})(\mathbf{A}_i + \mathbf{B}_i\mathbf{L}) \\ + (\mathbf{A}_i - \mathbf{F} + \mathbf{B}_i\mathbf{L} - \mathbf{K}\mathbf{C}_i)^\top \mathcal{E}_i(\underline{\mathbf{P}})(\mathbf{A}_i - \mathbf{F} + \mathbf{B}_i\mathbf{L} - \mathbf{K}\mathbf{C}_i) \ , \tag{3.47}$$

$$\underline{\mathbf{P}}_i = \mathbf{L}^\top \mathbf{R}_i \mathbf{L} + \mathbf{L}^\top \mathbf{B}_i^\top \mathcal{E}_i(\overline{\mathbf{P}})\mathbf{B}_i\mathbf{L} + (\mathbf{F} - \mathbf{B}_i\mathbf{L})^\top \mathcal{E}_i(\underline{\mathbf{P}})(\mathbf{F} - \mathbf{B}_i\mathbf{L}) \ , \tag{3.48}$$

$$\begin{bmatrix} \mathbf{\Phi}_F(\mathcal{D}) & \mathbf{\Gamma}_{FK}(\mathcal{D}) & \mathbf{\Gamma}_{FL}(\mathcal{D}) \\ \mathbf{\Gamma}_{FK}(\mathcal{D})^\top & \mathbf{\Phi}_K(\mathcal{D}) & \mathbf{\Gamma}_{KL}(\mathcal{D}) \\ \mathbf{\Gamma}_{FL}(\mathcal{D})^\top & \mathbf{\Gamma}_{KL}(\mathcal{D})^\top & \mathbf{\Phi}_L(\mathcal{D}) \end{bmatrix} \begin{bmatrix} \mathrm{vec}\,[\mathbf{F}] \\ \mathrm{vec}\,[\mathbf{K}] \\ \mathrm{vec}\,[\mathbf{L}] \end{bmatrix} + \begin{bmatrix} \underline{\rho}_F(\mathcal{D}) \\ \underline{\rho}_K(\mathcal{D}) \\ \underline{\rho}_L(\mathcal{D}) \end{bmatrix} = \underline{0} \ , \tag{3.49}$$

where $\mathcal{D} = \{(\overline{\mathbf{X}}_1, \underline{\mathbf{X}}_1, \overline{\mathbf{P}}_1, \underline{\mathbf{P}}_1), \ldots, (\overline{\mathbf{X}}_M, \underline{\mathbf{X}}_M, \overline{\mathbf{P}}_M, \underline{\mathbf{P}}_M)\}$, *and*

$$\begin{aligned}
\mathbf{\Phi}_F(\mathcal{D}) &= \sum_{i=1}^{M} \left[\underline{\mathbf{X}}_i \otimes \mathcal{E}_i(\underline{\mathbf{P}})\right] \ , \\
\mathbf{\Phi}_K(\mathcal{D}) &= \sum_{i=1}^{M} \left[(\mu_\infty^i \mathbf{J}_i\mathbf{J}_i^\top + \mathbf{C}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{C}_i^\top) \otimes \mathcal{E}_i(\underline{\mathbf{P}})\right] \ , \\
\mathbf{\Phi}_L(\mathcal{D}) &= \sum_{i=1}^{M} \left[\underline{\mathbf{X}}_i \otimes (\mathbf{R}_i + \mathbf{B}_i^\top \mathcal{E}_i(\overline{\mathbf{P}} + \underline{\mathbf{P}})\mathbf{B}_i)\right] \ , \\
\mathbf{\Gamma}_{FK}(\mathcal{D}) &= \sum_{i=1}^{M} \left[\underline{\mathbf{X}}_i\mathbf{C}_i^\top \otimes \mathcal{E}_i(\underline{\mathbf{P}})\right] \ , \\
\mathbf{\Gamma}_{FL}(\mathcal{D}) &= -\sum_{i=1}^{M} \left[\underline{\mathbf{X}}_i \otimes \mathcal{E}_i(\underline{\mathbf{P}})\mathbf{B}_i\right] \ , \\
\mathbf{\Gamma}_{KL}(\mathcal{D}) &= -\sum_{i=1}^{M} \left[\mathbf{C}_i\underline{\mathbf{X}}_i \otimes \mathcal{E}_i(\underline{\mathbf{P}})\mathbf{B}_i\right] \ , \\
\underline{\rho}_F(\mathcal{D}) &= -\mathrm{vec}\left[\sum_{i=1}^{M} \mathcal{E}_i(\underline{\mathbf{P}})\mathbf{A}_i\underline{\mathbf{X}}_i\right] \ , \\
\underline{\rho}_K(\mathcal{D}) &= -\mathrm{vec}\left[\sum_{i=1}^{M} \mathcal{E}_i(\underline{\mathbf{P}})\mathbf{A}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{C}_i^\top\right] \ , \\
\underline{\rho}_L(\mathcal{D}) &= \mathrm{vec}\left[\sum_{i=1}^{M} \mathbf{B}_i^\top \mathcal{E}_i(\overline{\mathbf{P}} + \underline{\mathbf{P}})\mathbf{A}_i\underline{\mathbf{X}}_i\right] \ .
\end{aligned} \tag{3.50}$$

Proof. The proof is given in Appendix A.1. □

As argued in Sec. 3.5, computing a solution to the equations (3.47)–(3.49) is computationally demanding. Thus, we propose to use an iterative method as formulated in Algorithm 3.2.

Algorithm 3.2 Computation of the controller parameters for infinite-horizon dynamic output feedback.

- *Step 1:* Set the counter $\eta = 0$ and initialize $\overline{\mathbf{X}}_i^{[\eta]}$, $\underline{\mathbf{X}}_i^{[\eta]}$, $\overline{\mathbf{P}}_i^{[\eta]}$ and $\underline{\mathbf{P}}_i^{[\eta]}$ with random symmetric positive definite $n_x \times n_x$ matrices.
- *Step 2:* Compute $\mathbf{F}_{[\eta]}$, $\mathbf{K}_{[\eta]}$, and $\mathbf{L}_{[\eta]}$ according to

$$\begin{bmatrix} \mathrm{vec}\begin{bmatrix}\mathbf{F}_{[\eta]}\end{bmatrix} \\ \mathrm{vec}\begin{bmatrix}\mathbf{K}_{[\eta]}\end{bmatrix} \\ \mathrm{vec}\begin{bmatrix}\mathbf{L}_{[\eta]}\end{bmatrix} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_{[\eta]}) \end{bmatrix}^{-1} \begin{bmatrix} \underline{\rho}_F(\mathcal{D}_{[\eta]}) \\ \underline{\rho}_K(\mathcal{D}_{[\eta]}) \\ \underline{\rho}_L(\mathcal{D}_{[\eta]}) \end{bmatrix} , \tag{3.51}$$

  where $\mathcal{D}_{[\eta]} = \{(\overline{\mathbf{X}}_1^{[\eta]}, \underline{\mathbf{X}}_1^{[\eta]}, \overline{\mathbf{P}}_1^{[\eta]}, \underline{\mathbf{P}}_1^{[\eta]}), \ldots, (\overline{\mathbf{X}}_M^{[\eta]}, \underline{\mathbf{X}}_M^{[\eta]}, \overline{\mathbf{P}}_M^{[\eta]}, \underline{\mathbf{P}}_M^{[\eta]})\}$, and reverse the vectorization operator.
- *Step 3:* If $\mathbf{F}_{[\eta]} = \mathbf{F}_{[\eta-1]}$, $\mathbf{K}_{[\eta]} = \mathbf{K}_{[\eta-1]}$, and $\mathbf{L}_{[\eta]} = \mathbf{L}_{[\eta-1]}$, stop the algorithm. Otherwise, proceed to *Step 5*.
- *Step 4:* Calculate

$$\begin{aligned}
\overline{\mathbf{X}}_j^{[\eta+1]} &= \sum_{i=1}^{M} p_{ij} \Big[\mu_\infty^i \mathbf{H}_i \mathbf{H}_i^\top + \mu_\infty^i \mathbf{K}_{[\eta]} \mathbf{J}_i \mathbf{J}_i^\top \mathbf{K}_{[\eta]}^\top + (\mathbf{A}_i - \mathbf{K}_{[\eta]}\mathbf{C}_i) \overline{\mathbf{X}}_i^{[\eta]} (\mathbf{A}_i - \mathbf{K}_{[\eta]}\mathbf{C}_i)^\top \\
&\qquad + (\mathbf{A}_i - \mathbf{F}_{[\eta]} + \mathbf{B}_i\mathbf{L}_{[\eta]} - \mathbf{K}_{[\eta]}\mathbf{C}_i) \underline{\mathbf{X}}_i^{[\eta]} (\mathbf{A}_i - \mathbf{F}_{[\eta]} + \mathbf{B}_i\mathbf{L}_{[\eta]} - \mathbf{K}_{[\eta]}\mathbf{C}_i)^\top \Big] , \\
\underline{\mathbf{X}}_j^{[\eta+1]} &= \sum_{i=1}^{M} p_{ij} \Big[\mu_\infty^i \mathbf{K}_{[\eta]} \mathbf{J}_i \mathbf{J}_i^\top \mathbf{K}_{[\eta]}^\top + \mathbf{K}_{[\eta]} \mathbf{C}_i \overline{\mathbf{X}}_i^{[\eta]} \mathbf{C}_i^\top \mathbf{K}_{[\eta]}^\top \\
&\qquad + (\mathbf{F}_{[\eta]} + \mathbf{K}_{[\eta]}\mathbf{C}_i) \underline{\mathbf{X}}_i^{[\eta]} (\mathbf{F}_{[\eta]} + \mathbf{K}_{[\eta]}\mathbf{C}_i)^\top \Big] ,
\end{aligned} \tag{3.52}$$

$$\begin{aligned}
\overline{\mathbf{P}}_j^{[\eta+1]} &= \mathbf{Q}_i + \mathbf{L}_{[\eta]}^\top \mathbf{R}_i \mathbf{L}_{[\eta]} + (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[\eta]})^\top \mathcal{E}_i(\overline{\mathbf{P}}^{[\eta]}) (\mathbf{A}_i + \mathbf{B}_i\mathbf{L}_{[\eta]}) \\
&\quad + (\mathbf{A}_i - \mathbf{F}_{[\eta]} + \mathbf{B}_i\mathbf{L}_{[\eta]} - \mathbf{K}_{[\eta]}\mathbf{C}_i)^\top \mathcal{E}_i(\underline{\mathbf{P}}^{[\eta]}) (\mathbf{A}_i - \mathbf{F}_{[\eta]} + \mathbf{B}_i\mathbf{L}_{[\eta]} - \mathbf{K}_{[\eta]}\mathbf{C}_i) , \\
\underline{\mathbf{P}}_j^{[\eta+1]} &= \mathbf{L}_{[\eta]}^\top \mathbf{R}_i \mathbf{L}_{[\eta]} + \mathbf{L}_{[\eta]}^\top \mathbf{B}_i^\top \mathcal{E}_i(\overline{\mathbf{P}}^{[\eta]}) \mathbf{B}_i \mathbf{L}_{[\eta]} + (\mathbf{F}_{[\eta]} - \mathbf{B}_i\mathbf{L}_{[\eta]})^\top \mathcal{E}_i(\underline{\mathbf{P}}^{[\eta]}) (\mathbf{F}_{[\eta]} - \mathbf{B}_i\mathbf{L}_{[\eta]}) .
\end{aligned} \tag{3.53}$$

- *Step 5:* Set $\eta = \eta + 1$ and go to *Step 2*.

The computation of the inverse in *Step 2* can be computationally very intense and prone to error. For this reason, we recommend to directly solve the SDP

$$\begin{aligned}
&\min \lambda \\
&\text{s. t. } \begin{bmatrix} \lambda \mathbf{I}_{n_x} & \begin{bmatrix} \mathbf{F}_{[\eta]} & \mathbf{K}_{[\eta]} & \mathbf{L}_{[\eta]}^\top \end{bmatrix} \\ \begin{bmatrix} \mathbf{F}_{[\eta]}^\top \\ \mathbf{K}_{[\eta]}^\top \\ \mathbf{L}_{[\eta]} \end{bmatrix} & \mathbf{I}_{n_x+n_y+n_u} \end{bmatrix} > 0 \\
&\begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}\begin{bmatrix} \mathbf{F}_{[\eta]} \end{bmatrix} \\ \operatorname{vec}\begin{bmatrix} \mathbf{K}_{[\eta]} \end{bmatrix} \\ \operatorname{vec}\begin{bmatrix} \mathbf{L}_{[\eta]} \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \underline{\rho}_F(\mathcal{D}_{[\eta]}) \\ \underline{\rho}_K(\mathcal{D}_{[\eta]}) \\ \underline{\rho}_L(\mathcal{D}_{[\eta]}) \end{bmatrix} = \underline{0} \ ,
\end{aligned}$$

using solvers presented, e.g., in [62], [152], or [146].

Whether Algorithm 3.2 converges to the global optimum is still an open question. Furthermore, convergence of Algorithm 3.2 is not sufficient for stability of the closed-loop system. Thus, it is necessary to check stability using one of the methods given in Sec. 3.3.

A numerical comparison of the proposed control approach is given in Sec. 3.7.3 together with the method presented in the next section for finite-horizon dynamic compensation of MJLS without mode observation.

3.7 Finite-Horizon Dynamic Output Feedback

In this section, we extend the results on control of time-invariant MJLS without mode observation from the previous section to control of time-variant MJLS without mode observation. Many concepts from the previous section carry over to the scenario considered here. Thus, only new concepts are described in full length.

3.7.1 Problem Formulation

The dynamics of the MJLS considered in this section are given by

$$\begin{aligned}
\underline{\boldsymbol{x}}_{k+1} &= \mathbf{A}_{\boldsymbol{\theta}_k,k}\underline{\boldsymbol{x}}_k + \mathbf{B}_{\boldsymbol{\theta}_k,k}\underline{u}_k + \mathbf{H}_{\boldsymbol{\theta}_k,k}\underline{\boldsymbol{w}}_k \ , \\
\underline{\boldsymbol{y}}_k &= \mathbf{C}_{\boldsymbol{\theta}_k,k}\underline{\boldsymbol{x}}_k + \mathbf{J}_{\boldsymbol{\theta}_k,k}\underline{\boldsymbol{v}}_k
\end{aligned} \tag{3.54}$$

where, in contrast to time-invariant sets as in (3.36), the system matrices are selected from time-variant sets $\{\mathbf{A}_{i,k}, \mathbf{B}_{i,k}, \mathbf{H}_{i,k}, \mathbf{C}_{i,k}, \mathbf{J}_{i,k}\}$, $i = 1, \ldots, M$. For system (3.54), we will derive an algorithm for computation of the parameters of the time-variant control law

$$\begin{aligned}
\underline{\hat{x}}_{k+1} &= \mathbf{F}_k\underline{\hat{x}}_k + \mathbf{K}_k\underline{\boldsymbol{y}}_k \ , \\
\underline{u}_k &= \mathbf{L}_k\underline{\hat{x}}_k
\end{aligned} \tag{3.55}$$

such that given the initial distribution of the state $\underline{\boldsymbol{x}}_0$ in form of a Gaussian mixture determined by the means $\underline{\xi}_0^i = \mathrm{E}\{\underline{\boldsymbol{x}}_0 \mathbb{1}_{\boldsymbol{\theta}_0=i}\}$ and the covariances $\boldsymbol{\Xi}_0^i = \mathrm{E}\left\{(\underline{\boldsymbol{x}}_0 - \underline{\xi}_0^i)(\underline{\boldsymbol{x}}_0 - \underline{\xi}_0^i)^\top \mathbb{1}_{\boldsymbol{\theta}_0=i}\right\}$, $i \in \{1, 2, \ldots, M\}$, it minimizes the finite-horizon cost function

$$\mathcal{J} = \mathrm{E}\left\{\underline{\boldsymbol{x}}_K^\top \mathbf{Q}_{\boldsymbol{\theta}_K,K} \underline{\boldsymbol{x}}_K + \sum_{k=0}^{K-1} \left[\underline{\boldsymbol{x}}_k^\top \mathbf{Q}_{\boldsymbol{\theta}_k,k} \underline{\boldsymbol{x}}_k + \underline{u}_k^\top \mathbf{R}_{\boldsymbol{\theta}_k,k} \underline{u}_k\right]\right\} , \tag{3.56}$$

where $\mathbf{Q}_{\boldsymbol{\theta}_k,k}$ are positive semidefinite and $\mathbf{R}_{\boldsymbol{\theta}_k,k}$ are positive definite, and $K \in \mathbb{N}$ denotes the length of the optimization horizon. The algorithm for computation of the parameters of (3.55) is derived in the next section.

3.7.2 Computation of the Control Law

In order to derive the proposed iterative algorithm for computation of the parameters of the control law (3.55), we will proceed as follows. First, we rewrite the considered control problem in terms of the second moment of the state of the closed-loop system formed by (3.54)–(3.55). Then, we formulate the costs (3.56) as a recursion that runs from the end of the optimization horizon $k = K$ to $k = 0$. Next, we bound this recursion pointwise from above as proposed in Proposition 3.26. With these prerequisites, we then exploit the necessary optimality conditions and present the proposed algorithm. Finally, we conclude this section by proving pointwise decrease of the costs-to-go in each iteration of our algorithm, which guarantees convergence to a local minimum of the cost function bound. The results presented in this section are adapted from [54] (own publication).

As mentioned above, we begin the derivation of the algorithm for computation of the controller parameters by reformulating the optimization problem described in Sec. 3.7 according to the following proposition.

Proposition 3.28. *The optimization problem*

$$\begin{aligned} \min_{\mathbf{F}_k,\mathbf{K}_k,\mathbf{L}_k} \quad & \mathrm{E}\left\{\underline{\boldsymbol{x}}_K^\top \mathbf{Q}_{\boldsymbol{\theta}_K,K} \underline{\boldsymbol{x}}_K + \sum_{k=0}^{K-1} \left[\underline{\boldsymbol{x}}_k^\top \mathbf{Q}_{\boldsymbol{\theta}_k,k} \underline{\boldsymbol{x}}_k + \underline{u}_k^\top \mathbf{R}_{\boldsymbol{\theta}_k,k} \underline{u}_k\right]\right\} \\ \text{s.t.} \quad & (3.54),\ (3.55),\ \underline{\xi}_0^i,\ \boldsymbol{\Xi}_0^i, \qquad\qquad i \in \{1, 2, \ldots, M\} \end{aligned}$$

is equivalent to

$$\min_{\mathbf{F}_k,\mathbf{K}_k,\mathbf{L}_k} \quad \mathrm{E}\left\{\mathrm{tr}\left[\widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_K,K} \widetilde{\mathbf{X}}_K^{\boldsymbol{\theta}_K} + \sum_{k=0}^{K-1} \widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_k,k} \widetilde{\mathbf{X}}_k^{\boldsymbol{\theta}_k}\right]\right\} \tag{3.57}$$

$$\text{s.t.} \quad \widetilde{\mathbf{X}}_{k+1}^j = \sum_{i=1}^{M} p_{ij,k} \left[\widetilde{\mathbf{A}}_{i,k} \widetilde{\mathbf{X}}_k^i \widetilde{\mathbf{A}}_{i,k}^\top + \mu_k^i \widetilde{\mathbf{H}}_{i,k} \widetilde{\mathbf{H}}_{i,k}^\top\right] , \tag{3.58}$$

$$\underline{\xi}_0^i,\ \boldsymbol{\Xi}_0^i ,$$

where

$$\widetilde{\mathbf{A}}_{\boldsymbol{\theta}_k,k} = \begin{bmatrix} \mathbf{A}_{\boldsymbol{\theta}_k,k} & \mathbf{B}_{\boldsymbol{\theta}_k,k}\mathbf{L}_k \\ \mathbf{K}_k\mathbf{C}_{\boldsymbol{\theta}_k,k} & \mathbf{F}_k \end{bmatrix} \quad , \quad \widetilde{\mathbf{H}}_{\boldsymbol{\theta}_k,k} = \begin{bmatrix} \mathbf{H}_{\boldsymbol{\theta}_k,k} & \mathbf{0}_{n_w \times n_v} \\ \mathbf{0}_{n_v \times n_w} & \mathbf{K}_k\mathbf{J}_{\boldsymbol{\theta}_k,k} \end{bmatrix} , \tag{3.59}$$

and

$$\widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_K,K} = \begin{bmatrix} \mathbf{Q}_{\boldsymbol{\theta}_K,K} & \mathbf{0}_{n_x} \\ \mathbf{0}_{n_x} & \mathbf{0}_{n_x} \end{bmatrix} \quad , \quad \widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_k,k} = \begin{bmatrix} \mathbf{Q}_{\boldsymbol{\theta}_k,k} & \mathbf{0}_{n_x} \\ \mathbf{0}_{n_x} & \mathbf{L}_k^\top \mathbf{R}_{\boldsymbol{\theta}_k,k} \mathbf{L}_k \end{bmatrix} \ . \tag{3.60}$$

Proof. As in the time-invariant problem considered in Sec. 3.6, we begin by constructing the closed-loop dynamics according to

$$\widetilde{\underline{\boldsymbol{x}}}_{k+1} = \widetilde{\mathbf{A}}_{\boldsymbol{\theta}_k,k} \widetilde{\underline{\boldsymbol{x}}}_k + \widetilde{\mathbf{H}}_{\boldsymbol{\theta}_k,k} \widetilde{\underline{\boldsymbol{w}}}_k \ , \tag{3.61}$$

where $\widetilde{\mathbf{A}}_{\boldsymbol{\theta}_k,k}$ and $\widetilde{\mathbf{H}}_{\boldsymbol{\theta}_k,k}$ are as given in (3.59). By introducing the closed-loop second moment $\widetilde{\mathbf{X}}_k^i$ as in (3.28), we obtain the dynamics (3.58). Next, by using (3.55) in (3.56), we can rewrite (3.56) in terms of the augmented state $\widetilde{\underline{\boldsymbol{x}}}_k$ as

$$\mathcal{J} = \mathrm{E}\left\{ \widetilde{\underline{\boldsymbol{x}}}_K^\top \widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_K,K} \widetilde{\underline{\boldsymbol{x}}}_K + \sum_{k=0}^{K-1} \widetilde{\underline{\boldsymbol{x}}}_k^\top \widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_k,k} \widetilde{\underline{\boldsymbol{x}}}_k \right\} \ ,$$

and using the identity $\mathrm{tr}\left[\mathbf{Q}\underline{x}\underline{x}^\top\right] = \underline{x}^\top \mathbf{Q} \underline{x}$, we obtain the cost function

$$\mathcal{J} = \mathrm{E}\left\{ \mathrm{tr}\left[\widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_K,K} \widetilde{\mathbf{X}}_K^{\boldsymbol{\theta}_K} + \sum_{k=0}^{K-1} \widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_k,k} \widetilde{\mathbf{X}}_k^{\boldsymbol{\theta}_k} \right] \right\} \ , \tag{3.62}$$

which concludes the proof. □

The next lemma allows us to formulate the cost function (3.62) as a recursion.

Lemma 3.29. *Given the controller parameters* $(\mathbf{F}_t, \mathbf{K}_t, \mathbf{L}_t)$ *for* $t = 0, 1, \ldots, K-1$*, it holds*

$$\begin{aligned} \mathcal{J}_t &= \mathrm{E}\left\{ \mathrm{tr}\left[\widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_K,k} \widetilde{\mathbf{X}}_K^{\boldsymbol{\theta}_K} + \sum_{k=t}^{K-1} \widetilde{\mathbf{Q}}_{\boldsymbol{\theta}_k,k} \widetilde{\mathbf{X}}_k^{\boldsymbol{\theta}_k} \right] \right\} \\ &= \sum_{i=1}^{M} \mathrm{tr}\left[\widetilde{\mathbf{P}}_t^i \widetilde{\mathbf{X}}_t^i \right] + \mu_t^i \omega_{i,t} \ , \end{aligned} \tag{3.63}$$

where

$$\begin{aligned} \widetilde{\mathbf{P}}_K^i &= \widetilde{\mathbf{Q}}_{i,K} \ , & \omega_{i,K} &= 0 \ , \\ \widetilde{\mathbf{P}}_k^i &= \widetilde{\mathbf{Q}}_{i,k} + \widetilde{\mathbf{A}}_{i,k}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{k+1}) \widetilde{\mathbf{A}}_{i,k} \ , & \omega_{i,k} &= \mathrm{tr}\left[\widetilde{\mathbf{H}}_{i,k}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{k+1}) \widetilde{\mathbf{H}}_{i,k} \right] + \mathcal{E}_i(\omega_{,k+1}) \ . \end{aligned} \tag{3.64}$$

Proof. We will show the result of Lemma 3.29 by induction. At time step K, we have

$$\begin{aligned} \mathcal{J}_K &= \sum_{i=1}^{M} \mathrm{tr}\left[\widetilde{\mathbf{Q}}_{i,K} \widetilde{\mathbf{X}}_K^i \right] \\ &= \sum_{i=1}^{M} \mathrm{tr}\left[\widetilde{\mathbf{P}}_K^i \widetilde{\mathbf{X}}_K^i \right] \ . \end{aligned}$$

Now, assume that (3.29) holds for some $k+1 \in \{1, ..., K-1\}$. Then, we obtain

$$\begin{aligned}
\mathcal{J}_k &= \sum_{i=1}^{M} \left[\operatorname{tr}\left[\widetilde{\mathbf{Q}}_{i,k}\widetilde{\mathbf{X}}_k^i\right] + \sum_{j=1}^{M} \left(\operatorname{tr}\left[\widetilde{\mathbf{P}}_{k+1}^j \widetilde{\mathbf{X}}_{k+1}^j\right] + \mu_{k+1}^j \omega_{j,k+1} \right) \right] \\
&= \sum_{i=1}^{M} \left[\operatorname{tr}\left[\widetilde{\mathbf{Q}}_{i,k}\widetilde{\mathbf{X}}_k^i\right] + \sum_{j=1}^{M} p_{ij} \left(\widetilde{\mathbf{P}}_{k+1}^j \left[\widetilde{\mathbf{A}}_{i,k}\widetilde{\mathbf{X}}_k^i \widetilde{\mathbf{A}}_{i,k}^\top + \mu_k^i \widetilde{\mathbf{H}}_{i,k}\widetilde{\mathbf{H}}_{i,k}^\top \right] + \mu_k^i \omega_{j,k+1} \right) \right] \\
&= \sum_{i=1}^{M} \operatorname{tr}\left[(\widetilde{\mathbf{Q}}_{i,k} + \widetilde{\mathbf{A}}_{i,k}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{k+1})\widetilde{\mathbf{A}}_{i,k})\widetilde{\mathbf{X}}_k^i \right] + \mu_k^i \left[\operatorname{tr}\left[\widetilde{\mathbf{H}}_{i,k}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{k+1})\widetilde{\mathbf{H}}_{i,k} \right] + \mathcal{E}_i(\omega_{,k+1}) \right] \\
&= \sum_{i=1}^{M} \operatorname{tr}\left[\widetilde{\mathbf{P}}_k^i \widetilde{\mathbf{X}}_k^i\right] + \mu_k^i \omega_{i,k} \ .
\end{aligned}$$

□

Our empirical evaluations have shown that a naïve minimization of (3.62) with respect to the controller parameters $(\mathbf{F}_k, \mathbf{K}_k, \mathbf{L}_k)$ for $k = 0, \ldots, K-1$ does not converge. For this reason, we propose a method that relies on an upper bound of the costs-to-go (3.29). In order to be able to formulate the proposed bound, we introduce the same substitutions as in Proposition 3.26 and use the following lemma.

Lemma 3.30. *Let* $\mathbf{X} \geq \mathbf{A}$ *and* $\mathbf{Y} \geq \mathbf{B}$ *be positive semidefinite matrices. Then,*

$$\operatorname{tr}[\mathbf{AB}] \leq \operatorname{tr}[\mathbf{XY}] \ .$$

Proof. Define $\mathbf{Q} = \mathbf{X} - \mathbf{A}$ and $\mathbf{R} = \mathbf{Y} - \mathbf{B}$. The matrices $\mathbf{Q}$ and $\mathbf{R}$ are positive semidefinite. Thus, we obtain

$$\operatorname{tr}[\mathbf{XY}] = \operatorname{tr}[\mathbf{AB}] + \operatorname{tr}[\mathbf{AR} + \mathbf{QB} + \mathbf{QR}] \geq \operatorname{tr}[\mathbf{AB}] \ .$$

□

Finally, we can state the main theorem of this section as follows.

Theorem 3.31. *For fixed controller parameters* $(\mathbf{F}_k, \mathbf{K}_k, \mathbf{L}_k)$ *with* $k = 0, 1, \ldots, K-1$*, the costs* (3.57) *at time step* k *are bounded from above according to*

$$\mathcal{J}_k \leq \overline{\mathcal{J}}_k = \sum_{i=1}^{M} \operatorname{tr}\left[\overline{\mathbf{P}}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) + \underline{\mathbf{P}}_{i,k}\overline{\mathbf{X}}_{i,k}\right] + \mu_k^i \underline{\omega}_{i,k} \ , \tag{3.65}$$

where $\overline{\mathbf{P}}_{i,k}$*,* $\underline{\mathbf{P}}_{i,k}$*,* $\overline{\mathbf{X}}_{i,k}$*, and* $\underline{\mathbf{X}}_{i,k}$ *follow the dynamics*

$$\begin{aligned}
\overline{\mathbf{X}}_{j,k+1} = \sum_{i=1}^{M} p_{ij} \Big[&\mu_k^i \mathbf{H}_{i,k}\mathbf{H}_{i,k}^\top + \mu_k^i \mathbf{K}_k \mathbf{J}_{i,k}\mathbf{J}_{i,k}^\top \mathbf{K}_k^\top + (\mathbf{A}_{i,k} - \mathbf{K}_k\mathbf{C}_{i,k})\overline{\mathbf{X}}_{i,k}(\mathbf{A}_{i,k} - \mathbf{K}_k\mathbf{C}_{i,k})^\top \\
&+ (\mathbf{A}_{i,k} - \mathbf{F}_k + \mathbf{B}_{i,k}\mathbf{L}_k - \mathbf{K}_k\mathbf{C}_{i,k})\underline{\mathbf{X}}_{i,k}(\mathbf{A}_{i,k} - \mathbf{F}_k + \mathbf{B}_{i,k}\mathbf{L}_k - \mathbf{K}_k\mathbf{C}_{i,k})^\top \Big] \ ,
\end{aligned} \tag{3.66}$$

$$\underline{\mathbf{X}}_{j,k+1} = \sum_{i=1}^{M} p_{ij} \left[\mu_k^i \mathbf{K}_k \mathbf{J}_{i,k} \mathbf{J}_{i,k}^\top \mathbf{K}_k^\top + \mathbf{K}_k \mathbf{C}_{i,k} \overline{\mathbf{X}}_{i,k} \mathbf{C}_{i,k}^\top \mathbf{K}_k^\top + (\mathbf{F}_k + \mathbf{K}_k \mathbf{C}_{i,k}) \underline{\mathbf{X}}_{i,k} (\mathbf{F}_k + \mathbf{K}_k \mathbf{C}_{i,k})^\top \right] , \tag{3.67}$$

$$\begin{aligned} \overline{\mathbf{P}}_{i,k} = {} & \mathbf{Q}_{i,k} + \mathbf{L}_k^\top \mathbf{R}_{i,k} \mathbf{L}_k + (\mathbf{A}_{i,k} + \mathbf{B}_{i,k} \mathbf{L}_k)^\top \mathcal{E}_i(\overline{\mathbf{P}}_{k+1}) (\mathbf{A}_{i,k} + \mathbf{B}_{i,k} \mathbf{L}_k) \\ & + (\mathbf{A}_{i,k} - \mathbf{F}_k + \mathbf{B}_{i,k} \mathbf{L}_k - \mathbf{K}_k \mathbf{C}_{i,k})^\top \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) (\mathbf{A}_{i,k} - \mathbf{F}_k + \mathbf{B}_{i,k} \mathbf{L}_k - \mathbf{K}_k \mathbf{C}_{i,k}) , \end{aligned} \tag{3.68}$$

$$\underline{\mathbf{P}}_{i,k} = \mathbf{L}_k^\top \mathbf{R}_{i,k} \mathbf{L}_k + \mathbf{L}_k^\top \mathbf{B}_{i,k}^\top \mathcal{E}_i(\overline{\mathbf{P}}_{k+1}) \mathbf{B}_{i,k} \mathbf{L}_k + (\mathbf{F}_k - \mathbf{B}_{i,k} \mathbf{L}_k)^\top \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) (\mathbf{F}_k - \mathbf{B}_{i,k} \mathbf{L}_k) , \tag{3.69}$$

initialized with $\overline{\mathbf{P}}_{i,K} = \mathbf{Q}_{i,K}$, $\underline{\mathbf{P}}_{i,K} = \mathbf{0}_{n_x}$, $\overline{\mathbf{X}}_{i,0} = \mathbf{\Xi}_0^i$, *and* $\underline{\mathbf{X}}_{i,0} = \mathbf{\Xi}_0^i + \underline{\xi}_0^i (\underline{\xi}_0^i)^\top$, *and*

$$\underline{\omega}_{i,k} = \operatorname{tr} \left[\mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1}) \mathbf{H}_{i,k} \mathbf{H}_{i,k}^\top + \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{K}_k \mathbf{J}_{i,k} \mathbf{J}_{i,k}^\top \mathbf{K}_k^\top \right] + \mathcal{E}_i(\underline{\omega}_{,k+1}) .$$

Proof. The proof is given in Appendix B.1. □

At this point, we can state the following theorem that yields conditions for controller parameters $(\mathbf{F}_k, \mathbf{K}_k, \mathbf{L}_k)$, $k = 0, \ldots, K-1$, which are optimal w.r.t. the upper bound (3.65).

Theorem 3.32. *Suppose that the cost function* (3.65) *attains a minimum for a given sequence* $(\mathbf{F}_k, \mathbf{K}_k, \mathbf{L}_k)$ *with* $k = 0, 1, \ldots, K-1$. *Then, for each* $k = 0, 1, \ldots, K-1$, $(\mathbf{F}_k, \mathbf{K}_k, \mathbf{L}_k)$ *satisfy the necessary optimality conditions*

$$\begin{bmatrix} \mathbf{\Phi}_F(\mathcal{D}_k) & \mathbf{\Gamma}_{FK}(\mathcal{D}_k) & \mathbf{\Gamma}_{FL}(\mathcal{D}_k) \\ \mathbf{\Gamma}_{FK}(\mathcal{D}_k)^\top & \mathbf{\Phi}_K(\mathcal{D}_k) & \mathbf{\Gamma}_{KL}(\mathcal{D}_k) \\ \mathbf{\Gamma}_{FL}(\mathcal{D}_k)^\top & \mathbf{\Gamma}_{KL}(\mathcal{D}_k)^\top & \mathbf{\Phi}_L(\mathcal{D}_k) \end{bmatrix} \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] \\ \operatorname{vec}[\mathbf{K}_k] \\ \operatorname{vec}[\mathbf{L}_k] \end{bmatrix} + \begin{bmatrix} \underline{\rho}_F(\mathcal{D}_k) \\ \underline{\rho}_K(\mathcal{D}_k) \\ \underline{\rho}_L(\mathcal{D}_k) \end{bmatrix} = \underline{0} , \tag{3.70}$$

where $\mathcal{D}_k = \{(\overline{\mathbf{X}}_{1,k}, \underline{\mathbf{X}}_{1,k}, \overline{\mathbf{P}}_{1,k+1}, \underline{\mathbf{P}}_{1,k+1}), \ldots, (\overline{\mathbf{X}}_{M,k}, \underline{\mathbf{X}}_{M,k}, \overline{\mathbf{P}}_{M,k+1}, \underline{\mathbf{P}}_{M,k+1})\}$ *and*

$$\begin{aligned} \mathbf{\Phi}_F(\mathcal{D}_k) &= \sum_{i=1}^{M} \left[\underline{\mathbf{X}}_{i,k} \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \right] , \\ \mathbf{\Phi}_K(\mathcal{D}_k) &= \sum_{i=1}^{M} \left[(\mu_k^i \mathbf{J}_{i,k} \mathbf{J}_{i,k}^\top + \mathbf{C}_{i,k} (\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) \mathbf{C}_{i,k}^\top) \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \right] , \\ \mathbf{\Phi}_L(\mathcal{D}_k) &= \sum_{i=1}^{M} \left[\underline{\mathbf{X}}_{i,k} \otimes (\mathbf{R}_{i,k} + \mathbf{B}_{i,k}^\top \mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1}) \mathbf{B}_{i,k}) \right] , \\ \mathbf{\Gamma}_{FK}(\mathcal{D}_k) &= \sum_{i=1}^{M} \left[\underline{\mathbf{X}}_{i,k} \mathbf{C}_{i,k}^\top \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \right] , \\ \mathbf{\Gamma}_{FL}(\mathcal{D}_k) &= -\sum_{i=1}^{M} \left[\underline{\mathbf{X}}_{i,k} \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{B}_{i,k} \right] , \\ \mathbf{\Gamma}_{KL}(\mathcal{D}_k) &= -\sum_{i=1}^{M} \left[\mathbf{C}_{i,k} \underline{\mathbf{X}}_{i,k} \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{B}_{i,k} \right] , \end{aligned}$$

$$\underline{\rho}_F(\mathcal{D}_k) = -\mathrm{vec}\left[\sum_{i=1}^{M} \mathcal{E}_i(\underline{\mathbf{P}}_{k+1})\mathbf{A}_{i,k}\underline{\mathbf{X}}_{i,k}\right] ,$$
$$\underline{\rho}_K(\mathcal{D}_k) = -\mathrm{vec}\left[\sum_{i=1}^{M} \mathcal{E}_i(\underline{\mathbf{P}}_{k+1})\mathbf{A}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k})\mathbf{C}_{i,k}^\top\right] ,$$
$$\underline{\rho}_L(\mathcal{D}_k) = \mathrm{vec}\left[\sum_{i=1}^{M} \mathbf{B}_{i,k}^\top\mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1})\mathbf{A}_{i,k}\underline{\mathbf{X}}_{i,k}\right] .$$

Proof. The proof is given in Appendix B.2. □

In what follows, we present a numerical method for computation of the sequence of controller gains $(\mathbf{F}_{0:K-1}, \mathbf{K}_{0:K-1}, \mathbf{L}_{0:K-1})$ in Algorithm 3.3 using the results of Theorem 3.32. We employ a variational approach that generates a sequence of monotonically decreasing cost functions $\overline{\mathcal{J}}_0^{[\eta]}$. The fixed point of this sequence yields a sequence $(\mathbf{F}_{0:K-1}^{[\infty]}, \mathbf{K}_{0:K-1}^{[\infty]}, \mathbf{L}_{0:K-1}^{[\infty]})$ that minimizes the cost function $\overline{\mathcal{J}}_0$. The proposed approach can be seen as a DDP method [118] with a control law assumption (see Sec. 2.5).

As argued in Sec. 3.6.2, computing $\mathbf{F}_k^{[\eta]}$, $\mathbf{K}_k^{[\eta]}$, and $\mathbf{L}_k^{[\eta]}$ in *Step 4* of Algorithm 3.3 may be numerically instable due to the matrix inversion. In this case, we propose to use the quadratic program

$$\begin{aligned}
\min \quad & \begin{bmatrix} \mathrm{vec}\left[\mathbf{F}_k^{[\eta]}\right] \\ \mathrm{vec}\left[\mathbf{K}_k^{[\eta]}\right] \\ \mathrm{vec}\left[\mathbf{L}_k^{[\eta]}\right] \end{bmatrix}^\top \begin{bmatrix} \mathrm{vec}\left[\mathbf{F}_k^{[\eta]}\right] \\ \mathrm{vec}\left[\mathbf{K}_k^{[\eta]}\right] \\ \mathrm{vec}\left[\mathbf{L}_k^{[\eta]}\right] \end{bmatrix} \\
\text{s. t.} \quad & \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \mathrm{vec}\left[\mathbf{F}_k^{[\eta]}\right] \\ \mathrm{vec}\left[\mathbf{K}_k^{[\eta]}\right] \\ \mathrm{vec}\left[\mathbf{L}_k^{[\eta]}\right] \end{bmatrix} + \begin{bmatrix} \underline{\rho}_F(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_K(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} = \underline{0} ,
\end{aligned} \tag{3.73}$$

where $\mathcal{D}_k^{[\eta]} = \{(\overline{\mathbf{X}}_{1,k}^{[\eta]}, \underline{\mathbf{X}}_{1,k}^{[\eta]}, \overline{\mathbf{P}}_{1,k+1}^{[\eta]}, \underline{\mathbf{P}}_{1,k+1}^{[\eta]}), \ldots, (\overline{\mathbf{X}}_{M,k}^{[\eta]}, \underline{\mathbf{X}}_{M,k}^{[\eta]}, \overline{\mathbf{P}}_{M,k+1}^{[\eta]}, \underline{\mathbf{P}}_{M,k+1}^{[\eta]})\}$. The quadratic program (3.73) can be solved using solvers presented, e.g., in [62], [146], and [152].

Also, the above algorithm can be implemented in a receding horizon framework, where the optimization horizon is usually much smaller than the operation horizon. This is especially beneficial for long operation horizons. In the receding horizon framework, the controller computes at each time step k of the operation horizon a sequence of controller parameters $(\mathbf{F}_{k:k+K}, \mathbf{K}_{k:k+K}, \mathbf{L}_{k:k+K})$ and uses this sequence only until it generates a new sequence $(\mathbf{F}_{k+t:k+t+K}, \mathbf{K}_{k+t:k+t+K}, \mathbf{L}_{k+t:k+t+K})$ at time step $k+t$, $t \in \{1, 2, ..., K\}$. In this framework, t can either be fixed or change dynamically depending on when additional information becomes available to the controller. Such information could be, e.g., a change of the operating point to some point $\underline{\zeta} \in \mathbb{R}^{n_x}$ which induces a shift of the origin according to the coordinate transform $\underline{\boldsymbol{x}}_k - \underline{\zeta}$.

Algorithm 3.3 Computation of the controller parameters for finite-horizon dynamic output feedback.

- *Step 1:* Set the iteration counter $\eta = 0$ and select an arbitrary sequence $(\mathbf{F}_{0:K-1}^{[0]}, \mathbf{K}_{0:K-1}^{[0]}, \mathbf{L}_{0:K-1}^{[0]})$. Furthermore, set the initial costs to $\overline{\mathcal{J}}_0^{[0]} = \infty$.
- *Step 2:* For $k = 0, 1, \ldots, K$ compute $\overline{\mathbf{X}}_{i,k}^{[\eta]}$ and $\underline{\mathbf{X}}_{i,k}^{[\eta]}$ using

$$\begin{aligned}
\overline{\mathbf{X}}_{j,k+1}^{[\eta]} &= \sum_{i=1}^{M} p_{ij} \Big[\mu_k^i \mathbf{H}_{i,k} \mathbf{H}_{i,k}^\top + \mu_k^i \mathbf{K}_k^{[\eta]} \mathbf{J}_{i,k} \mathbf{J}_{i,k}^\top (\mathbf{K}_k^{[\eta]})^\top + (\mathbf{A}_{i,k} - \mathbf{F}_k^{[\eta]} + \mathbf{B}_{i,k}\mathbf{L}_k^{[\eta]} - \mathbf{K}_k^{[\eta]}\mathbf{C}_{i,k}) \\
&\quad \times \underline{\mathbf{X}}_{i,k}^{[\eta]} (\mathbf{A}_{i,k} - \mathbf{F}_k^{[\eta]} + \mathbf{B}_{i,k}\mathbf{L}_k^{[\eta]} - \mathbf{K}_k^{[\eta]}\mathbf{C}_{i,k})^\top + (\mathbf{A}_{i,k} - \mathbf{K}_k^{[\eta]}\mathbf{C}_{i,k}) \overline{\mathbf{X}}_{i,k}^{[\eta]} (\mathbf{A}_{i,k} - \mathbf{K}_k^{[\eta]}\mathbf{C}_{i,k})^\top \Big] , \\
\underline{\mathbf{X}}_{j,k+1}^{[\eta]} &= \sum_{i=1}^{M} p_{ij} \Big[\mu_k^i \mathbf{K}_k^{[\eta]} \mathbf{J}_{i,k} \mathbf{J}_{i,k}^\top (\mathbf{K}_k^{[\eta]})^\top + \mathbf{K}_k^{[\eta]} \mathbf{C}_{i,k} \overline{\mathbf{X}}_{i,k}^{[\eta]} \mathbf{C}_{i,k}^\top (\mathbf{K}_k^{[\eta]})^\top \\
&\quad + (\mathbf{F}_k^{[\eta]} + \mathbf{K}_k^{[\eta]}\mathbf{C}_{i,k}) \underline{\mathbf{X}}_{i,k}^{[\eta]} (\mathbf{F}_k^{[\eta]} + \mathbf{K}_k^{[\eta]}\mathbf{C}_{i,k})^\top \Big] ,
\end{aligned}$$

  where $\overline{\mathbf{X}}_{i,0}^{[\eta]}$ and $\underline{\mathbf{X}}_{i,0}^{[\eta]}$ are initialized with Ξ_0^i and $\underline{\xi}_0^i (\underline{\xi}_0^i)^\top$, respectively.
- *Step 3:* Set $\overline{\mathbf{P}}_{i,K}^{[\eta]} = \mathbf{Q}_{i,K}$, $\underline{\mathbf{P}}_{i,K}^{[\eta]} = \mathbf{0}_{n_x}$, and $\underline{\omega}_{i,K}^{[\eta]} = 0$. Set the time counter $k = K - 1$.
- *Step 4:* Compute $\bar{\mathbf{F}}_k, \bar{\mathbf{K}}_k, \bar{\mathbf{L}}_k$ by solving the linear program

$$\begin{bmatrix} \mathbf{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \mathbf{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \mathbf{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \mathbf{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \mathbf{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \mathbf{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \mathbf{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \mathbf{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \mathbf{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}\begin{bmatrix} \bar{\mathbf{F}}_k \end{bmatrix} \\ \operatorname{vec}\begin{bmatrix} \bar{\mathbf{K}}_k \end{bmatrix} \\ \operatorname{vec}\begin{bmatrix} \bar{\mathbf{L}}_k \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \underline{\rho}_F(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_K(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} = \underline{0} , \tag{3.71}$$

  where $\mathcal{D}_k^{[\eta]} = \{(\overline{\mathbf{X}}_{1,k}^{[\eta]}, \underline{\mathbf{X}}_{1,k}^{[\eta]}, \overline{\mathbf{P}}_{1,k+1}^{[\eta]}, \underline{\mathbf{P}}_{1,k+1}^{[\eta]}), \ldots, (\overline{\mathbf{X}}_{M,k}^{[\eta]}, \underline{\mathbf{X}}_{M,k}^{[\eta]}, \overline{\mathbf{P}}_{M,k+1}^{[\eta]}, \underline{\mathbf{P}}_{M,k+1}^{[\eta]})\}$, and taking the inverse of the vectorization operator.
- *Step 5:* For $\eta \geq 1$, evaluate $\overline{\mathcal{J}}_k^{[\eta]}$ according to

$$\overline{\mathcal{J}}_k^{[\eta]} = \sum_{i=1}^{M} \operatorname{tr}\left[\overline{\mathbf{P}}_{i,k}^{[\eta]} (\overline{\mathbf{X}}_{i,k}^{[\eta]} + \underline{\mathbf{X}}_{i,k}^{[\eta]}) + \underline{\mathbf{P}}_{i,k}^{[\eta]} \overline{\mathbf{X}}_{i,k}^{[\eta]} \right] + \mu_k^i \underline{\omega}_{i,k} . \tag{3.72}$$

  If $\overline{\mathcal{J}}_k^{[\eta]} \leq \overline{\mathcal{J}}_k^{[\eta-1]}$, set $\mathbf{F}_k^{[\eta]} = \bar{\mathbf{F}}_k$, $\mathbf{K}_k^{[\eta]} = \bar{\mathbf{K}}_k$, and $\mathbf{L}_k^{[\eta]} = \bar{\mathbf{L}}_k$. Otherwise set $\mathbf{F}_k^{[\eta]} = \mathbf{F}_k^{[\eta-1]}$, $\mathbf{K}_k^{[\eta]} = \mathbf{K}_k^{[\eta-1]}$, and $\mathbf{L}_k^{[\eta]} = \mathbf{L}_k^{[\eta-1]}$.
- *Step 6:* Compute the time update

$$\begin{aligned}
\overline{\mathbf{P}}_{i,k}^{[\eta]} &= \mathbf{Q}_{i,k} + (\mathbf{L}_k^{[\eta]})^\top \mathbf{R}_{i,k} \mathbf{L}_k^{[\eta]} + (\mathbf{A}_{i,k} + \mathbf{B}_{i,k}\mathbf{L}_k^{[\eta]})^\top \mathcal{E}_i(\overline{\mathbf{P}}_{k+1}^{[\eta]}) (\mathbf{A}_{i,k} + \mathbf{B}_{i,k}\mathbf{L}_k^{[\eta]}) \\
&\quad + (\mathbf{A}_{i,k} - \mathbf{F}_k^{[\eta]} + \mathbf{B}_{i,k}\mathbf{L}_k^{[\eta]} - \mathbf{K}_k^{[\eta]}\mathbf{C}_{i,k})^\top \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}^{[\eta]}) (\mathbf{A}_{i,k} - \mathbf{F}_k^{[\eta]} + \mathbf{B}_{i,k}\mathbf{L}_k^{[\eta]} - \mathbf{K}_k^{[\eta]}\mathbf{C}_{i,k}) , \\
\underline{\mathbf{P}}_{i,k}^{[\eta]} &= (\mathbf{L}_k^{[\eta]})^\top \mathbf{R}_{i,k} \mathbf{L}_k^{[\eta]} + (\mathbf{L}_k^{[\eta]})^\top \mathbf{B}_{i,k}^\top \mathcal{E}_i(\overline{\mathbf{P}}_{k+1}^{[\eta]}) \mathbf{B}_{i,k} \mathbf{L}_k^{[\eta]} \\
&\quad + (\mathbf{F}_k^{[\eta]} - \mathbf{B}_{i,k}\mathbf{L}_k^{[\eta]})^\top \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}^{[\eta]}) (\mathbf{F}_k^{[\eta]} - \mathbf{B}_{i,k}\mathbf{L}_k^{[\eta]}) , \\
\underline{\omega}_{i,k}^{[\eta]} &= \operatorname{tr}\left[\mathcal{E}_i(\overline{\mathbf{P}}_{k+1}^{[\eta]} + \underline{\mathbf{P}}_{k+1}^{[\eta]}) \mathbf{H}_{i,k} \mathbf{H}_{i,k}^\top + \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}^{[\eta]}) \mathbf{K}_k^{[\eta]} \mathbf{J}_{i,k} \mathbf{J}_{i,k}^\top (\mathbf{K}_k^{[\eta]})^\top \right] + \mathcal{E}_i(\underline{\omega}_{,k+1}^{[\eta]}) .
\end{aligned}$$

  Then, set $k = k - 1$ and if $k > 0$ return to *Step 4*. Otherwise, proceed to *Step 7*.
- *Step 7:* Compute $\overline{\mathcal{J}}_0^{[\eta]}$ using (3.72). If $\overline{\mathcal{J}}_0^{[\eta-1]} - \overline{\mathcal{J}}_0^{[\eta]}$ is sufficiently small, stop the algorithm. Otherwise, set $\eta = \eta + 1$ and return to *Step 2*.

	$\mathbf{T}_{[1]}, \underline{\xi}_{0[1]}$	$\mathbf{T}_{[1]}, \underline{\xi}_{0[2]}$	$\mathbf{T}_{[2]}, \underline{\xi}_{0[1]}$	$\mathbf{T}_{[2]}, \underline{\xi}_{0[2]}$
optimal controller from [38]	0.7254	0.0999	0.7523	0.1081
proposed time-invariant controller	0.7423	0.1153	0.7587	0.1150
proposed time-variant controller	0.7400	0.1142	0.7564	0.1147

Table 3.3: Median costs of the Monte Carlo simulation for the three compared controllers.

To conclude this section, the next theorem presents results on the convergence of the proposed iterative algorithm.

Theorem 3.33. *The sequences* $(\mathbf{F}_{0:K-1}^{[\eta]}, \mathbf{K}_{0:K-1}^{[\eta]}, \mathbf{L}_{0:K-1}^{[\eta]})$ *computed using the iterative Algorithm 3.3 generate a sequence of (non-strictly) monotonically decreasing costs, i.e.,* $\overline{\mathcal{J}}_0^{[\eta+1]} \leq \overline{\mathcal{J}}_0^{[\eta]}$. *Furthermore, the sequence of controller parameters* $(\mathbf{F}_{0:K-1}^{[\eta]}, \mathbf{K}_{0:K-1}^{[\eta]}, \mathbf{L}_{0:K-1}^{[\eta]})$ *converges to* $(\mathbf{F}_{0:K-1}, \mathbf{K}_{0:K-1}, \mathbf{L}_{0:K-1}) = \lim_{\eta\to\infty}(\mathbf{F}_{0:K-1}^{[\eta]}, \mathbf{K}_{0:K-1}^{[\eta]}, \mathbf{L}_{0:K-1}^{[\eta]})$.

Proof. The proof is given in Appendix B.3. □

In the next section, we provide a numerical example, where we demonstrate the performance of both the infinite-horizon and finite-horizon controllers for MJLS without mode observations, whose parameters are computed using the iterative algorithms Algorithm 3.2 and Algorithm 3.3, respectively.

3.7.3 Numerical Example

In order to demonstrate the time-invariant control approach presented in Sec. 3.6 and the time-variant method from this section, we choose the parameters of (3.54) to

$$\mathbf{A}_1 = \begin{bmatrix} 1.2 & 1.2 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 1 & 0.8 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} 0 \\ 0.2 \end{bmatrix}, \quad \mathbf{H}_1 = \mathbf{H}_2 = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix},$$
$$\mathbf{C}_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} 0.8 & 0 \end{bmatrix}, \quad \mathbf{J}_1 = \mathbf{J}_2 = 0.1, \quad \mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{I}_2, \quad \mathbf{R}_1 = \mathbf{R}_2 = 1.$$

Two different transition matrices

$$\mathbf{T}_{[1]} = \begin{bmatrix} 0.7 & 0.3 \\ 0.6 & 0.4 \end{bmatrix} \quad \text{and} \quad \mathbf{T}_{[2]} = \begin{bmatrix} 0.9 & 0.1 \\ 0.9 & 0.1 \end{bmatrix}$$

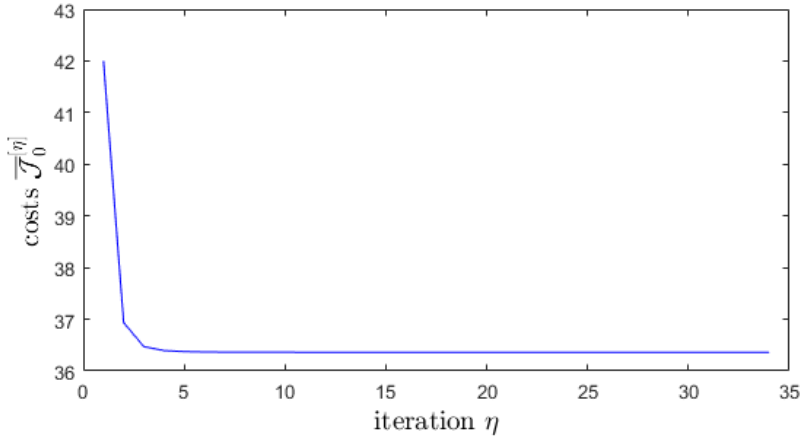


Figure 3.6: Convergence of the bound $\overline{\mathcal{J}}_0[\eta]$ of the optimal costs $\mathcal{J}_0$.

with the initial mode distribution $\underline{\mu}_0 = [1 \;\; 0]$ are assumed. The spectral radius of the open-loop system is 1.2970 for $\mathbf{T}_{[1]}$ and 1.3960 for $\mathbf{T}_{[2]}$, which means that the MJLS is unstable in both cases. For the initial system state, we choose two different means $(\underline{\xi}^1_{0,[1]} = [5 \;\; 0]^\top, \underline{\xi}^2_{0,[1]} = [0 \;\; 0]^\top)$ and $(\underline{\xi}^1_{0,[2]} = \underline{\xi}^2_{0,[2]} = [0 \;\; 0]^\top)$ with covariances $\mathbf{\Xi}^1_0 = 0.1^2 \cdot \mathbf{I}_2, \mathbf{\Xi}^2_0 = \mathbf{0}_2$. The simulation includes the control approaches presented in this section and Sec. 3.6, and the optimal control approach with mode feedback from [38]. The results of a Monte Carlo simulation with 1e5 runs à 100 time steps are summarized in Table 3.3, where the medians of the average costs, i.e., $\mathcal{J}/K$, can be seen. An example run with $\mathbf{T}_{[1]}$ and initial mean $\underline{\xi}^{0,[1]}_0$ is depicted in Fig. 3.8. A reference implementation of the proposed algorithms for computation of the controller parameters for both the infinite-horizon and the finite-horizon control of MJLS without mode observation is available on GitHub [53].

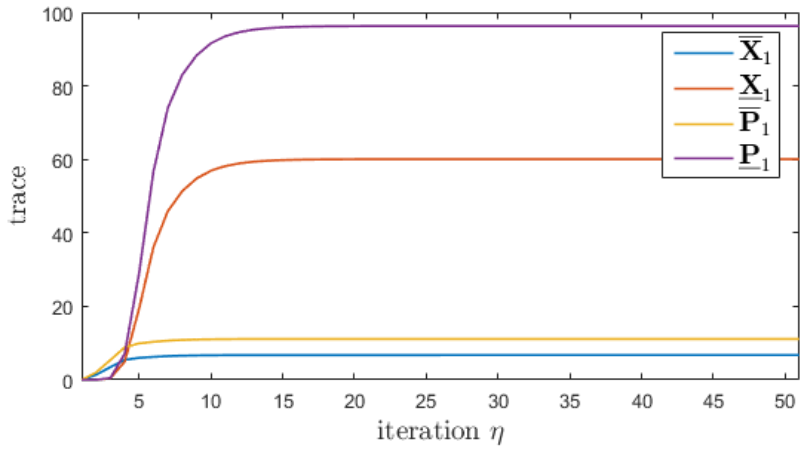


Figure 3.7: Depicted are the traces of $\overline{\mathbf{X}}_1$, $\underline{\mathbf{X}}_1$, $\overline{\mathbf{P}}_1$, and $\underline{\mathbf{P}}_1$, whose convergence indicates convergence of Algorithm 3.2.

According to Table 3.3, the time-variant approach performs slightly better than the time-invariant method. Furthermore, Fig. 3.6 demonstrates the convergence of the expected costs $\overline{\mathcal{J}}_0^{[\eta]}$ over iterations η for the same parameters as in the example run in Fig. 3.8. It can be seen, that the costs converge after approximately five iterations. The convergence of Algorithm 3.2 is demonstrated in Fig. 3.7, where the traces of $\overline{\mathbf{X}}_1$, $\underline{\mathbf{X}}_1$, $\overline{\mathbf{P}}_1$, and $\underline{\mathbf{P}}_1$ are depicted over the number of iteration of Algorithm 3.2. To generate this plot, we assumed the same parameters as used for the example run depicted in Fig. 3.8. It can be seen that the traces of $\overline{\mathbf{X}}_1$, $\underline{\mathbf{X}}_1$, $\overline{\mathbf{P}}_1$, and $\underline{\mathbf{P}}_1$ converge after approximately 15 iterations.

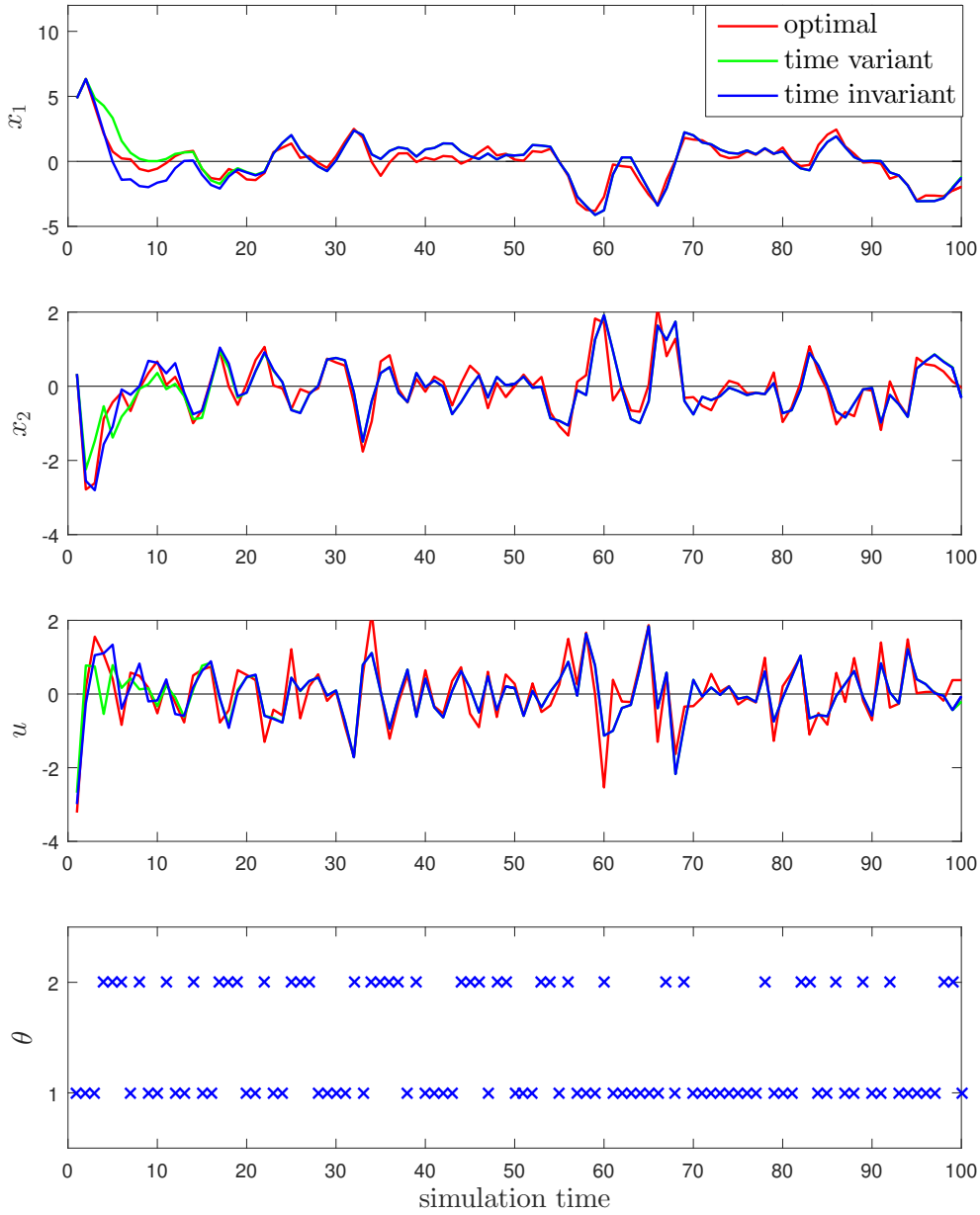


Figure 3.8: Trajectories of an example run with $\mathbf{T}_{[1]}$ and $\underline{\xi}_0^{0,[1]}$.

CHAPTER 4

Conclusion

In stochastic optimal control, dual effect is inherent to optimality except in a few very special cases. Thus, only closed-loop controllers can achieve optimality. However, their computation is usually intractable and requires approximations. In this thesis, we addressed approximate closed-loop stochastic optimal control of (i) nonlinear systems with smooth dynamics and (ii) MJLS with non-observed mode. In what follows, we summarize the contribution presented in this work and outline possible future research directions.

4.1 Contribution

Statistical Taylor Expansion and Linear Closed-loop Controller for Nonlinear Stochastic Systems

In this thesis, we proposed a method for statistical Taylor expansion of a scalar function at a probability distribution given in terms of particles. The second-order statistical Taylor expansion was then used in stochastic optimal control for approximation of the costs along a reference trajectory in order to derive a closed-loop linear controller for stochastic nonlinear systems (the controller linearity was induced by the approximation). Simulations and observations in nonlinear filtering indicate that the proposed approach may be more robust than state-of-the-art approaches which rely on EKF-based Taylor expansion of the costs around the mean of the Gaussian probability distribution that represents the current state. Moreover, the proposed approximation scheme also can be applied to other problems where local function approximation over the domain of probability distributions maintained in form of particles is required.

Framework for GP Regression over the Domain of General Probability Distributions

As an alternative to statistical Taylor expansion, we presented a framework for GP regression of functions whose arguments are general probability distributions. This approximation scheme can be used to obtain control laws that are nonlinear, which can improve the possible controller performance. The main notion of the proposed GP framework is to define the covariance functions that determine the GP in terms of the distance between the probability distributions provided as inputs. Different distance measures between probability distributions were reviewed and their applicability in the proposed framework was discussed. Numerical evaluations indicate

that the estimation of the GP parameters can be performed as in the classical GP framework with deterministic inputs. In a provided example, where a modified Rosenbrock function was approximated, the proposed approximation scheme outperformed a related state-of-the-art approach.

Sufficient Stabilizability Conditions for Linear Static Output-Feedback Control of MJLS without Mode Observation
In our analysis of static output-feedback control of MJLS without mode observation, we have demonstrated that it is possible to recast the BMI, whose solution determines the parameters of the optimal mode-independent regulator gain, into an LMI by introducing additional constraints. This insight allows for an efficient stabilizability check for MJLS with non-observed mode. Unfortunately, the presented criterion depends on the particular state space representation. Thus, we proposed an iterative algorithm for computation of the regulator gain and evaluated its performance in a numerical example. Moreover, we showed that the proposed algorithm converges if the MJLS is stabilizable via mode-independent static output feedback.

Computation of both a Finite-horizon and an Infinite-horizon Linear Dynamic Output-Feedback Controller for MJLS without Mode Observation
Finally, in the last two sections of this thesis, we addressed both finite-horizon and infinite-horizon dynamic output-feedback control of MJLS without mode observation. For each of the considered problems, we proposed an iterative algorithm for computation of the parameters of a linear control law. The main contribution herein consists in the proposition of a bound for the second moment of the closed-loop system state. Additionally, we showed that the proposed algorithm for computation of the parameters of the time-variant control law for finite-horizon control converges to a local minimum. Convergence of the algorithm for computation of the parameters of the time-invariant control law for infinite-horizon control remains an open question.

4.2 Future Research Directions

In the course of this thesis, we have addressed selected problems in control of nonlinear systems with smooth dynamics and systems with linear switching dynamics. Our insights and proposed approaches form a solid foundation for future research and development. In what follows, we discuss a selection of possible future extensions.

In the trajectory optimization approach that uses the proposed second-order statistical Taylor expansion, we optimized only a single nominal reference trajectory. However, to improve robustness and to increase the chance of reaching a global minimum, it is possible to optimize a bundle of reference trajectories. To generate such a bundle, we first need to sample the initial system state and trajectories of the process noise and the measurement noise. Then, these samples can be used together with the current control law in order to generate a bundle of reference trajectories. During improvement of the control law, we then have to perform a

statistical Taylor expansion at the state estimates of each of the trajectories and combine the approximations according to the weights of the trajectories. The investigation of this approach can also include the EKF-based approximation method published in [158, 159].

The proposed framework for GP regression on the space of general probability distributions is not limited to its application as a cost function approximation method in trajectory optimization. Moreover, it can be used for reconstruction of distributed phenomena, for global value function approximation in POMDPs [132], and in reinforcement learning [147] as an approximator for the so-called Q-function. However, for these applications, it is necessary to investigate how well the proposed framework can extrapolate the approximated function in the infinite space of general probability distributions. Nevertheless, the notion to define the covariance functions in terms of the distance between the probability distributions provided as inputs can be applied in other Bayesian function approximation approaches and in probabilistic machine learning.

In order to be able to derive convergent algorithms for computation of controller parameters for dynamic output-feedback control of MJLS without mode observation, we proposed to bound the second moment of the closed-loop system from above. However, this may result in conservatism, because by construction, the bound implies that the internal controller state corresponds to an unbiased estimate of the MJLS state for each mode. Thus, alternative approaches, where this confinement does not hold, may lead to a better control performance. A possible starting point for this research direction are projection-based gradient flow optimization algorithms [167, 88]. Their application to control of MJLS without mode observation would imply a projection of the second moment of the closed-loop system onto positive definite matrices that lie within the matrices formed by our approximation.

Often times, the state of a dynamical system is very large and comprises state variables that are not necessary for control. In order to account for this issue, the internal state of *reduced-order* dynamic controllers has a smaller dimension than the state of the process. Derivation of reduced-order optimal controllers for linear systems with deterministic and white parameters is already mature [91, 135, 13, 170, 161, 171, 86]. On the other hand, reduced-order control of MJLS is quite new and considers only problems where the mode is fed back to the controller [177, 121]. Thus, derivation of reduced-order versions of both the dynamic finite-horizon and the infinite-horizon controllers from Secs. 3.6 and 3.7 constitutes an interesting and relevant future research direction.

Finally, many networked control systems can be formalized as a Markov Jump System (MJS). In scenarios, where both the control inputs and the measurements or local state estimates are transmitted to the controller over a data network that introduces delays and losses, which are modeled as a stochastic process [64], the closed-loop system can be modeled as an MJS with two sets of modes. The first set of modes is observed and models the age of received measurements or local state estimates. On the other hand, the second set of modes is not observed, because it models the age of the control input that is actually applied to the process. Thus, an extension of the methods proposed in the second part of this thesis to such dual-mode MJS forms another possible research direction.

APPENDIX A

Proofs for Sec. 3.6

A.1 Proof of Theorem 3.27

Using Proposition 3.25, we can convert the optimization problem (3.42) into

$$\inf_{\mathbf{F},\mathbf{K},\mathbf{L},\widetilde{\mathbf{X}}_\infty,\widetilde{\mathbf{P}}_\infty} \mathcal{H} = \sum_{i=1}^{M} \mathrm{tr}\left[\widetilde{\mathbf{Q}}_i\widetilde{\mathbf{X}}^i_\infty + \mathcal{E}_i(\widetilde{\mathbf{P}}_\infty)\left[\widetilde{\mathbf{A}}_i\widetilde{\mathbf{X}}^i_\infty\widetilde{\mathbf{A}}_i^\top + \mu^i_\infty\widetilde{\mathbf{H}}_i\widetilde{\mathbf{H}}_i^\top\right] - \widetilde{\mathbf{P}}^i_\infty\widetilde{\mathbf{X}}^i_\infty\right] .$$

The necessary optimality conditions are

$$\frac{\partial\mathcal{H}}{\partial\mathbf{F}} = \mathbf{0}_{n_x} , \ \frac{\partial\mathcal{H}}{\partial\mathbf{K}} = \mathbf{0}_{n_x\times n_y} , \ \frac{\partial\mathcal{H}}{\partial\mathbf{L}} = \mathbf{0}_{n_u\times n_x} , \ \frac{\partial\mathcal{H}}{\partial\widetilde{\mathbf{X}}^i_\infty} = \mathbf{0}_{2n_x} , \ \frac{\partial\mathcal{H}}{\partial\widetilde{\mathbf{P}}^i_\infty} = \mathbf{0}_{2n_x} .$$

From $\partial\mathcal{H}/\partial\widetilde{\mathbf{X}}^i_\infty$ and $\partial\mathcal{H}/\partial\widetilde{\mathbf{P}}^i_\infty$, we obtain

$$\widetilde{\mathbf{X}}^j_\infty = \sum_{i=1}^{M} p_{ij}\left[\widetilde{\mathbf{A}}_i\widetilde{\mathbf{X}}^i_\infty\widetilde{\mathbf{A}}_i^\top + \mu^i_\infty\widetilde{\mathbf{H}}_i\widetilde{\mathbf{H}}_i^\top\right]$$

and

$$\widetilde{\mathbf{P}}^i_\infty = \widetilde{\mathbf{Q}}_i + \widetilde{\mathbf{A}}_i^\top\mathcal{E}_i(\widetilde{\mathbf{P}}_\infty)\widetilde{\mathbf{A}}_i ,$$

for whose partitions it holds

$$\begin{aligned}\mathbf{X}^j_{1,\infty} = \sum_{i=1}^{M} p_{ij}\Big[&\mathbf{A}_i\mathbf{X}^i_{1,\infty}\mathbf{A}_i^\top + \mathbf{A}_i\mathbf{X}^i_{12,\infty}\mathbf{L}^\top\mathbf{B}_i^\top + \mathbf{B}_i\mathbf{L}(\mathbf{X}^i_{12,\infty})^\top\mathbf{A}_i^\top + \mathbf{B}_i\mathbf{L}\mathbf{X}^i_{2,\infty}\mathbf{L}^\top\mathbf{B}_i^\top \\ &+\mu^i_\infty\mathbf{H}_i\mathbf{H}_i^\top\Big] ,\end{aligned} \tag{A.1}$$

$$\mathbf{X}^j_{12,\infty} = \sum_{i=1}^{M} p_{ij}\left[\mathbf{A}_i\mathbf{X}^i_{1,\infty}\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{A}_i\mathbf{X}^i_{12,\infty}\mathbf{F}^\top + \mathbf{B}_i\mathbf{L}(\mathbf{X}^i_{12,\infty})^\top\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{B}_i\mathbf{L}\mathbf{X}^i_{2,\infty}\mathbf{F}^\top\right] , \tag{A.2}$$

$$\begin{aligned}\mathbf{X}^j_{2,\infty} = \sum_{i=1}^{M} p_{ij}\Big[&\mathbf{K}\mathbf{C}_i\mathbf{X}^i_{1,\infty}\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{K}\mathbf{C}_i\mathbf{X}^i_{12,\infty}\mathbf{F}^\top + \mathbf{F}(\mathbf{X}^i_{12,\infty})^\top\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{F}\mathbf{X}^i_{2,\infty}\mathbf{F}^\top \\ &+\mu^i_\infty\mathbf{K}\mathbf{J}_i\mathbf{J}_i^\top\mathbf{K}^\top\Big] ,\end{aligned} \tag{A.3}$$

$$\begin{aligned}\mathbf{P}^i_{1,\infty} &= \mathbf{Q}_i + \mathbf{A}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{A}_i + \mathbf{A}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty})\mathbf{K}\mathbf{C}_i + \mathbf{C}_i^\top \mathbf{K}^\top (\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top \mathbf{A}_i \\ &\quad + \mathbf{C}_i^\top \mathbf{K}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{K}\mathbf{C}_i \ , \qquad \text{(A.4)} \\ \mathbf{P}^i_{12,\infty} &= \mathbf{A}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{B}_i\mathbf{L} + \mathbf{A}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty})\mathbf{F} + \mathbf{C}_i^\top \mathbf{K}^\top (\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top \mathbf{B}_i\mathbf{L} + \mathbf{C}_i^\top \mathbf{K}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{F} \ , \qquad \text{(A.5)} \\ \mathbf{P}^i_{2,\infty} &= \mathbf{L}^\top \mathbf{B}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{B}_i\mathbf{L} + \mathbf{L}^\top \mathbf{B}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty})\mathbf{F} + \mathbf{F}^\top (\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top \mathbf{B}_i\mathbf{L} + \mathbf{F}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{F} \\ &\quad + \mathbf{L}^\top \mathbf{R}_i\mathbf{L} \ . \qquad \text{(A.6)}\end{aligned}$$

Using substitutions (3.43) in (A.3), we obtain (3.45). Next, we show (3.45). To this end, observe that (3.43) yields

$$\begin{aligned}\overline{\mathbf{X}}_i &= \lim_{k\to\infty} \mathrm{E}\left\{\underline{\boldsymbol{x}}_k \underline{\boldsymbol{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} - \mathrm{E}\left\{\underline{\boldsymbol{x}}_k \underline{\hat{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} - \mathrm{E}\left\{\underline{\hat{x}}_k \underline{\boldsymbol{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} + \mathrm{E}\left\{\underline{\hat{x}}_k \underline{\hat{x}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \\ &= \mathbf{X}^i_{1,\infty} - \mathbf{X}^i_{12,\infty} - (\mathbf{X}^i_{12,\infty})^\top + \mathbf{X}^i_{2,\infty} \\ &= \sum_{i=1}^M p_{ij} \Big[\mathbf{A}_i\mathbf{X}^i_{1,\infty}\mathbf{A}_i^\top + \mathbf{A}_i\mathbf{X}^i_{12,\infty}\mathbf{L}^\top\mathbf{B}_i^\top + \mathbf{B}_i\mathbf{L}(\mathbf{X}^i_{12,\infty})^\top\mathbf{A}_i^\top + \mathbf{B}_i\mathbf{L}\mathbf{X}^i_{2,\infty}\mathbf{L}^\top\mathbf{B}_i^\top + \mu^i_\infty \mathbf{H}_i\mathbf{H}_i^\top \\ &\quad -\mathbf{A}_i\mathbf{X}^i_{1,\infty}\mathbf{C}_i^\top\mathbf{K}^\top - \mathbf{A}_i\mathbf{X}^i_{12,\infty}\mathbf{F}^\top - \mathbf{B}_i\mathbf{L}(\mathbf{X}^i_{12,\infty})^\top\mathbf{C}_i^\top\mathbf{K}^\top - \mathbf{B}_i\mathbf{L}\mathbf{X}^i_{2,\infty}\mathbf{F}^\top - \mathbf{K}\mathbf{C}_i\mathbf{X}^i_{1,\infty}\mathbf{A}_i \\ &\quad -\mathbf{F}(\mathbf{X}^i_{12,\infty})^\top\mathbf{A}_i^\top - \mathbf{K}\mathbf{C}_i\mathbf{X}^i_{12,\infty}\mathbf{L}^\top\mathbf{B}_i^\top - \mathbf{F}\mathbf{X}^i_{2,\infty}\mathbf{L}^\top\mathbf{B}_i^\top + \mathbf{K}\mathbf{C}_i\mathbf{X}^i_{1,\infty}\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{K}\mathbf{C}_i\mathbf{X}^i_{12,\infty}\mathbf{F}^\top \\ &\quad +\mathbf{F}(\mathbf{X}^i_{12,\infty})^\top\mathbf{C}_i^\top\mathbf{K}^\top + \mathbf{F}\mathbf{X}^i_{2,\infty}\mathbf{F}^\top + \mu^i_\infty \mathbf{K}\mathbf{J}_i\mathbf{J}_i^\top\mathbf{K}^\top\Big] \ ,\end{aligned}$$

where we used (A.1)–(A.3). Finally, with substitutions from (3.43), we obtain (3.45).

Next, we show (3.47) and (3.48). In analogy to control of LTI systems, we can interpret the Lagrange multiplier $\widetilde{\mathbf{P}}^i_\infty$ as the second moment of the costate of $\underline{\widetilde{\boldsymbol{x}}}_k$. Let denote this state as $\underline{\widetilde{\boldsymbol{\xi}}}_k = [\underline{\boldsymbol{\xi}}_k^\top \ \ \underline{\widehat{\xi}}_k^\top]^\top$. Then, in analogy to (3.44), the substitutions (3.43) imply

$$\underline{\mathbf{P}}_i = \lim_{k\to\infty} \mathrm{E}\left\{\underline{\widehat{\xi}}_k \underline{\widehat{\xi}}_k^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \text{ and } \overline{\mathbf{P}}_i = \lim_{k\to\infty} \mathrm{E}\left\{(\underline{\boldsymbol{\xi}}_k - \underline{\widehat{\xi}}_k)(\underline{\boldsymbol{\xi}}_k - \underline{\widehat{\xi}}_k)^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \ .$$

Thus, (3.43) used in (A.6) yields (3.48). From

$$\begin{aligned}\overline{\mathbf{P}}_i &= \lim_{k\to\infty} \mathrm{E}\left\{(\underline{\boldsymbol{\xi}}_k - \underline{\widehat{\xi}}_k)(\underline{\boldsymbol{\xi}}_k - \underline{\widehat{\xi}}_k)^\top \mathbb{1}_{\boldsymbol{\theta}_k=i}\right\} \\ &= \mathbf{P}^i_{1,\infty} - \mathbf{P}^i_{12,\infty} - (\mathbf{P}^i_{12,\infty})^\top + \mathbf{P}^i_{2,\infty}\end{aligned}$$

we get by using (A.4)–(A.6)

$$\begin{aligned}\overline{\mathbf{P}}_i &= \mathbf{A}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{A}_i + \mathbf{A}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty})\mathbf{K}\mathbf{C}_i + \mathbf{C}_i^\top \mathbf{K}^\top (\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top \mathbf{A}_i + \mathbf{C}_i^\top \mathbf{K}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{K}\mathbf{C}_i + \mathbf{Q}_i \\ &\quad - \mathbf{A}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{B}_i\mathbf{L} - \mathbf{A}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty})\mathbf{F} - \mathbf{C}_i^\top \mathbf{K}^\top (\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top \mathbf{B}_i\mathbf{L} - \mathbf{C}_i^\top \mathbf{K}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{F} \\ &\quad - \mathbf{L}^\top \mathbf{B}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{A}_i - \mathbf{F}^\top (\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top \mathbf{A}_i - \mathbf{L}^\top \mathbf{B}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty})\mathbf{K}\mathbf{C}_i - \mathbf{F}\mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{K}\mathbf{C}_i \\ &\quad + \mathbf{L}^\top \mathbf{B}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{B}_i\mathbf{L} + \mathbf{L}^\top \mathbf{B}_i^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty})\mathbf{F} + \mathbf{F}^\top (\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top \mathbf{B}_i\mathbf{L} + \mathbf{F}^\top \mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{F} + \mathbf{L}^\top \mathbf{R}_i\mathbf{L}\end{aligned}$$

that yields (3.47) if we apply (3.43).

Now, observe that the Hamiltonian in Proposition 3.25 can be written as

$$\begin{aligned}\mathcal{H} = \sum_{i=1}^{M} \mathrm{tr}\,\Big[&\mathbf{Q}_i\mathbf{X}^i_{1,\infty} + \mathbf{L}^\top\mathbf{R}_i\mathbf{L}\mathbf{X}^i_{2,\infty} - \mathbf{P}^i_{1,\infty}\mathbf{X}^i_{1,\infty} - 2(\mathbf{P}^i_{12,\infty})^\top\mathbf{X}^i_{12,\infty} - \mathbf{P}^i_{2,\infty}\mathbf{X}^i_{2,\infty}\\ &+\mu^i_\infty\mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{H}_i\mathbf{H}_i^\top + \mu^i_\infty\mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{K}\mathbf{J}_i\mathbf{J}_i^\top\mathbf{K}^\top + \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{A}_i\mathbf{X}^i_{1,\infty}\mathbf{A}_i^\top\\ &+2\mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{A}_i\mathbf{X}^i_{12,\infty}\mathbf{L}^\top\mathbf{B}_i^\top + \mathcal{E}_i(\widetilde{\mathbf{P}}_{1,\infty})\mathbf{B}_i\mathbf{L}\mathbf{X}^i_{2,\infty}\mathbf{L}^\top\mathbf{B}_i^\top + 2(\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top\mathbf{A}_i\mathbf{X}^i_{1,\infty}\mathbf{C}_i^\top\mathbf{K}^\top\\ &+2(\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top\mathbf{A}_i\mathbf{X}^i_{12,\infty}\mathbf{F}^\top + 2(\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top\mathbf{B}_i\mathbf{L}(\mathbf{X}^i_{12,\infty})^\top\mathbf{C}_i^\top\mathbf{K}^\top\\ &+2(\mathcal{E}_i(\widetilde{\mathbf{P}}_{12,\infty}))^\top\mathbf{B}_i\mathbf{L}\mathbf{X}^i_{2,\infty}\mathbf{F}^\top + \mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{K}\mathbf{C}_i\mathbf{X}^i_{1,\infty}\mathbf{C}_i^\top\mathbf{K}^\top + 2\mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{K}\mathbf{C}_i\mathbf{X}^i_{12,\infty}\mathbf{F}^\top\\ &+\mathcal{E}_i(\widetilde{\mathbf{P}}_{2,\infty})\mathbf{F}\mathbf{X}^i_{2,\infty}\mathbf{F}^\top\Big]\end{aligned}$$

that yields

$$\begin{aligned}\mathcal{H} = \sum_{i=1}^{M} \mathrm{tr}\,\Big[&\mathbf{Q}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i) + \mathbf{L}^\top\mathbf{R}_i\mathbf{L}\underline{\mathbf{X}}_i - \overline{\mathbf{P}}_i\overline{\mathbf{X}}_i - \underline{\mathbf{P}}_i\overline{\mathbf{X}}_i - \overline{\mathbf{P}}_i\underline{\mathbf{X}}_i + \mu^i_\infty(\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{H}_i\mathbf{H}_i^\top\\ &+\mu^i_\infty\underline{\mathbf{P}}_i\mathbf{K}\mathbf{J}_i\mathbf{J}_i^\top\mathbf{K}^\top + (\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{A}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{A}_i^\top + 2(\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{A}_i\underline{\mathbf{X}}_i\mathbf{L}^\top\mathbf{B}_i^\top\\ &+(\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{B}_i\mathbf{L}\underline{\mathbf{X}}_i\mathbf{L}^\top\mathbf{B}_i^\top - 2\underline{\mathbf{P}}_i\mathbf{A}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{C}_i^\top\mathbf{K}^\top - 2\underline{\mathbf{P}}_i\mathbf{A}_i\underline{\mathbf{X}}_i\mathbf{F}^\top - 2\underline{\mathbf{P}}_i\mathbf{B}_i\mathbf{L}\underline{\mathbf{X}}_i\mathbf{C}_i^\top\mathbf{K}^\top\\ &-2\underline{\mathbf{P}}_i\mathbf{B}_i\mathbf{L}\underline{\mathbf{X}}_i\mathbf{F}^\top + \underline{\mathbf{P}}_i\mathbf{K}\mathbf{C}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{C}_i^\top\mathbf{K}^\top + 2\underline{\mathbf{P}}_i\mathbf{K}\mathbf{C}_i\underline{\mathbf{X}}_i\mathbf{F}^\top + \underline{\mathbf{P}}_i\mathbf{F}\underline{\mathbf{X}}_i\mathbf{F}^\top\Big]\ ,\end{aligned}$$

if we use (3.43). Next, we use the identities

$$\mathrm{vec}\,[\mathbf{ABC}] = (\mathbf{C}^\top \otimes \mathbf{A})\mathrm{vec}\,[\mathbf{B}]$$

and

$$\mathrm{tr}\left[\mathbf{A}^\top\mathbf{B}\right] = \mathrm{vec}\,[\mathbf{A}]^\top\,\mathrm{vec}\,[\mathbf{B}]\ ,$$

which gives us

$$\begin{aligned}\mathcal{H} = \sum_{i=1}^{M} \mathrm{tr}\,\Big[&\mathbf{Q}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i) - \overline{\mathbf{P}}_i\overline{\mathbf{X}}_i - \underline{\mathbf{P}}_i\overline{\mathbf{X}}_i - \overline{\mathbf{P}}_i\underline{\mathbf{X}}_i + \mu^i_\infty(\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{H}_i\mathbf{H}_i^\top\\ &+(\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{A}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{A}_i^\top\Big] + \mathrm{vec}\,[\mathbf{L}]^\top\left(\sum_{i=1}^{M}\left[\underline{\mathbf{X}}_i \otimes (\mathbf{R}_i + \mathbf{B}_i^\top + (\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{B}_i)\right]\right)\mathrm{vec}\,[\mathbf{L}]\\ &+ 2\mathrm{vec}\left[\sum_{i=1}^{M}\mathbf{B}_i^\top(\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{A}_i\underline{\mathbf{X}}_i\right]^\top\mathrm{vec}\,[\mathbf{L}] + \mathrm{vec}\,[\mathbf{F}]^\top\left(\sum_{i=1}^{M}[\underline{\mathbf{X}}_i \otimes \underline{\mathbf{P}}_i]\right)\mathrm{vec}\,[\mathbf{F}]\\ &+ \mathrm{vec}\,[\mathbf{K}]^\top\left(\sum_{i=1}^{M}\left[(\mu^i_\infty\mathbf{J}_i\mathbf{J}_i^\top + \mathbf{C}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{C}_i^\top) \otimes \underline{\mathbf{X}}_i\right]\right)\mathrm{vec}\,[\mathbf{K}] - 2\mathrm{vec}\left[\sum_{i=1}^{M}\underline{\mathbf{P}}_i\mathbf{A}_i\underline{\mathbf{X}}_i\right]^\top\mathrm{vec}\,[\mathbf{F}]\\ &- 2\mathrm{vec}\left[\sum_{i=1}^{M}\underline{\mathbf{P}}_i\mathbf{A}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{C}_i^\top\right]^\top\mathrm{vec}\,[\mathbf{K}] - 2\mathrm{vec}\,[\mathbf{L}]^\top\left(\sum_{i=1}^{M}\left[\underline{\mathbf{X}}_i\mathbf{C}_i^\top \otimes \mathbf{B}_i^\top\underline{\mathbf{P}}_i\right]\right)\mathrm{vec}\,[\mathbf{K}]\\ &- 2\mathrm{vec}\,[\mathbf{L}]^\top\left(\sum_{i=1}^{M}\left[\underline{\mathbf{X}}_i \otimes \mathbf{B}_i^\top\underline{\mathbf{P}}_i\right]\right)\mathrm{vec}\,[\mathbf{F}] + 2\mathrm{vec}\,[\mathbf{K}]^\top\left(\sum_{i=1}^{M}[\mathbf{C}_i\underline{\mathbf{X}}_i \otimes \underline{\mathbf{P}}_i]\right)\mathrm{vec}\,[\mathbf{F}]\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{M} \operatorname{tr}\left[\mathbf{Q}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i) - \overline{\mathbf{P}}_i\overline{\mathbf{X}}_i - \underline{\mathbf{P}}_i\overline{\mathbf{X}}_i - \overline{\mathbf{P}}_i\underline{\mathbf{X}}_i + (\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{A}_i(\overline{\mathbf{X}}_i + \underline{\mathbf{X}}_i)\mathbf{A}_i^\top\right] \\
&+ \mu_\infty^i(\overline{\mathbf{\Lambda}}_i + \underline{\mathbf{P}}_i)\mathbf{H}_i\mathbf{H}_i^\top + \operatorname{vec}[\mathbf{L}]^\top \mathbf{\Phi}_L(\mathcal{D})\operatorname{vec}[\mathbf{L}] + \operatorname{vec}[\mathbf{K}]^\top \mathbf{\Phi}_K(\mathcal{D})\operatorname{vec}[\mathbf{K}] \qquad \text{(A.7)} \\
&+ \operatorname{vec}[\mathbf{F}]^\top \mathbf{\Phi}_F(\mathcal{D})\operatorname{vec}[\mathbf{F}] + 2\operatorname{vec}[\mathbf{K}]^\top \mathbf{\Gamma}_{KL}(\mathcal{D})\operatorname{vec}[\mathbf{L}] + 2\operatorname{vec}[\mathbf{F}]^\top \mathbf{\Gamma}_{FL}(\mathcal{D})\operatorname{vec}[\mathbf{L}] \\
&+ 2\operatorname{vec}[\mathbf{F}]^\top \mathbf{\Gamma}_{FK}(\mathcal{D})\operatorname{vec}[\mathbf{K}] + 2\operatorname{vec}[\mathbf{L}]^\top \underline{\rho}_L(\mathcal{D}) + 2\operatorname{vec}[\mathbf{K}]^\top \underline{\rho}_K(\mathcal{D}) + 2\operatorname{vec}[\mathbf{F}]^\top \underline{\rho}_F(\mathcal{D}) \ ,
\end{aligned}
$$

where we used the definitions (3.50). Differentiation of (A.7) with respect to $\operatorname{vec}[\mathbf{F}]$, $\operatorname{vec}[\mathbf{K}]$, and $\operatorname{vec}[\mathbf{L}]$, and setting the results equal to zero yields

$$
\begin{aligned}
&2\mathbf{\Phi}_F(\mathcal{D})\operatorname{vec}[\mathbf{F}] + 2\mathbf{\Gamma}_{FL}(\mathcal{D})\operatorname{vec}[\mathbf{L}] + 2\mathbf{\Gamma}_{FK}(\mathcal{D})\operatorname{vec}[\mathbf{F}] + 2\underline{\rho}_F(\mathcal{D}) = 0 \ , \\
&2\mathbf{\Phi}_K(\mathcal{D})\operatorname{vec}[\mathbf{K}] + 2\mathbf{\Gamma}_{KL}(\mathcal{D})\operatorname{vec}[\mathbf{L}] + 2\mathbf{\Gamma}_{FK}(\mathcal{D})\operatorname{vec}[\mathbf{F}] + 2\underline{\rho}_K(\mathcal{D}) = 0 \ , \\
&2\mathbf{\Phi}_L(\mathcal{D})\operatorname{vec}[\mathbf{L}] + 2\mathbf{\Gamma}_{KL}(\mathcal{D})\operatorname{vec}[\mathbf{K}] + 2\mathbf{\Gamma}_{FL}(\mathcal{D})\operatorname{vec}[\mathbf{F}] + 2\underline{\rho}_L(\mathcal{D}) = 0
\end{aligned}
$$

that can be written as the necessary optimality condition (3.49), thus concluding the proof.

APPENDIX B

Proofs for Sec. 3.7

B.1 Proof of Theorem 3.31

To prove the claim of Theorem 3.31, we first show by induction that the dynamics of the closed-loop second moment

$$\widetilde{\mathbf{X}}_{k+1}^{j} = \sum_{i=1}^{M} p_{ij,k} \left[\widetilde{\mathbf{A}}_{i,k} \widetilde{\mathbf{X}}_{k}^{i} \widetilde{\mathbf{A}}_{i,k}^{\top} + \mu_{k}^{i} \widetilde{\mathbf{H}}_{i,k} \widetilde{\mathbf{H}}_{i,k}^{\top} \right]$$

can be bounded pointwise from above in the positive semidefinite sense using Proposition 3.26. At time step $k = 0$, it follows directly from (3.43) that

$$\widetilde{\mathbf{X}}_{0}^{j} = \begin{bmatrix} \mathbf{X}_{1,0}^{i} & \mathbf{X}_{12,0}^{i} \\ (\mathbf{X}_{12,0}^{i})^{\top} & \mathbf{X}_{2,0}^{i} \end{bmatrix} \leq \begin{bmatrix} \overline{\mathbf{X}}_{i,0} + \underline{\mathbf{X}}_{i,0} & \underline{\mathbf{X}}_{i,0} \\ \underline{\mathbf{X}}_{i,0} & \underline{\mathbf{X}}_{i,0} \end{bmatrix} .$$

Now, let the bound hold for a $k = 0, \ldots, K-1$. Then, due to the linearity of the operator that maps $\widetilde{\mathbf{X}}_{k}^{i}$, $i = 1, \ldots, M$ to $\widetilde{\mathbf{X}}_{k+1}^{j}$ for $j = 1, \ldots, M$, we have

$$\widetilde{\mathbf{X}}_{k+1}^{j} = \sum_{i=1}^{M} p_{ij,k} \left[\widetilde{\mathbf{A}}_{i,k} \widetilde{\mathbf{X}}_{k}^{i} \widetilde{\mathbf{A}}_{i,k}^{\top} + \mu_{k}^{i} \widetilde{\mathbf{H}}_{i,k} \widetilde{\mathbf{H}}_{i,k}^{\top} \right] \leq \sum_{i=1}^{M} p_{ij,k} \left[\widetilde{\mathbf{A}}_{i,k} \hat{\mathbf{X}}_{k}^{i} \widetilde{\mathbf{A}}_{i,k}^{\top} + \mu_{k}^{i} \widetilde{\mathbf{H}}_{i,k} \widetilde{\mathbf{H}}_{i,k}^{\top} \right] , \quad \text{(B.1)}$$

where

$$\hat{\mathbf{X}}_{k}^{i} = \begin{bmatrix} \overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k} & \underline{\mathbf{X}}_{i,k} \\ \underline{\mathbf{X}}_{i,k} & \underline{\mathbf{X}}_{i,k} \end{bmatrix} .$$

Using that the dynamics of the cost-related matrices $\widetilde{\mathbf{P}}_{k}^{i}$ are determined by

$$\widetilde{\mathbf{P}}_{k}^{i} = \widetilde{\mathbf{Q}}_{i,k} + \widetilde{\mathbf{A}}_{i,k}^{\top} \mathcal{E}_{i}(\widetilde{\mathbf{P}}_{k+1}) \widetilde{\mathbf{A}}_{i,k} ,$$

and following the same argumentation via induction that starts at $k = K$ as above, we can obtain the pointwise upper bound in the positive semidefinite sense

$$\widetilde{\mathbf{P}}_{k}^{i} = \begin{bmatrix} \mathbf{P}_{k}^{i} & \mathbf{P}_{12,k}^{i} \\ (\mathbf{P}_{12,k}^{i})^{\top} & \mathbf{P}_{2,k}^{i} \end{bmatrix} \leq \begin{bmatrix} \overline{\mathbf{P}}_{i,k} + \underline{\mathbf{P}}_{i,k} & -\underline{\mathbf{P}}_{i,k} \\ -\underline{\mathbf{P}}_{i,k} & \underline{\mathbf{P}}_{i,k} \end{bmatrix} = \hat{\mathbf{P}}_{k}^{i} , \quad \text{(B.2)}$$

where $\overline{\mathbf{P}}_{i,k}$ and $\underline{\mathbf{P}}_{i,k}$ are defined as in Proposition 3.26.

With these prerequisites, we can show by induction that the pointwised upper bound (3.65) holds. To this end, we observe that from (B.1) and (B.2), we obtain using Lemma 3.30

$$\operatorname{tr}\left[\widetilde{\mathbf{P}}_k^i \widetilde{\mathbf{X}}_k^i\right] \leq \operatorname{tr}\left[\hat{\mathbf{P}}_k^i \hat{\mathbf{X}}_k^i\right] = \operatorname{tr}\left[\overline{\mathbf{P}}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) + \underline{\mathbf{P}}_{i,k}\overline{\mathbf{X}}_{i,k}\right] .$$

Finally, using the Proposition 3.26 in computation of $\omega_{i,k}$ in (3.29) concludes the proof.

B.2 Proof of Theorem 3.32

We begin by noting that the costs $\mathcal{J}_k$ at time step k of the optimization horizon are independent of the parameters $(\mathbf{F}_t, \mathbf{K}_t, \mathbf{L}_t)$ for $t < k$. Thus, we can use the recursion (3.65) in order to derive (3.70). To this end, assume that except $(\mathbf{F}_k, \mathbf{K}_k, \mathbf{L}_k)$, all other controller parameters $(\mathbf{F}_t, \mathbf{K}_t, \mathbf{L}_t)$ with $t \neq k$ are the minimizers of $\mathcal{J}$. Then, in order to be a minimizer of $\overline{\mathcal{J}}_k$, $(\mathbf{F}_k, \mathbf{K}_k, \mathbf{L}_k)$ must satisfy the necessary optimality conditions

$$\frac{\partial \overline{\mathcal{J}}_k}{\partial \mathbf{F}_k} = \mathbf{0}_{n_x} \ , \ \frac{\partial \overline{\mathcal{J}}_k}{\partial \mathbf{K}_k} = \mathbf{0}_{n_x \times n_y} \ , \ \frac{\partial \overline{\mathcal{J}}_k}{\partial \mathbf{L}_k} = \mathbf{0}_{n_u \times n_x} \ . \tag{B.3}$$

Next, we rewrite $\overline{\mathcal{J}}_k$ as

$$\begin{aligned}
\overline{\mathcal{J}}_k = \sum_{i=1}^{M} \operatorname{tr}\Big[& \mathbf{Q}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) + \mathbf{L}_k^\top \mathbf{R}_{i,k} \mathbf{L}_k \underline{\mathbf{X}}_{i,k} + \mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{i,k+1}) \mathbf{A}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) \mathbf{A}_{i,k}^\top \\
& - 2\mathbf{C}_{i,k}^\top \mathbf{K}_k^\top \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{A}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) + \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{K}_k \mathbf{C}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) \mathbf{C}_{i,k}^\top \mathbf{K}_k^\top \\
& + 2\mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1}) \mathbf{B}_{i,k} \mathbf{L}_k \underline{\mathbf{X}}_{i,k} \mathbf{A}_{i,k}^\top - 2\mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{B}_{i,k} \mathbf{L}_k \underline{\mathbf{X}}_{i,k} \mathbf{C}_{i,k}^\top \mathbf{K}_k^\top \\
& - 2\mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{F}_k \underline{\mathbf{X}}_{i,k} \mathbf{A}_{i,k}^\top + 2\mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{F}_k \underline{\mathbf{X}}_{i,k} \mathbf{C}_{i,k}^\top \mathbf{K}_k^\top - 2\mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{B}_{i,k} \mathbf{L}_k \underline{\mathbf{X}}_{i,k} \mathbf{F}_k^\top \\
& + \mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1}) \mathbf{B}_{i,k} \mathbf{L}_k \underline{\mathbf{X}}_{i,k} \mathbf{L}_k^\top \mathbf{B}_{i,k}^\top + \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{F}_k \underline{\mathbf{X}}_{i,k} \mathbf{F}_k^\top \Big] \\
& + \mu_k^i \mathcal{E}_i(\underline{\omega}_{,k+1}) + \mu_k^i \operatorname{tr}\left[\mathcal{E}_i(\underline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1}) \mathbf{H}_{i,k} \mathbf{H}_{i,k}^\top + \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}) \mathbf{K}_k \mathbf{J}_{i,k} \mathbf{J}_{i,k}^\top \mathbf{K}_k^\top \right] .
\end{aligned} \tag{B.4}$$

Then, we use the identities

$$\begin{aligned}
\operatorname{vec}[\mathbf{ABC}] &= (\mathbf{C}^\top \otimes \mathbf{A}) \operatorname{vec}[\mathbf{B}] \ , \\
\operatorname{tr}\left[\mathbf{A}^\top \mathbf{B}\right] &= \operatorname{vec}[\mathbf{A}]^\top \operatorname{vec}[\mathbf{B}]
\end{aligned}$$

in (B.4) and obtain

$$\begin{aligned}
\overline{\mathcal{J}}_k = \sum_{i=1}^{M} \operatorname{tr}\Big[& \mathbf{Q}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) + \mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{i,k+1}) \mathbf{A}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) \mathbf{A}_{i,k}^\top \\
& + \mu_k^i \mathcal{E}_i(\underline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1}) \mathbf{H}_{i,k} \mathbf{H}_{i,k}^\top \Big] + \mu_k^i \mathcal{E}_i(\underline{\omega}_{,k+1})
\end{aligned}$$

$$
\begin{aligned}
&+ \operatorname{vec}\left[\mathbf{L}_k\right]^\top \left(\sum_{i=1}^{M}\left[\underline{\mathbf{X}}_{i,k} \otimes (\mathbf{R}_{i,k} + \mathbf{B}_{i,k}^\top \mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1})\mathbf{B}_{i,k})\right]\right) \operatorname{vec}\left[\mathbf{L}_k\right] \\
&+ \operatorname{vec}\left[\mathbf{K}_k\right]^\top \left(\sum_{i=1}^{M}\left[(\mu_k^i \mathbf{J}_{i,k}\mathbf{J}_{i,k}^\top + \mathbf{C}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k})\mathbf{C}_{i,k}^\top) \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1})\right]\right) \operatorname{vec}\left[\mathbf{K}_k\right] \\
&- 2\operatorname{vec}\left[\mathbf{K}_k\right]^\top \left(\sum_{i=1}^{M}\left[\mathbf{C}_{i,k}\underline{\mathbf{X}}_{i,k} \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1}\mathbf{B}_{i,k})\right]\right) \operatorname{vec}\left[\mathbf{L}_k\right] \\
&+ 2\operatorname{vec}\left[\mathbf{L}_k\right]^\top \operatorname{vec}\left[\sum_{i=1}^{M}\mathbf{B}_{i,k}^\top \mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1})\mathbf{A}_{i,k}\underline{\mathbf{X}}_{i,k}\right] \\
&- 2\operatorname{vec}\left[\mathbf{K}_k\right]^\top \operatorname{vec}\left[\sum_{i=1}^{M}\mathcal{E}_i(\underline{\mathbf{P}}_{k+1})\mathbf{A}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k})\mathbf{C}_{i,k}^\top\right] \\
&- 2\operatorname{vec}\left[\mathbf{F}_k\right]^\top \operatorname{vec}\left[\sum_{i=1}^{M}\mathcal{E}_i(\underline{\mathbf{P}}_{k+1})\mathbf{A}_{i,k}\underline{\mathbf{X}}_{i,k}\right] \\
&- 2\operatorname{vec}\left[\mathbf{F}_k\right]^\top \left(\sum_{i=1}^{M}\left[\underline{\mathbf{X}}_{i,k} \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1})\mathbf{B}_{i,k}\right]\right) \operatorname{vec}\left[\mathbf{L}_k\right] \\
&+ 2\operatorname{vec}\left[\mathbf{F}_k\right]^\top \left(\sum_{i=1}^{M}\left[\underline{\mathbf{X}}_{i,k}\mathbf{C}_{i,k}^\top \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1})\right]\right) \operatorname{vec}\left[\mathbf{K}_k\right] \\
&+ \operatorname{vec}\left[\mathbf{F}_k\right]^\top \left(\sum_{i=1}^{M}\left[\underline{\mathbf{X}}_{i,k} \otimes \mathcal{E}_i(\underline{\mathbf{P}}_{k+1})\right]\right) \operatorname{vec}\left[\mathbf{F}_k\right] \\
&= \sum_{i=1}^{M} \operatorname{tr}\Big[\mathbf{Q}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k}) + \mathcal{E}_i(\overline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{i,k+1})\mathbf{A}_{i,k}(\overline{\mathbf{X}}_{i,k} + \underline{\mathbf{X}}_{i,k})\mathbf{A}_{i,k}^\top \\
&+ \mu_k^i \mathcal{E}_i(\underline{\mathbf{P}}_{k+1} + \underline{\mathbf{P}}_{k+1})\mathbf{H}_{i,k}\mathbf{H}_{i,k}^\top\Big] + \mu_k^i \mathcal{E}_i(\underline{\omega}_{,k+1}) + \operatorname{vec}\left[\mathbf{L}_k\right]^\top \mathbf{\Phi}_L(\mathcal{D}_k)\operatorname{vec}\left[\mathbf{L}_k\right] \\
&+ \operatorname{vec}\left[\mathbf{K}_k\right]^\top \mathbf{\Phi}_K(\mathcal{D}_k)\operatorname{vec}\left[\mathbf{K}_k\right] + \operatorname{vec}\left[\mathbf{F}_k\right]^\top \mathbf{\Phi}_F(\mathcal{D}_k)\operatorname{vec}\left[\mathbf{F}_k\right] + 2\operatorname{vec}\left[\mathbf{K}_k\right]^\top \mathbf{\Gamma}_{KL}(\mathcal{D}_k)\operatorname{vec}\left[\mathbf{L}_k\right] \\
&+ 2\operatorname{vec}\left[\mathbf{F}_k\right]^\top \mathbf{\Gamma}_{FL}(\mathcal{D}_k)\operatorname{vec}\left[\mathbf{L}_k\right] + 2\operatorname{vec}\left[\mathbf{F}_k\right]^\top \mathbf{\Gamma}_{FK}(\mathcal{D}_k)\operatorname{vec}\left[\mathbf{K}_k\right] + 2\operatorname{vec}\left[\mathbf{L}_k\right]^\top \underline{\rho}_L(\mathcal{D}_k) \\
&+ 2\operatorname{vec}\left[\mathbf{K}_k\right]^\top \underline{\rho}_K(\mathcal{D}_k) + 2\operatorname{vec}\left[\mathbf{F}_k\right]^\top \underline{\rho}_F(\mathcal{D}_k)\ .
\end{aligned}
\tag{B.5}
$$

Finally, evaluation of the necessary optimality conditions (B.3) yields the results of Theorem 3.32.

B.3 Proof of Theorem 3.33

We begin by defining the sequence of controller parameters $\mathcal{G}_{0:K-1}^{[\eta,t]}$ at iteration η of Algorithm 3.3 according to

$$
\begin{aligned}
\mathcal{G}_{0:K-1}^{[\eta,t]} = \{&(\mathbf{F}_0^{[\eta-1]}, \mathbf{K}_0^{[\eta-1]}, \mathbf{L}_0^{[\eta-1]}), \ldots, (\mathbf{F}_{t-1}^{[\eta-1]}, \mathbf{K}_{t-1}^{[\eta-1]}, \mathbf{L}_{t-1}^{[\eta-1]}), \\
&(\mathbf{F}_t^{[\eta]}, \mathbf{K}_t^{[\eta]}, \mathbf{L}_t^{[\eta]}), \ldots, (\mathbf{F}_K^{[\eta]}, \mathbf{K}_K^{[\eta]}, \mathbf{L}_K^{[\eta]})\}\ ,
\end{aligned}
$$

where the controller parameters from t to the end of the optimization horizon have been updated (highlighted in blue in the above equation) and the others are still the same as after having performed iteration $\eta - 1$. In order to show pointwise convergence of the costs-to-go at a time step k of the optimization horizon, we will evaluate the difference $\overline{\mathcal{J}}_k^{[\eta,k]} - \overline{\mathcal{J}}_k^{[\eta,k+1]}$ for $k = 0, \dots, K-1$, i.e., we compute the difference between the costs-to-go $\overline{\mathcal{J}}_k^{[\eta,k]}$ at iteration η of Algorithm 3.3, whose control law at time step k has been updated, and the costs $\overline{\mathcal{J}}_k^{[\eta,k+1]}$, whose control law at time step k is the same as at iteration $\eta - 1$ of Algorithm 3.3. We can formalize this definition of $\overline{\mathcal{J}}_k^{[\eta,k]}$ and $\overline{\mathcal{J}}_k^{[\eta,k+1]}$ by saying that $\overline{\mathcal{J}}_k^{[\eta,k]}$ was computed using $\mathcal{G}_{0:K-1}^{[\eta,k]}$ and $\overline{\mathcal{J}}_k^{[\eta,k+1]}$ using $\mathcal{G}_{0:K-1}^{[\eta,k+1]}$. The sequences $\mathcal{G}_{0:K-1}^{[\eta,k]}$ and $\mathcal{G}_{0:K-1}^{[\eta,k+1]}$ are identical from 0 to $k-1$ and from $k+1$ to the end of the optimization horizon. Thus, $\overline{\mathbf{X}}_{i,0:k}$ and $\underline{\mathbf{X}}_{i,0:k}$, and $\overline{\mathbf{P}}_{i,k+1:K}$ and $\underline{\mathbf{P}}_{i,k+1:K}$ with $i = 1, \dots, M$ are equal for $\mathcal{G}_{0:K-1}^{[\eta,k]}$ and $\mathcal{G}_{0:K-1}^{[\eta,k+1]}$. Also, the sequence of noise-associated costs $\omega_{i,k+1:K}$ is also equal for both sequences of controller parameters. Now, let $(\mathbf{F}_k, \mathbf{K}_k, \mathbf{L}_k)$ be the controller parameters of $\mathcal{G}_{0:K-1}^{[\eta,k]}$ at time step k, i.e., the updated parameters that satisfy (3.70). Further, let $(\mathbf{F}'_k, \mathbf{K}'_k, \mathbf{L}'_k)$ denote the controller parameters of $\mathcal{G}_{0:K-1}^{[\eta,k]}$ before the update. Then, for the difference of bounded costs-to-go (3.65), it holds

$$
\begin{aligned}
\overline{\mathcal{J}}_k^{[\eta,k]} - \overline{\mathcal{J}}_k^{[\eta,k+1]} = {} & \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] \\ \operatorname{vec}[\mathbf{K}_k] \\ \operatorname{vec}[\mathbf{L}_k] \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] \\ \operatorname{vec}[\mathbf{K}_k] \\ \operatorname{vec}[\mathbf{L}_k] \end{bmatrix} \\
& + 2 \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] \\ \operatorname{vec}[\mathbf{K}_k] \\ \operatorname{vec}[\mathbf{L}_k] \end{bmatrix}^\top \begin{bmatrix} \underline{\rho}_F(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_K(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} - 2 \begin{bmatrix} \operatorname{vec}[\mathbf{F}'_k] \\ \operatorname{vec}[\mathbf{K}'_k] \\ \operatorname{vec}[\mathbf{L}'_k] \end{bmatrix}^\top \begin{bmatrix} \underline{\rho}_F(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_K(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \\
& - \begin{bmatrix} \operatorname{vec}[\mathbf{F}'_k] \\ \operatorname{vec}[\mathbf{K}'_k] \\ \operatorname{vec}[\mathbf{L}'_k] \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}[\mathbf{F}'_k] \\ \operatorname{vec}[\mathbf{K}'_k] \\ \operatorname{vec}[\mathbf{L}'_k] \end{bmatrix},
\end{aligned}
\tag{B.6}
$$

where we used the vectorized formulation of $\overline{\mathcal{J}}_k^{[\eta,k]}$ and $\overline{\mathcal{J}}_k^{[\eta,k+1]}$ according to (B.5) and rewrote the difference $\overline{\mathcal{J}}_k^{[\eta,k]} - \overline{\mathcal{J}}_k^{[\eta,k+1]}$ in terms of the stacked vector $[\operatorname{vec}[\mathbf{F}_k]^\top \ \operatorname{vec}[\mathbf{K}_k]^\top \ \operatorname{vec}[\mathbf{L}_k]^\top]^\top$ of the updated controller parameters and the stacked vector $[\operatorname{vec}[\mathbf{F}'_k]^\top \ \operatorname{vec}[\mathbf{K}'_k]^\top \ \operatorname{vec}[\mathbf{L}'_k]^\top]^\top$ of controller parameters before the update.

Next, we observe that from the necessary optimality conditions (3.70), it follows

$$
\begin{bmatrix} \underline{\rho}_F(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_K(\mathcal{D}_k^{[\eta]}) \\ \underline{\rho}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} = - \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] \\ \operatorname{vec}[\mathbf{K}_k] \\ \operatorname{vec}[\mathbf{L}_k] \end{bmatrix}.
$$

Using this identity in (B.6), yields

$$\begin{aligned}
\overline{\mathcal{J}}_k^{[\eta,k]} - \overline{\mathcal{J}}_k^{[\eta,k+1]} = &- \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] \\ \operatorname{vec}[\mathbf{K}_k] \\ \operatorname{vec}[\mathbf{L}_k] \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] \\ \operatorname{vec}[\mathbf{K}_k] \\ \operatorname{vec}[\mathbf{L}_k] \end{bmatrix} \\
&+ 2 \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] \\ \operatorname{vec}[\mathbf{K}_k] \\ \operatorname{vec}[\mathbf{L}_k] \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}[\mathbf{F}'_k] \\ \operatorname{vec}[\mathbf{K}'_k] \\ \operatorname{vec}[\mathbf{L}'_k] \end{bmatrix} \\
&- \begin{bmatrix} \operatorname{vec}[\mathbf{F}'_k] \\ \operatorname{vec}[\mathbf{K}'_k] \\ \operatorname{vec}[\mathbf{L}'_k] \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}[\mathbf{F}'_k] \\ \operatorname{vec}[\mathbf{K}'_k] \\ \operatorname{vec}[\mathbf{L}'_k] \end{bmatrix} \\
= &- \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] - \operatorname{vec}[\mathbf{F}'_k] \\ \operatorname{vec}[\mathbf{K}_k] - \operatorname{vec}[\mathbf{K}'_k] \\ \operatorname{vec}[\mathbf{L}_k] - \operatorname{vec}[\mathbf{L}'_k] \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix} \begin{bmatrix} \operatorname{vec}[\mathbf{F}_k] - \operatorname{vec}[\mathbf{F}'_k] \\ \operatorname{vec}[\mathbf{K}_k] - \operatorname{vec}[\mathbf{K}'_k] \\ \operatorname{vec}[\mathbf{L}_k] - \operatorname{vec}[\mathbf{L}'_k] \end{bmatrix},
\end{aligned}$$

which proves that the difference $\overline{\mathcal{J}}_k^{[\eta,k]} - \overline{\mathcal{J}}_k^{[\eta,k+1]}$ is not positive, because the matrix

$$\begin{bmatrix} \boldsymbol{\Phi}_F(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FK}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_K(\mathcal{D}_k^{[\eta]}) & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]}) \\ \boldsymbol{\Gamma}_{FL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Gamma}_{KL}(\mathcal{D}_k^{[\eta]})^\top & \boldsymbol{\Phi}_L(\mathcal{D}_k^{[\eta]}) \end{bmatrix}$$

is the Hessian of the costs-to-go at time step k, and thus, at least positive semidefinite.

List of Figures

List of Algorithms

Bibliography

[1] E. L. Allgöwer and K. Georg. *Introduction to Numerical Continuation Methods*, volume 45 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, USA, 2003.

[2] M. A. Álvarez, L. Rosasco, and N. D. Lawrence. Kernels for vector-valued functions: A review. *Found. Trends Mach. Learn.*, 4(3):195–266, Mar. 2012.

[3] B. D. O. Anderson and J. B. Moore. *Optimal Control: Linear Quadratic Methods*. Prentice-Hall International, Inc., 1989.

[4] P. Antsaklis. Special Issue on Hybrid Systems: Theory and Applications - A Brief Introduction to the Theory and Applications of Hybrid Systems. *Proceedings of the IEEE*, 88(7), July 2000.

[5] I. Arasaratnam and S. Haykin. Cubature Kalman Filters. *IEEE Transactions on Automatic Control*, 54(6):1254–1269, June 2009.

[6] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A Tutorial on Particle Filters for Qnline Nonlinear/Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, Feb 2002.

[7] K. Aström and R. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2002.

[8] M. Athans. The Matrix Minimum Principle. *Information and Control*, 11(5–6):592 – 606, 1967.

[9] H. Bai, D. Hsu, W. S. Lee, and V. A. Ngo. Monte Carlo Value Iteration for Continuous-State POMDPs. In D. Hsu, V. Isler, J.-C. Latombe, and M. C. Lin, editors, *Algorithmic Foundations of Robotics IX: Selected Contributions of the Ninth International Workshop on the Algorithmic Foundations of Robotics*, pages 175–191. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[10] Y. Bar-Shalom, T. Kirubarajan, and X.-R. Li. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., New York, NY, USA, 2002.

[11] Y. Bar-Shalom and E. Tse. Dual Effect, Certainty Equivalence, and Separation in Stochastic Control. *IEEE Transactions on Automatic Control*, 19(5):494–500, Oct 1974.

[12] R. Bellman and R. Kalaba. *Dynamic Programming and Modern Control Theory*. Academic Press, New York, USA, 1965.

[13] D. Bernstein and D. Hyland. Optimal Projection Equations for Reduced-Order Modelling, Estimation and Control of Linear Systems with Multiplicative White Noise. *Journal of Optimization Theory and Applications*, 58(3):387–409, 1988.

[14] D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton University Press, Princeton, USA, 2 edition, 2009.

[15] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Belmont, Massachusetts, 2nd edition, 2000.

[16] D. P. Bertsekas. Dynamic Programming and Suboptimal Control: A Survey from ADP to MPC. *European Journal of Control*, 11(4–5):310–334, 2005.

[17] A. Beydoun, L. Wang, J. Sun, and S. Sivashanka. Hybrid Control of Automotive Powertrain Systems: A Case Study. In T. Henzinger and S. Sastry, editors, *Hybrid Systems: Computation and Control*, volume 1386 of *Lecture Notes in Computer Science*, pages 33–48. Springer Berlin Heidelberg, 1998.

[18] B. Bhaurucha. *On the Stability of Randomly Varying Systems*. PhD thesis, University of California, 1961.

[19] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, Secaucus, NJ, USA, 2006.

[20] L. Blackmore, A. Bektassov, M. Ono, and B. Williams. Robust, Optimal Predictive Control of Jump Markov Linear Systems Using Particles. In A. Bemporad, A. Bicchi, and G. Buttazzo, editors, *Hybrid Systems: Computation and Control*, volume 4416 of *Lecture Notes in Computer Science*, pages 104–117. Springer Berlin Heidelberg, 2007.

[21] E. Boukas and H. Yang. Stability of Discrete-time Linear Systems with Markovian Jumping Parameters. *Mathematics of Control, Signals and Systems*, 8(4):390–402, 1995.

[22] E. Boukas and H. Yang. Exponential Stabilizability of Stochastic Systems with Markovian Jumping Parameters. *Automatica*, 35(8):1437 – 1441, 1999.

[23] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, USA, 1994.

[24] S. Brechtel, T. Gindele, and R. Dillmann. Solving Continuous POMDPs: Value Iteration with Incremental Learning of an Efficient Space Representation. In *Proceedings of the International Conference on Machine Learning (ICML 2013)*, 2013.

[25] A. Brooks, A. Makarenko, S. Williams, and H. Durrant-Whyte. Parametric POMDPs for Planning in Continuous State Spaces. *Robotics and Autonomous Systems*, 54(11):887 – 897, 2006.

[26] A. Bry and N. Roy. Rapidly-exploring Random Belief Trees for Motion Planning Under Uncertainty. 2011.

[27] L. Campo and Y. Bar-Shalom. Control of Discrete-time Hybrid Stochastic Systems. *IEEE Transactions on Automatic Control*, 37(10):1522–1527, Oct 1992.

[28] J. Candela. *Learning with Uncertainty - Gaussian Processes and Relevance Vector Machines*. PhD thesis, Technical University of Denmark, 2004.

[29] J. Q. Candela, A. Girard, J. Larsen, and C. E. Rasmussen. Propagation of Uncertainty in Bayesian Kernel Models - Application to Multiple-step ahead Forecasting. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2003)*, Hong Kong, Apr. 2003.

[30] F. Casiello and K. A. Loparo. Optimal Control of Unknown Parameter Systems. *IEEE Transactions on Automatic Control*, 34(10), 1989.

[31] C. G. Cassandras and J. Lygeros. *Stochastic Hybrid Systems*. CRC Press, 2007.

[32] H. Chizeck and Y. Ji. Optimal Quadratic Control of Jump Linear Systems with Gaussian Noise in Discrete-Time. In *Proceedings of the 27th IEEE Conference on Decision and Control (CDC 1988)*, Austin, USA, Dec 1988.

[33] H. Chizeck and Y. Ji. Applying Jump Systems Control Theory to Discrete Event-Driven Hybrid Systems. In *Proceedings of the 1991 American Control Conference (ACC 1991)*, Boston, USA, June 1991.

[34] H. J. Chizeck and A. S. Willsky. Discrete-time Markovian Jump Linear Quadratic Optimal Control. *International Journal of Control*, 43(1):213 – 231, 1986.

[35] E. F. Costa and J. B. R. do Val. On the Detectability and Observability of Discrete-time Markov Jump Linear Systems. *Systems & Control Letters*, 44(2):135–145, 2001.

[36] O. Costa and M. Fragoso. Stability Results for Discrete-Time Linear Systems with Markovian Jumping Parameters. *Journal of Mathematical Analysis and Applications*, 179(1):154 – 178, 1993.

[37] O. Costa and R. Marques. Robust $H2$-Control for Discrete-time Markovian Jump Linear Systems. In *Proceedings of the 1998 American Control Conference (ACC 1998)*, Philadelphia, USA, Jun 1998.

[38] O. L. V. Costa, M. D. Fragoso, and R. P. Marques. *Discrete-Time Markov Jump Linear Systems*. Springer, London, 1 edition, 2005.

[39] O. L. V. Costa, M. D. Fragoso, and M. G. Todorov. *Continuous-Time Markov Jump Linear Systems*. Springer, Heidelberg, 1 edition, 2013.

[40] O. L. V. Costa, M. D. Fragoso, and M. G. Todorov. A New Approach for the H_2 control of Markov Jump Linear Systems with Partial Information. In *Proceedings of the 19th IFAC World Congress (IFAC 2014)*, Cape Town, South Africa, Aug. 2014.

[41] C. A. R. Crusius and A. Trofino. Sufficient LMI Conditions for Output Feedback Control Problems. *IEEE Transactions on Automatic Control*, 44(5):1053–1057, May 1999.

[42] P. Dallaire, C. Besse, and B. Chaib-draa. Learning Gaussian Process Models from Uncertain Data. In *Proceedings of the 16th International Conference on Neural Information Processing (ICONIP 2009)*, Bangkok, Thailand, Dec. 2009.

[43] P. Dallaire, C. Besse, S. Ross, and B. Chaib-draa. Bayesian Reinforcement Learning in Continuous POMDPs with Gaussian Processes. In *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2009)*, 2009.

[44] W. De Koning. Compensatability and Optimal Compensation of Systems with White Parameters. *IEEE Transactions on Automatic Control*, 37(5):579–588, May 1992.

[45] C. E. de Souza. Robust Stability and Stabilization of Uncertain Discrete-time Markovian Jump Linear Systems. *IEEE Transactions on Automatic Control*, 51(5):836–841, May 2006.

[46] J. D'Errico. Adaptive Robust Numerical Differentiation. http://www.mathworks.com/matlabcentral/fileexchange/13490-adaptive-robust-numerical-differentiation, 2006.

[47] P. Diggle and P. J. Ribeiro. *Model-based Geostatistics*. Springer Series in Statistics. Springer, 2007.

[48] J. do Val and E. Costa. Numerical Solution for Linear-Quadratic Control Problems of Markov Jump Linear Systems and Weak Detectability Concept. *Journal of Optimization Theory and Applications*, 114(1):69–96, 2002.

[49] J. do Val, J. Geromel, and O. Costa. Solutions for the Linear Quadratic Control Problem of Markov Jump Linear Systems. In *Proceedings of the 1999 American Control Conference (ACC 1999)*, San Diego, USA, jun 1999.

[50] J. do Val, J. Geromel, and O. Costa. Solutions for the Linear-Quadratic Control Problem of Markov Jump Linear Systems. *Journal of Optimization Theory and Applications*, 103(2):283–311, 1999.

[51] J. B. R. do Val and T. Başar. Receding Horizon Control of Jump Linear Systems and a Macroeconomic Policy Problem. *Journal of Economic Dynamics and Control*, 23(8):1099 – 1131, 1999.

[52] J. B. R. do Val, J. C. Geromel, and A. P. C. Gonçalves. The H_2-Control for Jump Linear Systems: Cluster Observations of the Markov State. *Automatica*, 38(2):343 – 349, 2002.

[53] M. Dolgov. Reference Implementations of Proposed Controllers. `https://github.com/Mxttak/MJLS`, Nov. 2016.

[54] M. Dolgov, G. Kurz, and U. D. Hanebeck. Finite-horizon Dynamic Compensation of Markov Jump Linear Systems without Mode Observation (to appear). In *Proceedings of the 55th IEEE Conference on Decision and Control (CDC 2016)*, Las Vegas, Nevada, USA, Dec. 2016.

[55] V. Dombrovskii and T. Obedko. Predictive Control of Systems with Markovian Jumps under Constraints and its Application to the Investment Portfolio Optimization. *Automation and Remote Control*, 72(5):989–1003, 2011.

[56] J. Duník, O. Straka, and M. Simandl. The Development of a Randomised Unscented Kalman Filter. *IFAC Proceedings Volumes*, 44(1):8–13, 2011.

[57] R. J. Elliott, T. K. Siu, L. Chan, and J. W. Lau. Pricing Options Under a Generalized Markov-Modulated Jump-Diffusion Model. *Stochastic Analysis and Applications*, 25(4):821–843, 2007.

[58] T. Erez and W. D. Smart. A Scalable Method for Solving High-Dimensional Continuous POMDPs Using Local Approximation. In *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence (UAI 2010)*, 2010.

[59] Y. Fang. *Stability Analysis of Linear Control Systems with Uncertain Parameters*. PhD thesis, Case Western Reserve University, Jan. 1994.

[60] Y. Fang and K. Loparo. Stochastic Stability of Jump Linear Systems. *IEEE Transactions on Automatic Control*, 47(7):1204–1208, Jul 2002.

[61] X. Feng, K. Loparo, Y. Ji, and H. Chizeck. Stochastic Stability Properties of Jump Linear Systems. *IEEE Transactions on Automatic Control*, 37(1):38–53, Jan 1992.

[62] J. Fiala, M. Kocvara, and M. Stingl. PENLAB: A MATLAB Solver for Nonlinear Semidefinite Optimization. *arXiv preprint: Optimization and Control (math.OC)*, Nov. 2013.

[63] A. R. Fioravanti, A. P. Goncalves, and J. C. Geromel. Discrete-time H_∞ output feedback for Markov jump systems with uncertain transition probabilities. *International Journal of Robust and Nonlinear Control*, 23(8):894–902, 2013.

[64] J. Fischer. *Optimal Sequence-Based Control of Networked Linear Systems*. PhD thesis, Karlsruhe Institute of Technology, July 2014.

[65] J. Fischer, A. Hekler, M. Dolgov, and U. D. Hanebeck. Optimal Sequence-Based LQG Control over TCP-like Networks Subject to Random Transmission Delays and Packet Losses. In *Proceedings of the 2013 American Control Conference (ACC 2013)*, Washington D. C., USA, June 2013.

[66] S. Foss and T. Konstantopoulos. An Overview of some Stochastic Stability Methods. *Journal of the Operations Research Society of Japan*, 47(4):275–292, 2004.

[67] J. C. Geromel, G. S. Deaecto, and P. Colaneri. Minimax control of markov jump linear systems. *International Journal of Adaptive Control and Signal Processing*, 2015.

[68] J. C. Geromel, A. P. C. Goncalves, and A. R. Fioravanti. Dynamic Output Feedback Control of Discrete-Time Markov Jump Linear Systems through Linear Matrix Inequalities. *SIAM Journal on Control and Optimization*, 48(2):573–593, 2009.

[69] I. Gilitschenski. *Deterministic Sampling for Nonlinear Dynamic State Estimation*. PhD thesis, Karlsruhe Institute of Technology, 2015.

[70] I. Gilitschenski and U. D. Hanebeck. Efficient Deterministic Dirac Mixture Approximation. In *Proceedings of the 2013 American Control Conference (ACC 2013)*, Washington D. C., USA, June 2013.

[71] I. Gilitschenski, J. Steinbring, U. D. Hanebeck, and M. Simandl. Deterministic Dirac Mixture Approximation of Gaussian Mixtures. In *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, July 2014.

[72] A. Girard. *Approximate Methods for Propagation of Uncertainty with Gaussian Process Models*. PhD thesis, University of Glasgow, 2004.

[73] A. Girard, J. Q. Candela, R. Murray-Smith, and C. E. Rasmussen. Gaussian Process Priors with Uncertain Inputs – Application to Multiple-Step Ahead Time Series Forecasting. In S. Thrun and K. Obermayer, editors, *Advances in Neural Information Processing Systems*. MIT Press, Cambridge, MA, 2002.

[74] A. Girard and R. Murray-Smith. Gaussian Processes: Prediction at a Noisy Input and Application to Iterative Multiple-step ahead Forecasting of Time-Series. In *Switching and Learning in Feedback Systems*, pages 158–184. Springer, 2005.

[75] C. R. Givens and R. M. Shortt. A Class of Wasserstein Metrics for Probability Distributions. *The Michigan Mathematical Journal*, 31(2):231–240, 1984.

[76] A. Goncalves, A. Fioravanti, and J. Geromel. Dynamic Output Feedback H_∞ Control of Discrete-time Markov Jump Linear Systems through Linear Matrix Inequalities. In *Proceedings of the 47th IEEE Conference on Decision and Control (CDC 2008)*, Cancun, Mexico, Dec 2008.

[77] W. S. Gray, O. R. Gonzalez, and S. Patilkulkarni. Stability of Digital Control Systems Subject to Jump‘ Linear Random Perturbations. In *Proceedings of the 39th IEEE Conference on Decision and Control (CDC 2000)*, Sydney, Australia, Dec 2000.

[78] B. Griffiths and K. A. Loparo. Optimal Control of Jump Linear Gaussian Systems. *International Journal of Control*, 43(4):792–819, 1983.

[79] C. M. Grinstead and J. L. Snell. *Introduction to Probability*. American Mathematical Society, 2 edition, 2003.

[80] V. Gupta, R. Murray, and B. Hassibi. On the Control of Jump Linear Markov Systems with Markov State Estimation. In *Proceedings of the 2003 American Control Conference (ACC 2003)*, Denver, USA, June 2003.

[81] U. D. Hanebeck. Optimal Reduction of Multivariate Dirac Mixture Densities. *at - Automatisierungstechnik*, 63(4), 2015.

[82] U. D. Hanebeck, M. F. Huber, and V. Klumpp. Dirac Mixture Approximation of Multivariate Gaussian Densities. In *Proceedings of the 2009 IEEE Conference on Decision and Control (CDC 2009)*, Shanghai, China, Dec. 2009.

[83] U. D. Hanebeck and V. Klumpp. Localized Cumulative Distributions and a Multivariate Generalization of the Cramér-von Mises Distance. In *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, Aug. 2008.

[84] E. A. Hansen. Solving POMDPs by Searching in Policy Space. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, pages 211–219, 1998.

[85] J. Hespanha, P. Naghshtabrizi, and Y. Xu. A Survey of Recent Results in Networked Control Systems. *Proceedings of the IEEE*, 95(1):138–162, 2007.

[86] G. Hilhorst, G. Pipeleers, and J. Swevers. An LMI Approach for Reduced-order H_2 LTI Controller Synthesis. In *Proceedings of the 2013 American Control Conference (ACC 2013)*, pages 2392–2396, June 2013.

[87] H. Hindi. A Tutorial on Convex Optimization. In *Proceedings of the 2004 American Control Conference (ACC 2004)*, volume 4, pages 3252–3265, Boston, USA, June 2004.

[88] M. W. Hirsch, S. Smale, and R. L. Devaney. *Differential equations, dynamical systems, and an introduction to chaos*, volume 60 of *Pure and Applied Mathematics*. Elsevier/Academic Press, Amsterdam, second edition, 2004.

[89] J. R. Hoffman and R. P. S. Mahler. Multitarget Miss Distance and its Applications. In *Proceedings of the Fifth International Conference on Information Fusion (Fusion 2002)*, July 2002.

[90] D. Hoitomt, P. Luh, E. Max, and K. Pattipati. Scheduling Jobs with Simple Precedence Constraints on Parallel Machines. *IEEE Control Systems Magazine*, 10(2):34–40, Feb 1990.

[91] D. Hyland and D. Bernstein. The Optimal Projection Equations for Fixed-order Dynamic Compensation. *IEEE Transactions on Automatic Control*, 29(11):1034–1037, Nov 1984.

[92] M. Jadaliha, Y. Xu, J. Choi, N. S. Johnson, and W. Li. Gaussian Process Regression for Sensor Networks Under Localization Uncertainty. *IEEE Transactions on Signal Processing*, 61(2):223–237, 2013.

[93] Y. Ji and H. J. Cchizeck. Controllability, Observability and Discrete-time Markovian Jump Linear Quadratic Control. *International Journal of Control*, 48(2):481–498, 1988.

[94] Y. Ji and H. J. Cchizeck. Bounded Sample Path Control of Discrete-time Jump Linear Systems. *IEEE Transactions on Systems, Man and Cybernetics*, 19(2):277–284, 1989.

[95] Y. Ji and H. J. Cchizeck. Controllability, Stabilizability, and Continuous-time Markovian Jump Linear Quadratic Control. *IEEE Transactions on Automatic Control*, 35(7):777–788, 1990.

[96] Y. Ji and H. J. Cchizeck. Jump Linear Quadratic Gaussian Control in Continuous Time. In *Proceedings of the 11th IFAC World Congress (IFAC 1990)*, Tallin, USSR, 1990.

[97] Y. Ji and H. J. Cchizeck. Jump Linear Quadratic Gaussian Control: Steady-state Solution and Testable Condition. *Control Theory and Advanced Technology*, 6(3):289–319, 1990.

[98] Y. Ji and H. J. Chizeck. Jump Linear Quadratic Gaussian Control in Continuous Time. *IEEE Transactions on Automatic Control*, 37(12):1884–1892, Dec 1992.

[99] Y. Ji, H. J. Chizeck, X. Feng, and K. A. Loparo. Stability and Control of Discrete-time Jump Linear Systems. *Control Theory and Advanced Technology*, 7(2):247–270, 1991.

[100] R. Jonker and A. Volgenant. A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems. *Computing*, 38(4):325–340, 1987.

[101] S. J. Julier and J. K. Uhlmann. Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE*, 92(3):401–422, Mar 2004.

[102] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[103] F. Kozin. A Survey of Stability of Stochastic Systems. *Automatica*, 5(1):95 – 112, 1969.

[104] N. N. Krasovskii and E. A. Lidskii. Analytic Design of Controller in Systems with Random Attributes - Part I. *Automation and Remote Control*, 22:1021–1025, 1961.

[105] N. N. Krasovskii and E. A. Lidskii. Analytic Design of Controller in Systems with Random Attributes - Part II. *Automation and Remote Control*, 22:1041–1046, 1961.

[106] N. N. Krasovskii and E. A. Lidskii. Analytic Design of Controller in Systems with Random Attributes - Part III. *Automation and Remote Control*, 22:1289–11294, 1961.

[107] R. Krtolica, U. Ozguner, H. Chan, H. Goktas, J. Winkelman, and M. Liubakka. Stability of Linear Feedback Systems with Random Communication Delays. In *Proceedings of the 1991 American Control Conference (ACC 1991)*, Boston, USA, June 1991.

[108] H. J. Kushner. *Stochastic Stability and Control*, volume 33 of *Mathematics in Science and Engineering*. Academic Press, New York, USA, 1967.

[109] T. Lefebvre, H. Bruyninckx, and J. de Schutter. *Nonlinear Kalman Filtering for Force-Controlled Robot Tasks*. Springer Tracts in Advanced Robotics. Springer, 2005.

[110] W. Li and E. Todorov. Iterative Linearization Methods for Approximately Optimal Control and Estimation of Non-linear Stochastic System. *International Journal of Control*, 80(9):1439–1453, 2007.

[111] J. Lin. Divergence Measures Based on the Shannon Entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 2006.

[112] Q. Ling and H. Deng. A New Proof to the Necessity of a Second Moment Stability Condition of Discrete-Time Markov Jump Linear Systems with Real States. *Journal of Applied Mathematics*, 2012.

[113] J. R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons Ltd., Chichester, GB, 3 edition, 2007.

[114] M. Mariton. Jump linear quadratic control with random state discontinuities. *Automatica*, 23(2):237–240, 1987.

[115] M. Mariton. *Jump Linear Systems in Automatic Control*. Marcel Dekker, Inc., New York, USA, 1990.

[116] M. Mariton and P. Bertrand. A Homotopy Algorithm for Solving Coupled Riccati Equations. *Optimal Control Applications and Methods*, 6(4):351–357, 1985.

[117] M. Mariton and P. Bertrand. Output Feedback for a Class of Linear Systems with Stochastic Jump Parameters. *IEEE Transactions on Automatic Control*, 30(9):898–900, 1985.

[118] D. Mayne. A Second-order Gradient Method for Determining Optimal Trajectories of Non-linear Discrete-time Systems. *International Journal of Control*, 3(1):85–95, 1966.

[119] A. McHutchon and C. E. Rasmussen. Gaussian Process Training with Input Noise. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 1341–1349. Curran Associates, Inc., 2011.

[120] C. Morais, M. Braga, R. Oliveira, and P. Peres. H2 Control of Discrete-time Markov Jump Linear Systems with Uncertain Transition Probability Matrix: Improved Linear Matrix Inequality Relaxations and Multi-simplex Modelling. *IET Control Theory and Applications*, 7(12):1665–1674, Aug 2013.

[121] C. F. Morais, M. F. Braga, R. C. L. F. Oliveira, and P. L. D. Peres. Reduced Order Dynamic Output Feedback Control of Uncertain Discrete-time Markov Jump Linear Systems (to appear). *International Journal of Control*, 2016.

[122] T. Morozan. Optimal Stationary Control for Dynamic Systems with Markov Perturbations. *Stochastic Analysis and Applications*, 1(3):299–325, 1983.

[123] T. Morozan. Stabilization of some Stochastic Discrete–time Control Systems. *Stochastic Analysis and Applications*, 1(1):89–116, 1983.

[124] K. P. Murphy. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012.

[125] Y. Nesterov and A. Nemirovski. *Interior-point Polynomial Methods in Convex Programming*, volume 13 of *Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, USA, 1994.

[126] R. Oliveira, A. Vargas, J. do Val, and P. Peres. Robust Stability, H_2 Analysis and Stabilisation of Discrete-time Markov Jump Linear Systems with Uncertain Probability Matrix. *International Journal of Control*, 82(3):470–481, 2009.

[127] R. Oliveira, A. Vargas, J. do Val, and P. Peres. Mode-Independent H_2-Control of a DC Motor Modeled as a Markov Jump Linear System. *IEEE Transactions on Control Systems Technology*, 22(5):1915–1919, Sept 2014.

[128] E. Parzen. On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076, 1962.

[129] J. Perkins and P. Kumar. Stable Distributed Real-time Scheduling of Flexible Manufacturing/Assembly/Disassembly Systems. In *Proceedings of the 27th IEEE Conference on Decision and Control (CDC 1988)*, Tampa, USA, Dec 1988.

[130] J. Pineau, G. Gordon, and S. Thrun. Point-based Value Iteration: An Anytime Algorithm for POMDPs. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2003.

[131] R. Platt, R. Tedrake, L. Kaelbling, and T. Lozano-Perez. Belief Space Planning Assuming Maximum Likelihood Observations. In *Proceedings of Robotics: Science and Systems*, Zaragoza, Spain, June 2010.

[132] J. M. Porta, N. Vlassis, M. T. Spaan, and P. Poupart. Point-Based Value Iteration for Continuous POMDPs. *Journal of Machine Learning Research*, 7:2329–2367, 2006.

[133] W. D. Powell. *Approximate Dynamic Programming*. John Wiley & Sons, 2007.

[134] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2005.

[135] S. Richter. A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence and Global Optimality. In *American Control Conference, 1987*, pages 1527–1531, June 1987.

[136] R. A. L. S. Kullback. On information and sufficiency. *Annals of Mathematical Statistics*, 21(1):79–86, 1951.

[137] S. Särkkä. *Bayesian Filtering and Smoothing*. Cambridge University Press, New York, NY, USA, 2013.

[138] A. V. Savkin and R. J. Evans. *Hybrid Dynamical Systems: Controller and Sensor Switching Problems*. Birkauser, Boston, USA, 2002.

[139] D. Schuhmacher, B. T. Vo, and B. N. Vo. A Consistent Metric for Performance Evaluation of Multi-Object Filters. *IEEE Transactions on Signal Processing*, 56(8):3447–3457, 2008.

[140] M. Shen and S. Fei. A Constructive Method to Static Output Stabilisation of Markov Jump Systems. *International Journal of Control*, 88(5):990–1000, 2015.

[141] R. Smallwod and E. J. Sondik. The Optimal Control of Partially Observable Markov Processes over a Finite Horizon. *Operations Research*, 21(5):1071–1088, 1973.

[142] E. J. Sondik. The Optimal Control of Partially Observable Markov Processes over the Infinite Horizon: Discounted Costs. *Operations Research*, 26(2):282–304, 1978.

[143] Y. Song, S. Liu, and G. Wei. Constrained Robust Distributed Model Predictive Control for Uncertain Discrete-time Markovian Jump Linear System. *Journal of the Franklin Institute*, 352(1):73–92, 2015.

[144] J. Steinbring. Nonlinear Estimation Toolbox.

[145] J. Steinbring and U. D. Hanebeck. LRKF Revisited: The Smart Sampling Kalman Filter (S2KF). *Journal of Advances in Information Fusion*, 9(2):106–123, Dec. 2014.

[146] J. F. Sturm. Using SeDuMi 1.02, A Matlab Toolbox for Optimization over Symmetric Cones. *Optimization Methods and Software*, 11(1-4):625–653, 1999.

[147] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, USA, 1 edition, 1998.

[148] D. D. Sworder. Feedback Control of a Class of Linear Systems with Jump Parameters. *IEEE Transactions on Automatic Control*, 14:9–14, 1989.

[149] S. Thrun. Monte Carlo POMDPs. In S. Solla, T. Leen, and K.-R. Müller, editors, *Advances in Neural Information Processing Systems*, pages 1064–1070. MIT Press, 2000.

[150] E. Todorov and W. Li. A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proceedings of the 2005 American Control Conference (ACC 2005)*, June 2005.

[151] M. G. Todorov and M. D. Fragoso. On the Robust Stability, Stabilization, and Stability Radii of Continuous-Time Infinite Markov Jump Linear Systems. *SIAM Journal on Control and Optimization*, 49(3):1171–1196, 2011.

[152] K. C. Toh, M. J. Todd, and R. H. Tutuncu. SDPT3 — a Matlab Software Package for Semidefinite Programming. *Optimization Methods and Software*, 11:545–581, 1999.

[153] O. Toker and H. Ozbay. On the NP-Hardness of Solving Bilinear Matrix Inequalities and Simultaneous Stabilization with Static Output Feedback. In *Proceedings of the 1995 American Control Conference (ACC 1995)*, Seattle, USA, Jun 1995.

[154] C. Tomlin, G. Pappas, and S. Sastry. Conflict Resolution for Air Traffic Management: a Study in Multiagent Hybrid Systems. *IEEE Transactions on Automatic Control*, 43(4):509–521, Apr 1998.

[155] E. Tse and Y. Bar-Shalom. An Actively Adaptive Control for Linear Systems with Random Parameters via the Dual Control Approach. *IEEE Transactions on Automatic Control*, 18(2):109–117, Apr 1973.

[156] E. Tse, Y. Bar-Shalom, and L. Meier. Wide-sense Adaptive Dual Control for Nonlinear Stochastic Systems. *IEEE Transactions on Automatic Control*, 18(2):98–108, Apr 1973.

[157] C. E. Tyner, J. P. Sutherland, and W. R. J. Gould. Solar Two: A Molten Salt Power Tower Demonstration. UNT Digital Library, USA. `http://digital.library.unt.edu/ark:/67531/metadc613717/` Accessed October 2, 2015.

[158] J. van den Berg, S. Patil, and R. Alterovitz. Motion Planning under Uncertainty using Differential Dynamic Programming in Belief Space. In *Proceedings of the 15th International Symposium on Robotics Research (ISRR 2011*, 2011.

[159] J. van den Berg, S. Patil, and R. Alterovitz. Motion planning under uncertainty using iterative local optimization in belief space. *The International Journal of Robotics Research*, 31(11):1263–1278, 2012.

[160] A. J. van der Schaft and H. Schumacher. *An Introduction to Hybrid Dynamical Systems*. Lecture Notes in Control and Information Sciences. Springer, 2000.

[161] L. G. Van Willenburg and W. L. De Koning. Numerical Algorithms and Issues Concerning the Discrete-Time Optimal Projection Equations. *European Journal of Control*, 6(1):93 – 110, 2000.

[162] J. G. VanAntwerp and R. D. Braatz. A Tutorial on Linear and Bilinear Matrix Inequalities. *Journal of Process Control*, 10(4):363–385, 2000.

[163] A. N. Vargas, L. Acho, G. Pujol, E. F. Costa, J. Y. Ishihara, and J. B. R. do Val. Output Feedback of Markov Jump Linear Systems with no Mode Observation: An Automotive Throttle Application. *International Journal of Robust and Nonlinear Control*, 2015.

[164] A. N. Vargas, D. C. Bortolin, E. F. Costa, and J. B. R. do Val. Gradient-based Optimization Techniques for the Design of Static Controllers for Markov Jump Linear Systems with Unobservable Modes. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 2014.

[165] A. N. Vargas, E. F. Costa, and J. B. R. do Val. 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*, 23(10):1136–1150, 2013.

[166] A. N. Vargas, W. Furloni, and J. B. R. do Val. Constrained Model Predictive Control of Jump Linear Systems with Noise and non-observed Markov State. In *Proceedings of the 2006 American Control Conference (ACC 2006)*, Minneapolis, USA, June 2006.

[167] B. W. Wah and T. Wang. Efficient and Adaptive Lagrange-Multiplier Methods for Nonlinear Continuous Global Optimization. *Journal of Global Optimization*, 14(1):1–25, 1999.

[168] E. A. Wan and R. van der Merwe. *The Unscented Kalman Filter*, pages 221–280. John Wiley & Sons, Inc., 2002.

[169] L. N. Wasserstein. Markov Processes Over Denumerable Products of Spaces, Describing Large Systems of Automata. *Problems of Information Transmission*, 5(3):47–52, 1969.

[170] L. G. V. Willigenburg and W. L. D. Koning. Optimal Reduced-order Compensation of Time-varying Discrete-time Systems with Deterministic and White Parameters. *Automatica*, 35(1):129 – 138, 1999.

[171] L. V. Willigenburg and W. D. Koning. Minimal and Non-minimal Optimal Fixed-order Compensators for Time-varying Discrete-time Systems. *Automatica*, 38(1):157 – 165, 2002.

[172] H. S. Witsenhausen. A Counterexample in Stochastic Optimum Control. *SIAM Journal on Control*, 6(1):131–147, 1968.

[173] W. M. Wonham. Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering. *Journal of the SIAM Series A Control*, 2:106–109, 1964.

[174] W. M. Wonham. Random Differential Equations in Control Theory. *Probabilistic Methods in Applied Mathematics*, 2:131–213, 1972.

[175] R. K. Yedavalli. *Robust Control of Uncertain Dynamic Systems: A Linear State Space Approach*. Springer, New York, USA, 2014.

[176] L. Zhang and E.-K. Boukas. Stability and Stabilization of Markovian Jump Linear Systems with Partly Unknown Transition Probabilities. *Automatica*, 45(2):463–468, 2009.

[177] L. Zhang, E.-K. Boukas, and P. Shi. H_∞ Model Reduction for Discrete-time Markov Jump Linear Systems with Partially Known Transition Probabilities. *International Journal of Control*, 82(2):343–351, 2009.

[178] V. M. Zolotarev. Probability Metrics (in Russian). *Teoriya Veroyatnostei i ee Primeneniya*, 28(1):278–302, 1983.

Own Publications

[179] M. Dolgov, C. Chlebek, and U. D. Hanebeck. Dynamic Compensation of Markov Jump Linear Systems without Mode Observation. In *Proceedings of the 2016 European Control Conference (ECC 2016)*, Aalborg, Denmark, June 2016.

[180] M. Dolgov, J. Fischer, and U. D. Hanebeck. Event-based LQG Control over Networks Subject to Random Transmission Delays and Packet Losses. In *Proceedings of the 4th IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys 2013)*, Koblenz, Germany, Sept. 2013.

[181] M. Dolgov, J. Fischer, and U. D. Hanebeck. Sequence-based LQG Control with Linear Integral Constraints over Stochastic Networks. In *Proceedings of the 53rd IEEE Conference on Decision and Control (CDC 2014)*, Los Angeles, California, USA, Dec. 2014.

[182] M. Dolgov, J. Fischer, and U. D. Hanebeck. Infinite-horizon sequence-based networked control without acknowledgments. In *Proceedings of the 2015 American Control Conference (ACC 2015)*, Chicago, Illinois, USA, July 2015.

[183] M. Dolgov and U. D. Hanebeck. Infinite-horizon Linear Optimal Control of Markov Jump Systems without Mode Observation via State Feedback. *arXiv preprint: Systems and Control (cs.SY)*, July 2015.

[184] M. Dolgov and U. D. Hanebeck. Static Output-Feedback Control of Markov Jump Linear Systems without Mode Observation. Sept. 2015.

[185] M. Dolgov, G. Kurz, and U. D. Hanebeck. State Estimation for Stochastic Hybrid Systems Based on Deterministic Dirac Mixture Approximation. In *Proceedings of the 2014 American Control Conference (ACC 2014)*, Portland, Oregon, USA, June 2014.

[186] M. Dolgov, G. Kurz, and U. D. Hanebeck. Chance-constrained Model Predictive Control based on Box Approximations. In *Proceedings of the 54th IEEE Conference on Decision and Control (CDC 2015)*, Osaka, Japan, Dec. 2015.

[187] M. Dolgov, G. Kurz, and U. D. Hanebeck. Finite-horizon Dynamic Compensation of Markov Jump Linear Systems without Mode Observation (to appear). In *Proceedings of the 55th IEEE Conference on Decision and Control (CDC 2016)*, Las Vegas, Nevada, USA, Dec. 2016.

[188] F. Faion, M. Dolgov, A. Zea, and U. D. Hanebeck. Closed-form bias reduction for shape estimation with polygon models. In *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, Heidelberg, Germany, 2016.

[189] J. Fischer, M. Dolgov, and U. D. Hanebeck. On Stability of Sequence-Based LQG Control. In *Proceedings of the 52st IEEE Conference on Decision and Control (CDC 2013)*, Florence, Italy, Dec. 2013.

[190] J. Fischer, M. Dolgov, and U. D. Hanebeck. Optimal Sequence-Based Tracking Control over Unreliable Networks. In *Proceedings of the 19th IFAC World Congress (IFAC 2014)*, Cape Town, South Africa, Aug. 2014.

[191] J. Fischer, A. Hekler, M. Dolgov, and U. D. Hanebeck. Optimal Sequence-Based LQG Control over TCP-like Networks Subject to Random Transmission Delays and Packet Losses. *arXiv preprint: Systems and Control (cs.SY)*, Nov. 2012.

[192] J. Fischer, A. Hekler, M. Dolgov, and U. D. Hanebeck. Optimal Sequence-Based LQG Control over TCP-like Networks Subject to Random Transmission Delays and Packet Losses. In *Proceedings of the 2013 American Control Conference (ACC 2013)*, Washington D. C., USA, June 2013.

[193] U. D. Hanebeck and M. Dolgov. Adaptive Lower Bounds for Gaussian Measures of Polytopes. In *Proceedings of the 18th International Conference on Information Fusion (Fusion 2015)*, Washington D. C., USA, July 2015.

[194] G. Kurz, M. Dolgov, and U. D. Hanebeck. Nonlinear Stochastic Model Predictive Control in the Circular Domain. In *Proceedings of the 2015 American Control Conference (ACC 2015)*, Chicago, Illinois, USA, July 2015.

[195] G. Kurz, M. Dolgov, and U. D. Hanebeck. Progressive closed-loop chance-constrained control. In *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, Heidelberg, Germany, 2016.

[196] G. Kurz, I. Gilitschenski, M. Dolgov, and U. D. Hanebeck. Bivariate Angular Estimation Under Consideration of Dependencies Using Directional Statistics. In *Proceedings of the 53rd IEEE Conference on Decision and Control (CDC 2014)*, Los Angeles, California, USA, Dec. 2014.