Risk-aware Model Based Control and Applications on Whey Separation Processes

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op dinsdag 9 januari 2018 om 16:00 uur

door

Muhammed Bahadır Saltık

geboren te Lüleburgaz, Turkije

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt: :

voorzitter:
prof.dr.ir. A.B. Smolders

1e promotor:
prof.dr. S. Weiland

2e promotor:
prof.dr.ir. P.M.J. Van den Hof

Copromotor:
dr. L. Özkan

leden:
prof.dr.ir. I.J.B.F. Adan

prof.dr. J.F.M. van Impe (KU Leuven, België)

dr.ir. A.J.J. van den Boom (TU Delft)

dr.ir. A.J.B. van Boxtel (WUR)

Het onderzoek dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

Risk-aware Model Based Control
and
Applications on Whey Separation Processes

M. Bahadır Saltık

The research presented in this thesis forms a part of the research program of the Dutch Institute of Systems and Control (DISC). The author has successfully completed the educational program of the DISC Graduate School.

The research presented in this thesis was supported by the project “Improved Process Operation via Rigorous Simulation Models (IMPROVISE)”, funding provided by the organization Institute for Sustainable Process Technology (ISPT) and FrieslandCampina.

This thesis was prepared using the LaTeX typesetting system.

Risk-aware Model Based Control and Applications on Whey Separation Processes
by M. Bahadır Saltık
Technische Universiteit Eindhoven
Eindhoven, 2017. Proefschrift.

A catalogue record is available from the Eindhoven University of Technology Library.
ISBN: 978-90-386-4413-4, NUR: 992

Cover design: Kendal Varolgüneş, Eindhoven, The Netherlands.
The cover image contains measurement data from Huygens-Cassini mission of ESA-NASA. Courtesy NASA/JPL-Caltech.
Printed by: Gildeprint Drukkerijen, Enschede, The Netherlands.
Subject headings: Stochastic linear systems / Robust operation with constraints / Whey protein separation / Ultrafiltration membranes / Scheduling of subprocesses.

Contents

Summary

Because of the available computational power, processes that manifest complex behavior can be represented with large-scale and nonlinear models containing differential-algebraic equations. These models are able to represent the physical phenomena for a wide range of operating conditions by using the conservation laws and the physical/geometric structure of the process. Hence, one general goal for current model based applications is to incorporate these rigorous first principles based models into design and daily operation of processes, which is in direct conflict with the computational complexity constraints caused by the real-time operation requirements.

Over the last few decades, model predictive control (MPC) algorithms have become an accepted control approach in the process industry. In control applications, MPC technology achieves desired quality specifications on the outputs by making use of the future predictions of the system evolution. These predictions are generated from a mathematical model. Although increasingly better models are being developed, the true process behavior always differs from the predictions due to disturbances, unmodelled dynamics or unexpected changes in the upstream. To overcome the detrimental effects, one needs to incorporate elements from uncertain future into the predictions. With these predictions, by solving an optimization problem in every decision instant, one can achieve optimal operation, thus higher savings in costs and less utilization of resources, while adhering to the physical or economic constraints. By this way, we effectively control processes for some (or all) possible uncertain elements.

The first part of this dissertation, covered in Chapters 2-3-4-5, addresses the problem of synthesizing computationally tractable and stochastically robust predictive controllers based on large-scale models. To achieve a risk-averse controller, centralized moments of the uncertain predictions are incorporated into the predictive control problem. The dissertation first presents analysis results on the effect of various descriptions of moment based MPC algorithms constructed for different classes of uncertainties. Detailed simulation results are used to discuss the effect of design parameters on the time and frequency domain characteristics of the closed-loop system. Furthermore, for risk-aware MPC problems, we provide explicit reformulation (tightening) of constraints for various classes (bound, affine or quadratic) of constraints. Our results indicate that, an uncertainty-free predictive control problem with reformulating several optimization parameters can guarantee robust operation even if the unknown effects are present in the operation. Furthermore, one should quantify and model the uncertain effects to the finest detail which reflects in the control law formulation. By this way, one can reduce the unnecessary pessimism induced for guaranteeing

robust operation.

The second topic that is discussed in this dissertation, Chapters 6-7-8, is on the modelling, scheduling, monitoring and (low level and batch-to-batch) control of a whey protein separation process. Whey contains high levels of proteins that are extracted via (membrane) separation process apart from other necessary unit operations. A control relevant model of ultrafiltration (UF) membrane units is proposed and validated with the operation data gathered from an industrial plant. The presented UF membrane model uses the series of resistance concept to describe the fouling phenomena, which is the main cause of performance deterioration in separation processes. Through the use of this model, one can track the fouling, and therefore improve the operation efficiency by adjusting the inlet variables accordingly for better operating strategies of UF membranes. We demonstrate that one needs to distribute the required filtration amount among different membrane stages to decrease the accumulation of fouling. By using data based techniques, observability and identifiability analysis of the model is also reported. This allows the practitioners to select sensors to gather measurements, or saved simulation data, that yield the highest information content about the system at hand. Another conducted simulation study is the comparison between the classical and advanced (model-based) control structures. A crucial discussion presented in the dissertation is the learning aspect that has been incorporated into the model-based controllers, meaning that across the distinct batches of operation, we are able to improve the control performance, by using the errors observed in the past batches. Lastly, we provide (high-level) optimal operating schedules of unit operations and the optimal input trajectories for the whey processes.

List of abbreviations

The following abbreviations are used throughout this thesis:

MPC	Model predictive control	RMPC	Robust MPC
OMBA	Online model based applications	SMPC	Stochastic MPC
RTO	Real-time Optimization	DP	Dynamic programming
LMI	Linear matrix inequality	CC	Chance constrained
FPM	First-principles (based) model	LS	Least Squares
BBM	Black-box model	TV	Time varying
pdf	Probability distribution function	TI	Time invariant
(B)WC	(Budgetted) Worst-case	O.P.	Operating point
MMPC	Moment based MPC	RO	Reverse osmosis
UO	Unit operation	NF	Nanofiltration
UF	Ultrafiltration	WPC	Whey-protein condensate
TMP	Transmembrane pressure	VRF	Volume reduction factor
KPI	Key performance indicator	OSC	Optimal sensor configuration
MHE	Moving horizon estimator	DT	Discrete time
EKF	Extended Kalman Filter	IL	Iterative learning
MIMO	Multi input-output	MILP	Mixed integer linear program
RPI	Robust positively invariant	prot	Protein
ISS	Input-to-state stable	comp	Component
(A)RO	(Adjustable) Robust optimization	atm	Atmospheric
DAE	Differential-algebraic equations	emp	Empirical
LTI	Linear time invariant	dyn	Dynamical
LPV	Linear parameter varying	poly	Polynomial
MV	Mean variance	log	Logarithmic
MVS	Mean variance skewness	exp	Exponential
MS-Stability	Mean Square Stability	ret	Retentate
(C)VaR	(Conditional) Value at Risk	perm	Permeate
CSTR	Continuously Stirred Tank Reactor	des	Desired
MSD	Mass-spring-damper	ref	Reference
SOCP	Second order cone program	lin	Linearized
PID	Proportional-integral-derivative	Ent	Enterprise

Nomenclature

t, k	Time indices
x, y, u	Elements of vector spaces
Σ	(Linear) Systems, Covariance matrices
A, B, Q, R	Real-valued matrices
$\mathcal{B}^n$	Hypersphere in $\mathbb{R}^n$
$w, v, \delta, \gamma, \zeta$	Uncertain variables
$\mathbb{P}$	Probability function
$f_{\tilde{w}}(w)$	pdf of random variable w
$c_{ij}(\cdot)$	Constraint function
$\alpha,\ \beta$	Fouling coefficients (-)
a_i, b, c	Static resistance coefficients (-)
$a_{\mathcal{R}}, b_{\mathcal{R}}$	Retention factor coefficients (-)
$\mathcal{C}$	Covariance matrix of parameters (-)
$\mathbb{C}$	Correlation matrix of variables(-)
R^2	R^2 value of estimation problem (-)

$\mathbb{R}$	Real field
$\mathbb{Z}$	Set of integers
$\mathcal{X}, \mathcal{Y}, \mathcal{W}$	Sets
$\mathcal{P}^{(\cdot)}$	MPC problem
$\mathcal{R}(\cdot)$	Risk-mapping
$J(\cdot)$	Cost function
T	Temperature (K)
P	Pressure (Pa)
x	Mass fraction (kg/kg)
ρ	Density (kg/m^3)
m	Mass hold up (kg)
F	Mass flow (kg/s)
A	Area (m^2)
J	Mass flux ($kg/m^2/s$)
R	Membrane resistance (-)
$\mathcal{R}$	Retention factor (-)

Super and Subscripts

M	Mean
MV	Mean-variance
MVS	Mean-variance-skewness
N_p	Prediction stages
adf	Affine disturbance feedback
O	Output
E	Even
I	Initial
f	Feed
r	Retentate
p	Permeate
$drop$	Drop over membrane
$Total$	Total
in	Inlet
w	Water

Chapter 1

Prologue

> As I laid the book down there was a knock at the door, and my stranger came in. I gave him a pipe and a chair, and made him welcome. I also comforted him with a hot Scotch whisky; gave him another one; then still another - hoping always for his story. After a fourth persuader, he drifted into it himself, in a quite simple and natural way.
>
> Mark Twain - A Connecticut Yankee in King Arthur's Court

This thesis presents the research activities conducted on different methods to improve the use of rigorous simulation models in robust online model based applications for process control systems. This chapter initiates the dissertation by motivating the research activities on the general topic of online model based applications and system theoretic properties of these applications in closed-loop operation. Moreover, this chapter provides introductory content on the two structurally separate problems that are commonly faced in many model based optimization activities. The first problem is on the topic of risk-awareness in model predictive control routines. Specifically, this research theme addresses the conflict between the robustness properties of controlled systems versus the computational complexity of the control algorithms. The second theme is directed towards a specific process, a whey protein separation process, in which, separation of whey proteins from other organic components occurs. We demonstrate a number of possible improvements in the operation of this process by incorporating a complex rigorous dynamical model into the offline and online operation. The chapter concludes with the outline of the thesis.

1.1 Research Motivation

For humankind, the demand for prosperous life standards is at the core of the daily activities for many millennium. From clocks for timing the days to the use of robots as dominant labor force, governed mechanisms are devised and used to improve the life standards of

people. Process industry, similar to many other engineering fields, is a massive industry providing products based mainly on the chemical interactions. An essential issue that we discuss in this thesis is the efficient control of processes concerning diverse physical interactions in process control systems to achieve desired goals as outputs under the adverse effects of uncertainty. To provide some context, a process is a physical mechanism that couples some inputs, actuators driving the system called manipulated variables, to some performance or measurement outputs, called controlled variables. These outputs can be the eventual end products released to economic market or intermediate outcomes to be used in another process operation. Furthermore, by control, we mean the rigorous way of analyzing and designing the manipulated variables so as to achieve desired goals in the controlled variables. Control goals are achieved generally through *regulation* (around an operating point) or *servo* action (to track a reference), with a satisfactory robust operation against the known or unknown disturbances or uncertainties. Here robust operation means that the controlled systems[1] are able to reject the disturbances and to compensate inherent modeling biases for achieving the specifications on the outputs. In this context, process control applications are desired to be operating in a resilient way against unknown or unconsidered effects while provide large quantities of products with sustainability concerns. Control synthesis amounts to constructing the input strategy that meets the operational demands, meaning that the realized *controller* performs as desired for a set of specifications or goals on the outputs ([234]). In general, a feedback controller is designed via models representing the processes. Here, models are mathematical abstractions of self-reinforced patterns that are observed from (physical) processes[2].

The computational and communication power allocated for monitoring and control of processes has increased almost exponentially in the recent decades. The inefficient manual production and also the competitive business environment necessitate a highly efficient operation under tighter specifications. Besides concerns about the effect of human activities on the environment are rising, which amounts to regulating many operations to achieve certain safety, reliability and sustainability levels. Traditionally, mathematical models are used to design controllers for describing the behaviour of processes under consideration. To describe the processes better, which, generally, leads to better control solutions, it is expected that more and more complex models will be used in control applications.

Over the last half century, there has been an increased attention on the model-based decision making and operational algorithms. Great achievements in fundamental and applied sciences lead to a much broader understanding of the optimization theory ([61])[3]. This drastic change in theoretical understanding is also assisted by various technological breakthroughs. In this half century, humankind managed to construct increasingly robust and efficient computing and communication devices. This allowed digitalization and coopera-

[1] A controlled system is an interconnection of at least two subsystems, generally called as the controller and the process itself. The interconnection constraints the time evolution of process, thus, with rigorous control design, allows the closed-loop system to react as desired.

[2] In here we do not provide an in-depth discussion on the historical trajectory of control theory, however interested reader may reach to [11, 249, 343] from openly available sources.

[3] Some examples can be given as the concept of convexity, the KKT theorem for constrained optimization problems, Pontryagin's Maximum Principle and Bellman's Principle of Optimality for dynamical optimization problems and *Simplex* and *Interior Point* methods for the numerical techniques to conduct the optimization.

tive computing, which are crucial for handling complex problems commonly resulting from large-scale optimization based algorithms that include multiple decision makers apart from uncertain factors. A recent and radical example can be given as the successful optimization based design of efficient fusion reactors ([56]), see Figure 1.1. Control practitioners

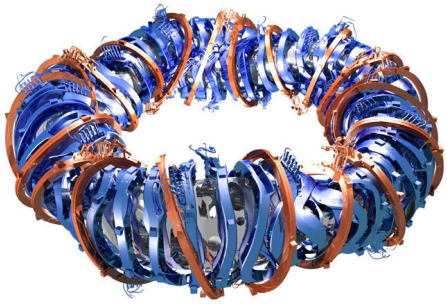


Figure 1.1: Model-based optimization routines are used to locate coils of the magnetic cage system of Wendelstein 7-X. The image is taken from the website of Max-Planck-Institut für Plasmaphysik (IPP).

and researchers, commonly acting under business drives, are using these tools to operate interconnected networks of processes, each of which can be described by complex dynamical behaviour, under tight error margins. It is imperative for control practitioner to make use of high-fidelity models to develop monitoring routines, *the observers*, or automatic controllers that are able to reach these tight specifications by diminishing undesired sensitivities undesired effects.

1.1.1 Online Model-based Applications in Process Control Area

In the process control area, the online model-based applications (OMBAs) include, but are not limited to, the soft-sensor or advanced process control implementations ([266]). Some OMBA examples are:

- The model predictive control (MPC) technology (or its predecessors) is used to formulate a practically elegant solution to the question: "how to practically control processes to achieve (sub-)optimal operation under actuation saturation or operational limits, such as computational constraints?" Relying heavily on the predictions generated from the process model, MPC technology optimizes over future trajectories while explicitly incorporating the limitations as constraints. Furthermore, MPC technology allows the practitioner to decide on the problem's computational complexity, by adjusting the problem size, the number of prediction stages. MPC based methods have already been applied in many practical environments, e.g. [288, 289, 290].

- Monitoring of processes, in its general form, amounts to the problem of inferring trajectories of (un)measured variables from the sensory data. Soft-sensor implementations keep track of various performance indices which are then used for control purposes ([299],

[154]). The Kalman filter, with only two explicit algebraic update equations, is a highly effective example of model-based process monitoring tools.

- Tactical decision making routines or schedulers are also examples of OMBAs. The scheduling decisions are commonly used in interconnections of processes to optimize the production and further diminish the (raw material, transportation or energy) costs by using large-scale and static models.

Some OMBA examples relevant to the topic of this dissertation are;

- The use of reservoir models in decision making environments within the oil-extraction process to maximize the economic gains ([336]) or;
- The batch crystallization processes to minimize the effect of disturbances and mismatches in the initial conditions ([359]).

Model-based decision making routines are also used in many other practical instances that are not strictly under the control domain, such as the emergency scheduling ([134]) or portfolio management ([331]).

One common aspect of the mentioned OMBA activities is the separation of a centralized and global problem into several different subproblems. Within the control domain applications, control hierarchy is used to separate the *universal* problem into multiple problems associated with different aspects of the plant operation in an efficient way ([226]). These layers deal with the problems that have different economic incentives, sampling/decision frequencies or models. Among the same layer, different subprocesses might communicate and interact with each other to cooperate and improve the process operation. In Figures 1.2a-1.2b we visualize and shortly note some observations on the commonly perceived control hierarchy in process control implementations.

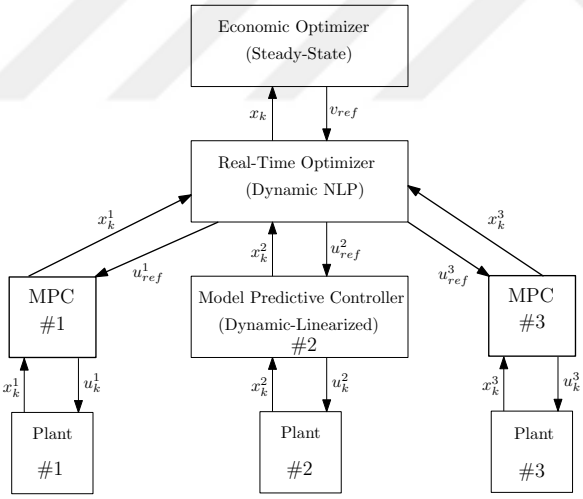
In all mentioned cases of OMBAs, one common and essential element is the implicit process model. As the models describing the process are further improved and validated, we receive and process better predictions of the process. In the current line of reasoning, the models are distinguished by the model structure selected as our priors. As an ideal case, models that are described by fundamental laws of physics are referred to as the white or first principles (based) models (FPMs). On the other hand, if the model is solely relying on the inherent causal/correlated relations of recorded input and output data, we call these models black-box models (BBMs). In fact, there are many more successful applications of BBMs, due to reliable and computationally cheap system identification methods ([216]).

In this dissertation, we design controllers for processes by making use of globally valid input-output (operating) space, and, if applicable, switch between different specifications efficiently within the run-time ([266]). The BBMs are describing processes locally which might lead to problems in controlling systems that demonstrate nonlinear, time-varying, fast-and-slow (with coupling in between) behaviour. The FPMs, by their construction, is able to exploit the detailed process knowledge within the OMBA algorithms[4]. This allows FPM-based OMBAs to be used in batch type of processes[5], which are common in process

[4] In many cases FPMs rely on different physical variables which are naturally used also in controllers.

[5] The batch processes, in this dissertation, have a start and a final time for the operation. The design of observers or controllers for batch processes is known to be notoriously difficult, [224], since the behaviour of the process varies quite drastically during the operating window.

(a) Conceptual control hierarchy.



(b) An implementation example of control hierarchy.

Figure 1.2: Representation of process control hierarchy depicting the different goals and resources at each layer. Throughout the control hierarchy, mainly economic benefits but sometimes also safety or performance requirements pushes the process operation to become more reliable and reproducible.

control industries. Contrary to the presented reasoning and due to various reasons, extensive and expensive model development effort, lack of validation tools or computational issues, FPMs are not preferred in OMBAs frequently. Instead, to overcome these drawbacks grey-box models are used in practice. These models can effectively incorporate the complex interactions between the physical variables through various number of black-box components, generally by (nonlinear) regression functions to represent fast or complex phenomena while containing also (simple) first principle laws. A table highlighting the complexity of model

versus the regression quality between BBMs, FPMs and grey-box models is provided in Table 1.1.

	Fitting Accuracy	Computational Complexity	Modeling Effort
BBM	Low	Low	Low
Grey-box	Low or High	Low	Low or High
FPM	High	High	High

Table 1.1: Comparison in between the model types.

Then, constructing computationally efficient applications using physics-based models in real-time operation is a goal of this thesis. We base our discussion on various offline and online model-based applications for an industrial case study, a whey protein separation process, as a research theme of this thesis.

1.1.2 Uncertainty in Models and Model-based Applications

The effect of uncertainty is an important aspect in OMBAs and hence the resulting control performance. In any case, the prediction model will be different from the true process at hand. This means that the predicted trajectories from the uncertainty-free (nominal) model will be wrong, i.e., we assume that the uncertainty is inherent to the nature of the process ([71]). In order to take action against the unknown effects, one can model the uncertainty to reflect it in the predictions[6]. Once the predictions, incorporating uncertain factors, are generated, the control action is evaluated according to the risk associated with these predictions. One way to evaluate the risk is to assume the worst possible outcome out of all scenarios ([115]). The controllers designed with this type of deterministic guarantees against uncertain effects, in general, lead to a heavy loss in performance of the controlled system. Hence modeling decisions taken on the effective uncertainty and the distribution of its realizations is an important design step ([253]). The analysis and treatment of uncertain effects in optimization problems are open research areas in various domains, see [35] for a technical introduction to the topic.

We know that the set of possible uncertain effects changes the resulting closed-loop behaviour. Incorporating unrealistic uncertainty scenarios into the prediction models cause the controller to reduce the process sensitivities from inputs to outputs, which is hampering the performance, and therefore highly undesired in industrial operations[7]. Operating processes more efficiently amounts to steering the operation towards economic or operational constraints. As the operation approaches the constraint region, the importance of risk-aware control methods is increasing, see Figure 1.3. This figure visualizes the distribution of predicted outputs for three different controller scenarios, in case (a) a pessimistic controller is used, since the distribution of outputs are far away from the constraint; while for case

[6]The elusive nature of uncertainty is also effective on the modeling and quantification of it in the dynamical evolution. However, here we do not discuss this problem.

[7]Since business objectives, in general, do not specify the robustness requirements and there is no incentive from industry for *complicated* control design, there is a highly contrasting gap in between the theoretical achievements and the simplicity of the practical implementations.

(b) possibly an optimization based linear controller is used which shifts the distribution towards the constraints. The last case, (c), makes use of risk-aware decision making routines, thus is able to improve the performance of safe[8] tail of the distribution while reducing the possibility of constraint violation for the tail realizations that violate the constraint. One

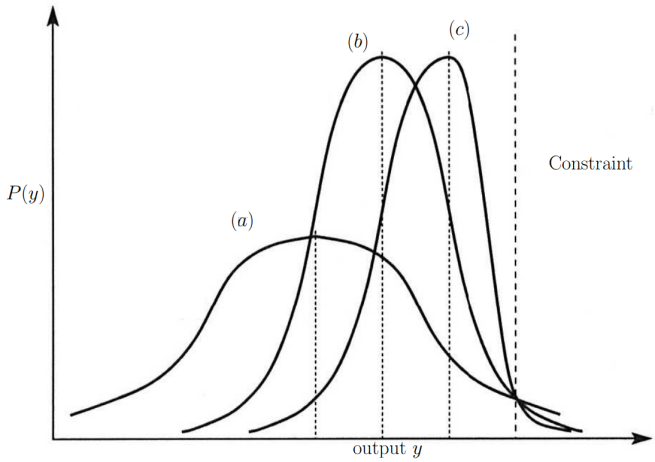


Figure 1.3: Probability density function of an output variable controlled with three different controllers. Taken from [226].

can describe these distributions by (reachable) intervals or probability distribution functions (pdfs). However, high-performance risk-aware control methods are not developed yet for general risk functions. This is due to *pessimistic* performances and the high computational complexity of the current methods. Within this thesis, we address the possible shortcomings of robust predictive control methods and propose a different strategy for evaluating the uncertain effects in MPC problems.

1.2 The Research Goal, Themes and Questions

Complexity of a model-based optimization problem with uncertain elements depends on the complexity of the nominal (uncertainty-free) model[9], the uncertainty model[10] and how the uncertainty is evaluated in decisions[11]. Development of computationally simple and reliable applications based on rigorous models for process control systems determines the research goal of this thesis;

Research Goal: To improve the use of rigorous models in robust model-based applications for process control systems.

[8] Here, safe describes the realizations, or instances, that occur far away from the constraints.

[9] A rigorous process model which does not incorporate any affect of uncertainties can be already too complex for OMBAs.

[10] Once the uncertain effects are incorporated into the model, the optimization problems grows considerably in size.

[11] Re-expressing the robust optimization problem depends on the uncertainty measures, called as *price of robustness* ([44]), which amounts to balancing the reliability against uncertain effects versus the performance.

In this dissertation we discuss two possible ways to address the research goal. The first theme of this dissertation is on the risk-aware evaluation of process predictions in MPC applications. The second theme is aligned with a practical application and addresses the use of rigorous models in a whey protein separation process. By making use of the control hierarchy and a master model, Figure 1.2a, we consider scheduling, monitoring, and control problems to improve the separation operation performance. In the next two subsections, we discuss these two themes separately and present research questions that are addressed during the research.

1.2.1 Robust Predictive Control and Risk-aware Operation

In this dissertation, we focus the discussion on the development of *practical* and *risk-aware* model predictive controllers with desired closed-loop performance and computational properties. MPC is commonly used in process industry for constrained multivariable optimal control ([21]). The performance of these controllers depends to a large extent on the quality/validity of process models ([290]). In addition to the inaccurate identification of model parameters, noise or model simplifications ([164] and references therein) also introduce mismatch between the process observations and the model predictions. Yet in many cases, MPC problems derived through the use of nominal dynamics are sufficiently effective. Hence here we seek to develop efficient methods to include uncertain effects without deteriorating the control performance. The essential problem within this first research theme is the conflict between the risk allocation of uncertain predictions and the computational complexity to describe this risk function[12]. The risk-awareness concept is the ability of the controller to assess the detrimental effects of uncertain predictions and incorporate some of these trajectories into the control law formulation.

In the last couple of decades, different MPC algorithms have been introduced to achieve the desired control objectives while reducing the effect of uncertainty ([133]). We can broadly classify these algorithms into two classes;
i) The worst case based techniques (WC-MPC) ([185]) where under any effect of predefined maximum level of uncertainty, the process variables do not violate the desired specifications or;
ii) The chance based MPC algorithms ([326]), where the specifications are softened depending on the chance of being off the desired levels. By this way violations are cast to be rare events.

Both of these formulations have their own drawbacks; the common disadvantage is the complexity of the robust counterpart problem[13]. Since the probabilistic interpretation, such as the chance based MPC, incorporates WC-MPC as a special case, a stochastic approach towards robustness is preferred in this dissertation.

One major issue with these methods is that the distribution functions of uncertain variables are difficult to calculate in large-scale systems. Thus we address ways to construct robust MPC problems that are computationally simple;

[12] The deterministic problem which already incorporates uncertainty into its formulation is named as robust counterpart problem.

[13] Issues such as pessimism in the closed-loop operation are introduced and discussed in Chapter 2

1st **Research question:** What type of methods can be used to construct predictive controllers based on rigorous large-scale models with adjustable robustness and desirable computational properties?

We detail more on the research question with some subquestions as follows:

- Incorporating uncertain effects into the control applications is to measure the effect of uncertain predictions in the outputs by evaluating their risks. This step determines both the pessimism associated with desired robustness and the computational complexity. Then a valid question is:

  Research Question 1.1: What are the techniques that transform uncertain MPC problems into robust counterpart MPC problems? What are the computational complexity properties of these techniques?

  We address this question in Chapter 2, in which we classify the methods, and also in Chapters 3, 4 and 5 we address this question for a specific technique, the moment-based MPC formulation.

- In general, we aggregate various different sources of uncertainties on the outputs. However, these effects causing the uncertain predictions, such as perturbations in predictions, lack of sufficient measurements or process-prediction model mismatches cause structurally different prediction errors. Thus;

  Research Question 1.2.1: What are the control theoretic interpretations of the robust MPC controllers that guarantee robust operation for different types of uncertainty sources?

  Closely related to the previous question, analyzing and evaluating the distribution of uncertain effects requires an extra step in the uncertainty modeling phase. The main reason is that there are no tools to describe the relatively inconsequential[14] uncertain effects;

  Research Question 1.2.2: What is the effect of detailed modeling of the uncertainty space to the control actions in MPC driven closed-loop operation?

  We address these questions in Chapters 3 and 4.

- MPC, being an optimization based control strategy, requires a cost function that needs to be minimized. Many different possibilities exist for selecting the cost function, quadratic or polytopic (norm based) functions being the control theory oriented ones. We parameterize these functions by weighting terms that adjust the relative importance of variables. Commonly, these weighting terms are selected in an *ad-hoc* way, heuristic techniques are used to cross-weight different variables. We expect that selecting weighting terms in an algorithmic way might lead to increased closed-loop performance at the expense of robustness properties. Furthermore, any technique addressing the cost selection problem should also be accessible to the operators. In process control applications, operators prefer to use long-known and reliable MPC routines to control the processes. The experience developed

[14] In this dissertation we assume that high consequence realizations of uncertainty, *uncertain impacts*, are not existent within the operation. Hence, in the context of this thesis, the uncertain effects are not causing harsh or unrecoverable events, but act as mere perturbations along the dominant and known process dynamics.

along the MPC implementations also pushes practitioners to rely on basic MPC algorithms depending on nominal models.

> **Research Question 1.3:** Which cost functions (or optimization parameters) improve the robustness or disturbance rejection properties in MPC problems? Are there nominal MPC problems that guarantee desired robustness properties for uncertain dynamical systems?

We address these questions in Chapters 3 and 4.

- One of the most important aspects of MPC based closed-loop operation is the explicit constraint handling capabilities. However addressing the constraint satisfaction under the effect of uncertainties necessitates a risk-based re-evaluation.

> **Research Question 1.4:** What are the robust constraint satisfaction properties of MPC based closed-loop operation by including stochastic models of uncertainty?

We address this question in Chapter 5.

In the light of these questions, this thesis advocates the use of statistics, the finite order centralized moments, of the state predictions to calculate the MPC control actions. We present a novel MPC strategy, the so-called moment-based MPC (MMPC), which considers the expectations and variances of the predicted trajectories. By making use of variance (or higher order statistics) of the state predictions, one can effectively back-off the operating conditions as a function of the process dynamics and the uncertainty model. This approach improves the robustness or disturbance rejection properties of the closed-loop system. Even if different types of uncertainty models are considered, it is shown that the MMPC has computationally desirable properties.

1.2.2 Online Model-based Applications in a Practical Case Study: Whey Protein Separation Process

The second theme of this dissertation is directed to improve the use of rigorous models for OMBAs by investigating possible issues and drawbacks in whey protein separation process. This process consists of a network of unit operations (UOs), in general including membranes of different pore sizes, evaporators, and dryers, see Figure 1.4. Our investigation of

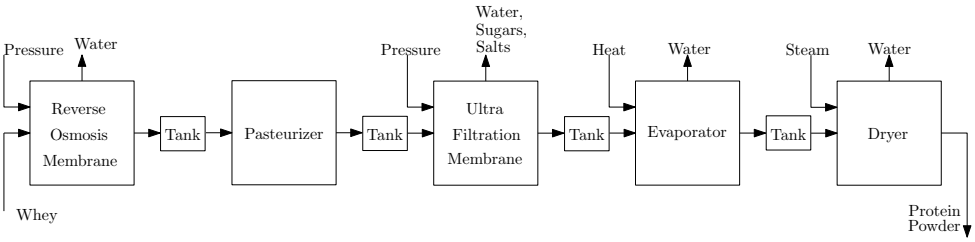


Figure 1.4: The block scheme of whey protein separation process.

using rigorous models for the whey protein separation process mainly deals with the ultrafiltration (UF) membrane UO. Specifically, we address the dynamical modelling and optimal

operation of the UF membranes in this process. We also direct attention towards the optimal operation of the rest of the UOs, i.e., we consider plant-wide optimization problems that address the safe operation properties. Next, a short introduction on the process and the identified research directions are stated.

The whey (protein separation) process consists of the reverse osmosis membranes, UF membranes, evaporators, and dryers. Among these UOs, the dynamics of evaporators and dryers are known to be considerably faster than the dynamics of membrane units, hence static behaviour from inputs to outputs is assumed to be representing their effects. However, the input-output operating points of membrane units vary harshly during the operation. Due to this reason, we need to incorporate dynamical effects into the rigorous model of UF membranes.

The UF membranes are pressure driven separation processes, in which the pressure difference, the transmembrane pressure (TMP), between the two mediums separated by the membrane wall causes some particles to freely pass through the wall while the larger particles in the inlet stream can not move in between these mediums. More precisely, relatively small molecular size components, such as the water or sugars in whey, are driven from the feed inlet to the so-called permeate outlet of the membrane, see Figure 1.5. The remain-

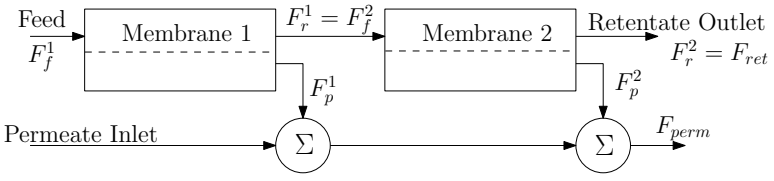


Figure 1.5: The interconnection scheme of two UF membrane stacks.

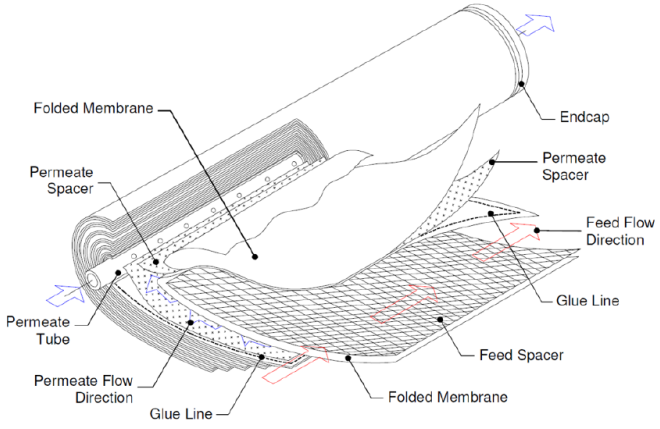


Figure 1.6: A visual picturing the membrane unit's physical parts.

ing components are relatively large size components that are not able to pass through the

membrane wall such as the dairy/cheese proteins in whey. These components leave the membrane unit from the retentate port. Hence the concentration of large molecular size components at the retentate port is high while the permeate is rich in small molecule sized components. Due to the gradual deterioration of separation efficiency of the UF membrane unit, the process is turned off for cleaning purposes, casting the process to be a batch process.

The whey components and their concentrations are effective at the resulting product's properties. Furthermore, the deterioration of membrane separation performance, called as fouling, directly affects the eventual output. Due to these effects on the end product, the whey protein powder, the OMBAs that we discuss in this dissertation are targeting the monitoring and control of fouling accumulation. For high-efficiency operation of the whey protein separation process and the UF membranes, we formulate the following research question;

2nd **Research question:** What are the potential benefits of using offline and online model-based applications for the UF membrane units in whey protein separation process to achieve specified performance goals?

Some research questions and practical goals that are considered in this thesis are stated as follows;

- The UF membranes have complex, nonlinear and highly coupled fast-and-slow dynamics. During design models that contain partial differential equations, which express the pressure distribution along the membrane unit in terms of spatial coordinates, are used. Another common way of describing membrane units is using the static regression functions for capturing the nonlinear behaviour. In this thesis, we desire to use a rigorous UF model that is sufficiently complex to incorporate different membrane operating points. Furthermore, we need to keep the model complexity at a certain level to allow us to design real-time model-based controllers.

  Research Question 2.1.1.: What are the physical laws that are needed to describe the complex behaviour of UF membranes while keeping the model size and complexity at a minimum which allows the designer to use the model in monitoring or control applications?

  Similarly, fouling is the main reason of gradual degradation of process efficiency in UF membranes. Thus, it is of high importance to keep track of fouling during the operation. However, membrane fouling is not a physical variable but an aggregate effect that is deduced from the performance deterioration. It is difficult to directly describe or measure the 'total' fouling of a membrane during the operation. Here we make use of the developed model, the Research Question 2.1.1, to implement soft-sensors monitoring the accumulated fouling during the operation.

  Research Question 2.1.2: Which measurable variables should be recorded to correctly reconstruct the fouling?

- An important aspect of UF membrane operation is the optimal operating trajectories along the batch time window, the so-called 'golden batch'. With the rigorous UF membrane

model, one can analyze the operating input-output trajectories that are able to reach desired specifications for the output products while decreasing the accumulation of fouling and distributing the required workload of whey separation in between different operating membrane units. Thus;

Research Question 2.2: What are the best operating strategies with respect to on-off conditions of membranes, the inlet flow and pressure values for each and every membrane unit in the overall process to achieve the desired goals at the output with reduced fouling?

- In this thesis we desire to implement OMBAs for UF membranes using the rigorous model. Hence, we address various online applications concerning the UF membrane process, such as;

Research Question 2.3.1: What are the best monitoring (soft-sensor) strategies that can be implemented in real time and provide information on variables and parameters of the model?

Research Question 2.3.2: What type of low-level control structure should be selected for efficient operation with UF membrane units?

Research Question 2.3.3: Can we improve the operation efficiency and decrease the effects of plant-model mismatch by incorporating a learning action that considers errors among the previous (recorded) batches?

- Since the products of whey protein separation process are organic materials, the process is subject to strict regulations, requiring the processing to reliably dinish quickly. The whey protein separation process consists of multiple batch units and these UOs can be independently shut down for various reasons such as cleaning or maintenance. These aspects lead to using buffer units to accumulate the processed material after each UO, which is an important cause of increased residence time. Thus the scheduling of the unit operations in the process has a considerable importance. In this thesis, we approach the scheduling of UOs in whey process from a safety perspective, meaning that we seek for operation such that the buffer tanks are not overflowing or depleted and the throughput is as large as possible. However, many scheduling routines to achieve this goal are quite computationally demanding to solve since multiple UOs need to be optimized over on-off actions.

Research Question 2.4: What type of scheduling models can be used to generate computationally simple scheduling problems which guarantee safe operation for a network of UOs that show on and off behavior in a cyclic fashion?

1.3 Thesis Outline

This dissertation addresses two distinct phases of model-based control and estimation applications in the direction of questions formulated in the previous section. For the first theme,

the risk-aware model predictive control algorithms with desirable computational complexity properties, we dedicate a substantial content of this dissertation to address, first, the drawbacks in the current robust MPC strategies in the academic literature and then we develop a novel approach, the moment-based MPC formulations, to diminish the associated pessimism; and lastly we analyse the theoretical properties of these MPC formulations. In detail:

- In ***Chapter 2***, we present an extensive outlook on the available literature on MPC, comparing the robustness properties of nominal, robust and stochastic MPC formulations. These different strategies are then implemented in different simulation examples.

- In ***Chapter 3***, building up from the observations stated in the literature review, we present a computationally efficient robust MPC method by making use of the statistical information of uncertainties affecting the dynamical system to explicitly reformulate the robust counterpart problems of cost functions[15]. We first address the case of linear systems that are perturbed with additive noise, for which we provide the robust counterpart MPC problems for the first three centralized moments which are computationally equivalent to nominal MPC problems with different cost functions. In this part we treat the cases with different assumptions on the structure of dynamics or uncertainty or different control goals, such as we consider (a) state regulation and stability problem; (b) state/output tracking performance evaluation; (c) the output feedback MPC case; (d) rate MPC formulation; (e) non-Gaussian uncertainty characteristics for the uncertainties. A detailed analysis of closed-loop properties is also presented.

- In ***Chapter 4***, we direct our attention towards more challenging type of uncertainties, the plant-model mismatch for linear systems case, which is also called as multiplicative (in state-space representation) uncertainty. The robust counterpart problem for plant-model mismatch case is formulated for linear systems with time-varying or time-invariant uncertain effects. Similar to the previous case a detailed analysis of closed-loop properties is addressed. Since the resulting nominal MPC problem lacks some desirable numerical properties, we provide two heuristic MPC formulations, with observed behaviour similar to the true moment-based MPC formulation.

- In ***Chapter 5***, we expand our discussion on moment-based predictive control approach towards the robust reformulation of the uncertain constraints. We treat different classes of constraint functions that are commonly used in MPC applications, such as bound constraints on states or control actions, or zone (polytopic) type of constraints on outputs. We discuss the effect of uncertainties with an unbounded domain on the (probabilistically satisfied) recursive feasibility property of MPC controlled system.

- The second theme of this dissertation starts with ***Chapter 6*** which is on the development of a dynamical simulation model to be used in model-based applications for an industrial whey protein separation plant. We start our discussion on whey protein separation process by addressing the modelling problem for the UF membrane process. To keep the resulting model simple enough for real-time operations while providing crucial latent information

[15] In this chapter, we do not yet consider constrained MPC problems.

of the process that is not easily accessible, we develop a grey-box model for UF membranes with the help of research group in *FrieslandCampina* Wageningen and industrial data provided by *FrieslandCampina* Workum groups. Furthermore, we make use of this UF membrane model for offline analysis of the process itself. First, we compare different operating strategies to demonstrate that the optimal operation (in the sense of longest batch time) of membranes is achieved when all of the membranes are kept operational all throughout the batch, with an even distribution among membranes of filtration to reach specifications. Secondly, we turn our attention to soft-sensing and sensor configuration problem, in which we make use of data obtained from numerous simulations to select the measurement channels to cast the process under investigation observable and/or identifiable.

- ***Chapter 7*** addresses the problem of online model-based applications for the UF membrane process. Using the developed simulation model, we implement observers and controllers that are crucial for high-efficiency operations. First, we demonstrate the effectiveness of selected sensor channels and compare different filter design techniques, such as extended Kalman filter or moving horizon estimation (MHE) based filters. Then we move towards the low-level control design and construct different control techniques for UF membrane system. We compare PID based control methods designed in different controller structures with the MPC based controllers. Since the membrane processes are inherently (semi-)batch hence dynamic, we show that high-efficiency operation can be achieved by operating the membranes in different pressure levels. Lastly, we incorporate a learning scheme into the control loop, in order to further eliminate the possible effects of modeling issues and persistent disturbances across the batches that deteriorate the closed-loop performance.

- In ***Chapter 8*** we return to the full scale whey protein separation plant by the presenting our results for safe scheduling of the UOs within the process. We demonstrate an efficient way of formulating safe indefinite or steering schedules for a generic plant with interconnected processes.

- We finalize the dissertation with ***Chapter 9*** which contains the reflections on the theoretical or practical results and future research suggestions on the topics within this thesis.

1.4 List of Publications Based on the Research Activities

- M.B. Saltik, L. Özkan, M. Jacobs, A. van der Padt. Dynamic modeling of ultrafiltration membranes for whey separation processes. Computers and Chemical Engineering, 99, 280-295, 2017.

- M.B. Saltik, L. Özkan, J.H.A. Ludlage, S. Weiland and P.M.J. Van den Hof. An Outlook on Robust Model Predictive Control Algorithms: Reflections on Performance and Computational Aspects. Accepted for publication in *Journal of Process Control*.

- M.B. Saltik, L. Özkan, S. Weiland. Moment-based Model Predictive Control for Linear Systems Part 1: Additive Perturbations Case. Submitted and under review for publication in Mathematics of Control Signals and Systems.

- M.B. Saltik, L. Özkan, S. Weiland. Moment-based Model Predictive Control for Linear Systems Part 2: Multiplicative Perturbations Case. Submitted and under review for publication in Mathematics of Control Signals and Systems.
- M.B. Saltik, L. Özkan, S. Weiland. Constraint Tightening in Moment-based Model Predictive Control. Submitted and under review for publication in International Journal of Robust and Nonlinear Control.
- X. Cao, M.B. Saltik, S. Weiland. Optimal Hankel Norm Model Reduction for Discrete-Time Descriptor Systems. Submitted and under review for publication in Journal of the Franklin Institute.
- X. Cao, M.B. Saltik, and S. Weiland. Hankel model reduction for descriptor systems. In proceedings of IEEE 54th Annual Conference on Decision and Control (CDC). IEEE, 2015.
- M.B. Saltik, N. Athanasopoulos, L. Özkan and S. Weiland. Safety Analysis for a Class of Graph Constrained Scheduling Problems. In proceedings of IEEE 54th Annual Conference on Decision and Control (CDC). IEEE, 2015.
- M.B. Saltik, N. Athanasopoulos, L. Özkan. Enterprise-wide Optimization. In proceedings of 34th Benelux Meeting on Systems and Control, 2015, Lommel, Belgium.
- M.B. Saltik, L. Özkan, M. Jacobs, A. van der Padt. Optimal Start-Up and Operation Policy for an Ultrafiltration Membrane Unit in Whey Separation. In proceedings of 26th European Symposium on Computer-Aided Process Engineering, 2016.
- M.B. Saltik, L. Özkan, S. Weiland and P.M.J. Van den Hof. Optimal Sensor Selection Problem for Membrane Separation Systems. In proceedings of 11th IFAC Symposium on Dynamics and Control of Process Systems, 2016.
- M.B. Saltik, L. Özkan, J.H.A. Ludlage, S. Weiland, P.M.J. Van den Hof. On the Moment-based Robust MPC Formulations. In proceedings of 2016 Annual meeting of AIChE. AIChE, 2016.
- M.B. Saltik, L. Özkan, S. Weiland, J.H.A. Ludlage. Moment-based Model Predictive Control for Systems with Additive Uncertainty. In proceedings of 2017 American Control Conference. IEEE, 2017.
- R. Zhang, M.B. Saltik, L. Özkan, Study of Moment-Based MPC formulations and their Connection to Classical Control, In proceedings of 2017 Annual meeting of AIChE. AIChE, 2017.
- S. van Gameren, M.B. Saltik, L. Özkan, S. Weiland. Recovery scheduling for industrial processes using graph constraints. In proceedings of 27th European Symposium on Computer-Aided Process Engineering, 2017.
- X. Cao, M.B. Saltik, S. Weiland. Optimal Hankel Norm Approximation for Continuous-Time Descriptor Systems. Submitted to 2018 American Control Conference.

1.5 Information on the Project

The research discussed in this dissertation is part of the project named "Improved Process Operation via Rigorous Simulation Models (IMPROVISE)". This project is funded by the Institute for Sustainable Process Technology (ISPT) project cluster "Process System Engineering/Process Control Cluster". The following partners took a collaborative part in the IMPROVISE project:

- Institute for Sustainable Process Technology (ISPT),
- FrieslandCampina,
- DSM,
- Corbion,
- TU/e.

Chapter 2

Risk-aware Model Predictive Control in Literature

> To understand the future to the point of being able to predict it, you need to incorporate elements from the future itself.
>
> Nassim Nicholas Taleb - Black Swan

We start the technical content of this dissertation with presenting a generic MPC problem, introducing the common ways of incorporating and quantifying the uncertainty in MPC problems and stating the different methods to find the robust counterpart problem in Section 2.1. Section 2.2.1 discusses the robust MPC methods and contributions that are introduced with deterministic treatment of uncertainty, either worst-case or uncertainty budget approaches. The stochastic MPC approaches towards uncertain dynamics are discussed in Section 2.2.2. We present the moment, probabilistic and randomized MPC contributions and show the possibility of incorporating the theory of risk into the MPC setting. The effectiveness (closed-loop performance) of the methods from literature are demonstrated by means of simulation examples in Section 2.2.3.

2.1 A Short Introduction on Risk-aware Model Predictive Control Problem

Model predictive control (MPC) technology is a mature research field developed over four decades both in industry and academia addressing the question of (practical) optimal control of dynamical systems under process constraints and economic incentives. Its popularity is mainly attributed to two significant properties of MPC algorithms; first one is the (explicit) constraint handling capabilities while providing (sub-)optimal operation, see, e.g., [199, 237, 240]; and the second superiority is the ease of extending the algorithms to multi-

[0]Substantial content of this chapter is also published or presented in 'M.B. Saltik, L. Özkan, J.H.A. Ludlage, S. Weiland and P.M.J. Van den Hof. An Outlook on Robust Model Predictive Control Algorithms: Reflections on Performance and Computational Aspects.'

input multi-output (MIMO) systems. Many different approaches were developed, such as; Model Algorithmic Control in 1978 ([307]), with finite impulse response models, Dynamic Matrix Control in 1980 ([97]), with step response models, Generalized Predictive Control in 1987 ([91]), with transfer function models. Lately, MPC methods developed by considering the state-space models have become the standard way of formulating predictive control problems. Throughout the different algorithms, however, the essence of predictive control is the same and can be stated as, [301], optimizing over manipulated inputs to control the forecasts of future process behaviour. Stated rigorously, [242, 250], MPC is a form of control in which the current control action is obtained by solving, at each decision instant, a finite (or infinite) horizon open-loop optimal control problem. In this technique an optimal control sequence is obtained by using the current state of the plant as the initial state of the plant and the first control in this sequence is applied to the plant, while at the next decision instant the whole procedure is repeated.

The process of selecting an optimal control action can be summarized in two distinct steps ([229, 230]),

i) shaping the beliefs of future output performances (forecasts);

ii) the choice of to-be-applied control action as a function of these forecasts.

A general approach to obtain output forecasts is through dynamic models describing the process behaviour. During the initial development of MPC, empirical linear input-output models were utilized. If the operating window is relatively small, such models are proved to be sufficient. However, if the operating conditions vary drastically, e.g., batch processes, then nonlinear models should be used, which effects the complexity of the MPC problem[1]. In either case the developed models will be far from perfect; leading to mismatch between the forecasts and the true behaviour. As a result, the commissioned MPC controllers are kept non-operational frequently due to the model deterioration or lack of maintenance of the model, ([7]). It is both natural and logical to include the effect of (modeled) uncertainty into the prediction model, hence into the optimal control action[2]. However, uncertainty also radically effects the optimal control actions in closed-loop predictions, casting them to become pessimistic (or aggressive), hence the resulting performance levels are also effected ([33]).

A well established way to overcome or reduce the effects of uncertainty is by applying feedback techniques. In many instances, robust control theory ([100]) provides *sufficient* tools for achieving robust operation. However, this design choice often leads to over-utilization of the available resources as it might not be necessary to execute a pessimistic control law at each time instant. For industrial applications, especially in process control industry where economic concerns are directly effecting the operation decisions, the pessimistic control methods are in general rejected and robustness is achieved in an ad-hoc manner ([236]). In recent years, a huge effort has been put in developing computationally

[1] Here we do not consider the difficult questions of how and at which complexity level the process model should be constructed. We refer the interested reader to [101, 102] as introductory discussion on modelling the uncertain behaviour.

[2] In different words, selecting a control action on the basis of the nominal forecasts leads to undesired operation due to definite dispersion from the expectations in the controlled variables.

efficient (or tractable) and less pessimistic (or adjustable) robust optimization tools that have parameter ambiguity and stochastic uncertainties within the formulation of the optimization (equivalently MPC) problems ([128]).

It is important to distinguish three different robustness aspects of MPC algorithms in the way of treating uncertain effects,

1. robust feasibility,
2. robust stability,
3. robust (closed-loop) performance.

The robust feasibility is about the constraint satisfaction in the face of uncertainty, while the robust stability is tracked via the cost function through Lyapunov based stability arguments. We have a considerable understanding on robust constraint satisfaction or robust stability while the interplay between the uncertainty and the closed-loop performance is yet to be rigorously analyzed. Although there exist some methods to synthesize predictive controllers that operate in a computationally acceptable way ([180]), many of the current robust MPC methods lead to computationally challenging optimization problems, while causing unacceptable levels of performance deterioration. The performance deterioration, or even the total absence of performance, due to overly conservative methods is causing a gap between academic works and industrial implementations. High performance is achieved if the uncertain effects are compensated when it is required, while robustness requirements demand to act in a pre-emptive manner. Hence incorporating *only the necessary* uncertain process predictions into the control action by incorporating risk management techniques is of great interest for predictive control applications.

Combining robust control and predictive control regarding the robust constraint satisfaction, stability and performance aspects with quantitative guarantees is still an open problem. There are a multitude of techniques, detailed in the next sections, to reshape robust MPC (RMPC) (or similarly stochastic MPC (SMPC)) problems. The main dilemma is due to the open-loop nature of predictions, leading to loss of incorporation of *future* uncertainty into the control actions. Dynamic programming (DP) techniques provide a way out of this problem, however the *curse of dimensionality*, specifically for moderate or large-scale systems or uncertainty spaces, drastically effects the computational aspects, see [207].

Another important point regarding the industrial acceptance of RMPC algorithms is the computational aspect. It is a requirement that RMPC problems should be consisting of relatively simple and reliable algorithms resulting in desired performance levels of industrial needs ([236]). In the case of industrial implementation, the algorithms should be ([132]);

- easy to interpret and interact by the operator with sufficiently large amount of information on controlled variables and/or constraint violation risks, ([126]);
- relatively simple to solve computationally in repeated fashion, since the industrial platforms are generally not tailored for highly complex and efficient numerical algorithms.

The RMPC techniques developed within the academic practice are severely lacking both of these properties, ([235]). Hence, it is no surprising that RMPC has not found applicability in process industry. Then, we set the goal of this chapter as;

- a comprehensive review, discussion and classification of the current RMPC formulations in literature and signify the connections with risk theory and robust optimization research areas;

- a comparison of these formulations regarding their closed loop performance and computational load.

2.1.1 Notation for Robust MPC Problems

The field of reals and sets of nonnegative reals, integers and nonnegative integers are denoted by $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{Z}$, $\mathbb{Z}_{\geq 0}$, respectively. For a vector $x \in \mathbb{R}^{n\times 1}$ (or in short $x \in \mathbb{R}^n$), $x^\top$ denotes the transpose of that vector. A sequence of vectors $x_i \in \mathbb{R}^n$, until (time) index k, is defined as $x_{[0,k]}^\top := \begin{bmatrix} x_0^\top & x_1^\top & x_2^\top & \dots & x_k^\top \end{bmatrix} \in \mathbb{R}^{n(k+1)}$. For a (time) sequence of vectors $x_{[0,t]}$, $x_{i|k}$ denotes the (predicted) vector x_{k+i} for $i, k, t \in \mathbb{Z}_{\geq 0}$ and $k + i \leq t$. Furthermore, $x_{[a,b]|k}$ denotes the sequence of vectors $x_{[k+a,k+b]}$.

For normed vector spaces, $||\cdot||_p$ denotes the standard $p-$norm in $\mathbb{R}^n$. The unit ball in $\mathbb{R}^n$ corresponding to the $p-$norm is denoted with $\mathcal{B}_p^n := \{x \in \mathbb{R}^n | \, ||x||_p \leq 1\}$. If p is not explicitly specified then $||\cdot||$ is taken as the Euclidean norm $||\cdot||_2$, i.e., $\|\cdot\| = \|\cdot\|_2$.

The spectral radius of a matrix A, i.e., $\rho(A)$, is defined as $\rho(A) := \max_i |\lambda_i(A)|$, where λ_i is the i^{th} eigenvalue of matrix A. The identity matrix, with dimension $n \times n$, is denoted by I_n. The Kronecker product of matrices $A \in \mathbb{R}^{m_1\times n_1}$ and $B \in \mathbb{R}^{m_2\times n_2}$ is defined from the weighted block matrix concatenation

$$A \otimes B = \begin{bmatrix} A_{(1,1)}B & \dots & A_{(1,n_1)}B \\ \vdots & \ddots & \vdots \\ A_{(m_1,1)}B & \dots & A_{(m_1,n_1)}B \end{bmatrix},$$

or element-wise defined as

$$A \otimes B_{(m_2(k_1-1)+k_2,n_2(l_1-1)+l_2)} = A_{(k_1,l_1)}B_{(k_2,l_2)},$$

where $(i, j)^{th}$ element of matrix A is denoted with $A_{(i,j)}$.

Given two sets in a vector space $\mathcal{X},\ \mathcal{Y} \subseteq \mathbb{R}^n$, the Minkowski sum of these sets is defined by

$$\mathcal{X} \oplus \mathcal{Y} = \{x + y \in \mathbb{R}^n | \, x \in \mathcal{X},\ y \in \mathcal{Y}\},$$

while the Pontryagin difference of these two sets, assuming $\mathcal{X} \subset \mathcal{Y}$,

$$\mathcal{Y} \ominus \mathcal{X} = \{y \in \mathcal{Y} | \, y + x \in \mathcal{Y},\ x \in \mathcal{X}\}.$$

For stochastic variables, we assume that there is an underlying probability space $(\mathbb{R}^{n_\varsigma}, \mathcal{F}, \mathbb{P})$ equipped with the event space $\mathbb{R}^{n_\varsigma}$, the $\sigma-$algebra $\mathcal{F}$ defined over the Borel sets of $\mathbb{R}^{n_\varsigma}$ and well defined probability measure $\mathbb{P} : \mathfrak{P}^{\mathbb{R}^{n_\varsigma}} \to [0, 1]$, where $\mathfrak{P}^{\mathbb{R}^{n_\varsigma}}$ is the power set of $\mathbb{R}^{n_\varsigma}$, see [42, 267] for rigorous treatment of probability spaces. Here, for brevity of discussion, we assume that there is no measurability issues with the stochastic variables evolving over dynamics and used operators. We denote the probability of a random variable $\tilde{w} \in \mathbb{R}$ to take

values between $\underline{w}$ and $\bar{w}$ as $\mathbb{P}\{\underline{w} \leq \tilde{w} \leq \bar{w}\}$, which is equal to the integral of the probability density function (pdf), $f_{\tilde{w}}(w)$, over the interval $[\underline{w}, \bar{w}]$, i.e.,

$$\mathbb{P}\{\underline{w} \leq \tilde{w} \leq \bar{w}\} = \int_{\underline{w}}^{\bar{w}} f_{\tilde{w}}(w)dw,$$

while for the multi-dimensional case, $\tilde{w} \in \mathbb{R}^{n_\zeta}$, $n_\zeta \geq 1$, the probability is defined over multiple integrals over the considered region $\bar{\mathcal{W}}$. Lastly the mean, the variance, the moment of order n and the covariance matrix of random variables $\tilde{x}$ and $\tilde{y}$ are denoted with $\mu_{\tilde{x}}$, $\sigma^2_{\tilde{x}}$, $\mathbb{E}\{\tilde{x}^n\}$, $\Sigma_{\tilde{x},\tilde{y}}$, respectively.

2.1.2 Dynamical Systems and Optimization Based Control

In this work, we detail our discussion on robust MPC for discrete-time linear uncertain systems, denoted with Σ, together with its nominal counterpart Σ^{nom}.

$$\begin{aligned} \Sigma: \quad & \begin{cases} x_{k+1} = & A(\delta_k)x_k + B(\delta_k)u_k + Fw_k, \\ y_k = & C(\delta_k)x_k + D(\delta_k)u_k + v_k, \end{cases} \\ \Sigma^{nom}: \quad & \begin{cases} \bar{x}_{k+1} = & A_0\bar{x}_k + B_0u_k, \\ \bar{y}_k = & C_0\bar{x}_k + D_0u_k, \end{cases} \end{aligned} \tag{2.1}$$

where x_k (or $\bar{x}_k$) $\in \mathbb{R}^n$, and $u_k \in \mathbb{R}^{n_u}$, and y_k (or $\bar{y}_k$) $\in \mathbb{R}^{n_y}$ are the uncertain (or nominal) state, the control input and the uncertain (or nominal) output at discrete time instant $k \in \mathbb{Z}_{\geq 0}$, respectively. The plant Σ is subject to three types of uncertainties denoted by $\delta_k \in \Delta \subseteq \mathbb{R}^{n_\delta}$, $w_k \in \mathcal{W} \subseteq \mathbb{R}^{n_w}$ and $v_k \in \mathcal{V} \subseteq \mathbb{R}^{n_v}$ for $k \in \mathbb{Z}_{\geq 0}$. These uncertainties are either the model uncertainties or the disturbances effecting the state and output equation and can take values from bounded sets, i.e., $w_k \in \mathcal{W}$ for $k \in \mathbb{Z}_{\geq 0}$, or they can be stochastic vector sequences with known pdfs, e.g., $f_{\tilde{w}_k}(w_k) : \mathfrak{P}^{\mathbb{R}^{n_w}} \to \mathbb{R}$or they can be depending on the instantaneous values states x_k or inputs u_k implicitly, e.g., $\delta_k(x_k, u_k)$, see [292]. The matrices $A(\delta_k)$, $B(\delta_k)$, $C(\delta_k)$, for all δ_k realizations, are real matrices with dimensions $\mathbb{R}^{n\times n}$, $\mathbb{R}^{n\times n_u}$, $\mathbb{R}^{n_y\times n}$, respectively. We assume that the uncertain and the nominal systems Σ and Σ^{nom} are stabilizable and observable, see [160]. We use ζ_k for all uncertain variables, i.e., $\zeta_k^\top = \begin{bmatrix}\delta_k^\top & w_k^\top & v_k^\top\end{bmatrix}$. The i^{th}-step prediction of the state at the k^{th} time instant is denoted with $x_{i|k}$ which, by Equation (2.1), depends on the initial point x_k, the system dynamics Σ, the exogenous inputs $u_{[0,i-1]|k}$, $w_{[0,i-1]|k}$ and the internal uncertainties $\delta_{[0,i-1]|k}$.

Here we consider that the system is subject to hard or soft state (or output) and hard input constraints[3]. The constraints are represented here as inequalities,

$$c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k}, y_{j|k}) \leq 0, \quad \text{for } j \in \mathbb{Z}_{[0,N_p-1]}, \text{ for } i \in \mathbb{Z}_{[1,N_c^j]}, \tag{2.2}$$

where N_p denotes the prediction horizon of the MPC controller and N_c^j is the number of constraints for the time step j. In general these constraints are much more explicit, such as set (or zone) membership constraints, i.e., $x_{j|k} \in \mathbb{X}_{j|k} \subseteq \mathbb{R}^n$, $u_{j|k} \in \mathbb{U}_k \subseteq \mathbb{R}^{n_u}$.

[3]The reason for allowing state (or output) constraints to be soft is that these constraints are in general performance requirements while the input constraints are induced from actuator limitations.

The computational complexity of the MPC problem is highly dependent on the prediction horizon (N_p), the total number of constraints $N_c := \sum_j N_c^j$ and the convexity properties of the constraints $c_{ij}(\cdot)$, hence each of them effect the resulting closed-loop performance and computational properties of the MPC algorithm.

We cast two distinct MPC problems for systems Σ and Σ^{nom}, denoted with $\mathcal{P}(k)$ and $\bar{\mathcal{P}}(k)$, respectively, at time $k \in \mathbb{Z}_{\geq 0}$. The mismatch between the solutions (or the resulting trajectories) of $\mathcal{P}(k)$ and $\bar{\mathcal{P}}(k)$ is the price paid for robustness:

$$\mathcal{P}(k): \begin{cases} \min\limits_{u_{[0,N_p-1]|k}} & \mathcal{R}^{cost}(J(x_k, u_{[0,N_p-1]|k}, \zeta_{[0,N_p-1]|k})) \\ \text{s.t.} & x_{j+1|k} = A(\delta_{j|k})x_{j|k} + B(\delta_{j|k})u_{j|k} + Fw_{j|k}, \\ & y_{j|k} = C(\delta_{j|k})x_{j|k} + D(\delta_{j|k})u_{j|k} + v_{j|k}, \\ & \mathcal{R}^{const}(c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k}, y_{j|k})) \leq 0, \\ & j \in \mathbb{Z}_{[0,N_p-1]}, \;\; i \in \mathbb{Z}_{[1,N_c^j]}, \;\; x_{0|k} = x_k, \end{cases} \tag{2.3a}$$

$$\bar{\mathcal{P}}(k): \begin{cases} \min\limits_{u_{[0,N_p-1]|k}} & J(\bar{x}_k, u_{[0,N_p-1]|k}, 0) \\ \text{s.t.} & \bar{x}_{j+1|k} = A_0\bar{x}_{j|k} + B_0u_{j|k}, \\ & \bar{y}_{j|k} = C_0\bar{x}_{j|k} + D_0u_{j|k}, \\ & c_{ij}(\bar{x}_{j|k}, u_{j|k}, \bar{y}_{j|k}) \leq 0, \\ & j \in \mathbb{Z}_{[0,N_p-1]}, \;\; i \in \mathbb{Z}_{[1,N_c^j]}, \;\; \bar{x}_{0|k} = x_k, \end{cases} \tag{2.3b}$$

where $J(\cdot)$ is assumed to be the standard quadratic cost function for convenience,

$$\begin{aligned} J(x_k, u_{[0,N_p-1]|k}, \zeta_{[0,N_p-1]|k}) := & \sum_{j=0}^{N_p-1} J_r(x_{j|k}, u_{j|k}) + J_f(x_{k+N_p}), \\ = & \sum_{j=0}^{N_p-1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k} + x_{N_p|k}^\top Q_f x_{N_p|k}, \end{aligned} \tag{2.3c}$$

In Equation (2.3c), J_r is named as the *running cost* or the *cost to go* and J_f is the terminal cost, i.e., [240, 241]. In $\mathcal{P}(k)$ we introduce a (risk) functional $\mathcal{R}^{cost} : \mathcal{J}(x, u, \zeta) \to \bar{\mathcal{J}}(x, u)$ (or similarly $\mathcal{R}^{const}(\cdot) : \mathcal{C}(x, u, \zeta, y) \to \bar{\mathcal{C}}(x, u, y)$), which is mapping an uncertain function $\mathcal{J}$ to a deterministic function $\bar{\mathcal{J}}$ which is the robust counterpart of the cost function ($\mathcal{J}$) (equivalently for the constraints ($\mathcal{C}$)). For linear systems, due to possible parametrization, there are easy ways for expressing the robust counterparts of cost and constraint functions, i.e., $\mathcal{R}^{cost}(J)$ or $\mathcal{R}^{const}(c_{ij})$, which is not true for nonlinear system dynamics.

The RMPC problems are dealing with robust stabilization of the system Σ while satisfying the constraints $c_{ij}(\cdot)$ in a risk-aware manner, by solving $\mathcal{P}(k)$. It is expected from a robust predictive controller to adjust the average performance levels versus the constraint violation possibility. Consider the hypothetical scenario of operating a process close to a constrain as depicted in Figure 2.1. Then the distance from the constraint boundary, i.e., q_c, to operating point 1 (O.P.1) or O.P.2, i.e., d_1 and d_2 is a nominal performance metric. Furthermore the inherent standard deviation in operating conditions σ_{OP1} or σ_{OP2} are the relevant to the robustness of the operation, as the controller suppresses the dispersion from O.P.2, then the nominal operation can be pushed towards the constraint, such as operating at O.P.1 on average. Lastly, as also mentioned in [226], nonlinear predictive control techniques can be preferred to control the process, since some uncertainty realizations are not

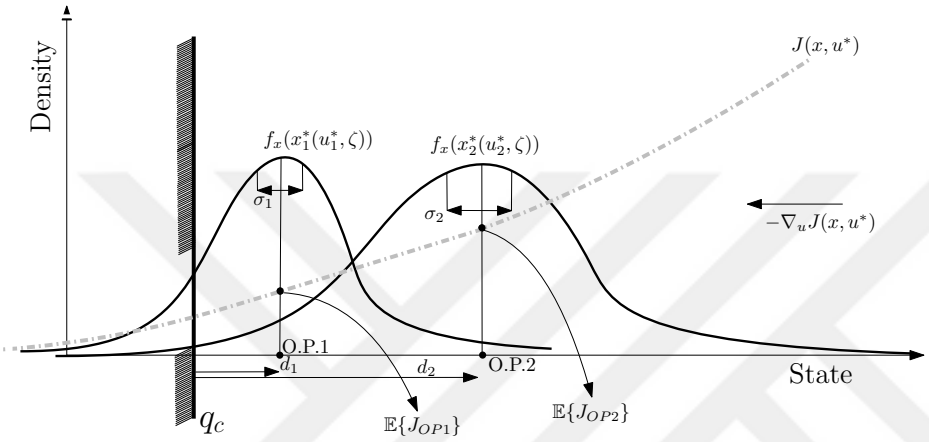


Figure 2.1: Selection of the steady state operating point.

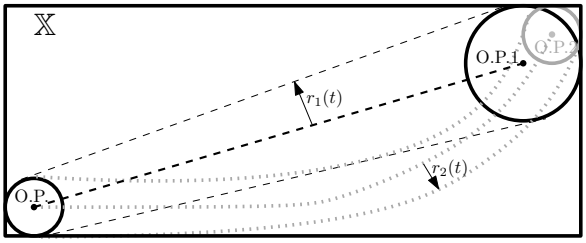


Figure 2.2: Steering in between two operating points.

necessarily adversary, instead many realizations are pushing the operation to the desired direction in input-state-output space. For the case of transitions between two operating points as depicted in Figure 2.2, the time window that the servo action to achieve its goals versus the robust constraint satisfaction is another robustness metric. As the controller becomes more aggressive, which is expected to decrease the transition time, the spread of possible trajectories in general grows considerably[4], i.e., for $r_1(t), t \in \mathbb{R}_{[0,t_f^1]}$ and $r_2(t), t \in \mathbb{R}_{[0,t_f^2]}$, in general one satisfies $t_f^1 \leq t_f^2$ only if $r_1(t') \geq r_2(t')$, for $t' \in \bar{t}, \bar{t} = \min(t_f^1, t_f^2)$.

Regarding the robust stability property, three meta-approaches are used in the literature, [242], to achieve guaranteed stability. These approaches are;
(i) designing an MPC controller for the nominal system by neglecting uncertainty and achieving robustness in an ad-hoc manner. This approach has close connections with the certainty equivalence reasoning where the uncertainty realizations are assumed to be equal to the mean value over the prediction stages;
(ii) considering some (or all) of the possible realizations in the uncertainty set to generate

[4]In Section 2.2.3 we implement various robust MPC techniques for a batch reactor and a CSTR system in simulation environment to compare the closed-loop performance metrics such as the ones mentioned here.

the possible trajectories and cast the system to be robust with respect to all of these realizations;
(iii) solving the closed-loop MPC problem through using a dynamic feedback over the predictions. This case eventually leads to a dynamic programming problem ([23]) which is the most promising case from the perspective of robustness however the computational demand to solve the associated problem increases exponentially, so much that, this approach is not valid for many of the real world examples ([246]).

In the case of the first approach, the uncertain variables are replaced with fixed instances of them thus we handle a (simpler) nominal MPC problem, however, the effect of other realizations are not taken into account, hence no provisions of risk are taken. This leads to possibly frequent and highly effective constraint violations or instability[5]. Furthermore one can make use of the larger MPC problems (with more constraints or scenarios) since nominal MPC problems are computationally simpler than the cases where uncertainties are present. To compensate for the effects of the uncertainty, one can incorporate some realizations of uncertainty into the predictions, case (ii). However, an undesired effect of this choice is that it leads to poor control actions, hence poor responses, since the optimal input depends heavily on the selected uncertainty realizations, i.e., $u_k^*(\bar{\zeta}^1_{[0,N_p-1]|k}) \neq u_k^*(\bar{\zeta}^2_{[0,N_p-1]|k})$, where $u_k^*(\bar{\zeta}^i_{(\cdot)})$ is the optimal solution of MPC problem $\mathcal{P}(k)$ with respect to the uncertain variables $\bar{\zeta}^i_{[0,N_p-1]|k}$.

2.1.3 Uncertainty Descriptions and Uncertainty Quantification

The internal model principle states that the uncertainty model should resemble the true uncertainty effecting the system to counteract adverse effects of it. However, the complexity to describe the uncertainty is a crucial factor on the feasibility and the practical applicability of the RMPC algorithms.

Here we group the uncertainties as model-based (internal) uncertainties and environment based (exogeneous) uncertainties. The mismatch between the process and the mathematical model (internal uncertainty) can be induced from, [100], unmodeled dynamics, time varying effect in the process or changing loads, while the external uncertainties, the disturbance signals and the output noise, are effecting the control input, the state evolution or the measurement signals.

Modelling the Internal Uncertainties

We classify the model mismatch possibilities as;

- **Multi-model Uncertainty:** Equally acceptable class of models, a countable set of models likely to represent the true system, see [263] as an example,

$$\left[\begin{array}{c|c} A(\delta) & B(\delta) \\ \hline C(\delta) & D(\delta) \end{array}\right] \in \left\{ \left[\begin{array}{c|c} A(\delta_i) & B(\delta_i) \\ \hline C(\delta_i) & D(\delta_i) \end{array}\right] | \delta_i \in \Delta, i = 1, 2, \ldots, N_{MM} \right\}. \tag{2.4}$$

- **Unknown-but-bounded uncertainty description:** By assigning a nominal model with describing the uncertainty set Δ in terms of; (i) affine relations on the system matrices,

[5] In practice these drawbacks are suppressed by improving the prediction model, which shrinks the uncertainty set.

such as, for the polytopic affine uncertainties case, which is quite common in control relevant literature

$$\left[\begin{array}{c|c} A(\delta) & B(\delta) \\ \hline C(\delta) & D(\delta) \end{array}\right] \in \Delta := \left\{ \left[\begin{array}{c|c} A_0 & B_0 \\ \hline C_0 & D_0 \end{array}\right] + \sum_{i=1}^{N_v} \delta_i \left[\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right], \sum_{i=1}^{N_v} \delta_i = 1, \delta_i \geq 0 \right\},$$

or for the ellipsoidal affine uncertainty case, which is not commonly preferred in MPC literature,

$$\left[\begin{array}{c|c} A(\delta) & B(\delta) \\ \hline C(\delta) & D(\delta) \end{array}\right] \in \Delta := \left\{ \left[\begin{array}{c|c} A_0 & B_0 \\ \hline C_0 & D_0 \end{array}\right] + \sum_{i=1}^{N_v} \delta_i \left[\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right], (\delta - \delta_c)^\top Q_\delta (\delta - \delta_c) \leq 1, \right\},$$

(ii) parameteric (structured or unstructured, dynamic or static) uncertainty description, ([133]), such as

$$\Sigma : \left\{ \begin{array}{rl} x_{k+1} = & A_0 x_k + B_0 u_k + B^\delta \delta_k, \\ y_k = & C_0 x_k, \\ q_k = & C^\delta x_k + D^\delta u_k, \\ \delta_k = & \Delta q_k, \end{array} \right.$$

where Δ is a(n) (un-)structured, dynamic (or static) operator with norm bounds on the (possibly time-varying) elements; (iii) uncertainty in impulse response coefficients ([30]) such as $\Delta^h = \{h_k | h_k^l \leq h_k \leq h_k^u\}$, where $y_k = \sum_{\tau=0}^{k} h_{k-\tau} u_\tau$, or uncertainty in frequency response values $\Delta(j\omega) = \{\delta(j\omega) \in \mathbb{C}(j\omega), \omega \in [0, \infty)\}$, where $Y(j\omega) = G(j\omega)(I + \Delta(j\omega))U(j\omega)$ and $G(j\omega)$ is the nominal transfer function. In some cases uncertain systems modelled via the linear fractional representation (LFR) are preferred in comparison to other uncertainty models, due to the ease of incorporation of a larger set of possible dynamics, see the discussion in [274, 321] or [100].

- **probability distribution functions for uncertainties:** by assigning probabilistic information to the parameters of uncertain models to describe the internal uncertainties. One possible benefit of utilizing stochastic descriptions is that the models/parameters identified from data yield statistical information which is difficult to express in deterministic uncertainty models such as the previous cases.

Modelling the External Uncertainties

External uncertainties are the unknown signals affecting the system in an exogenous way, w_k and v_k vectors in Equation (2.1). We note three distinct approaches for modelling the external uncertainty while considering only the effect of w_k, since a similar reasoning is valid also for noise vector v_k.

- **The uncertainties with discrete set of realizations:** The uncertainties with an event space of finite number of realizations can effect the plant exogenously, such as different load configurations,

$$\Sigma : \left\{ \begin{array}{rl} x_{k+1} = & Ax_k + Bu_k + Fw_k, \\ y_k = & Cx_k + Du_k, \end{array} \right. \tag{2.5}$$

where $w_k \in \mathcal{W} := \{w^1, w^2, \ldots, w^{N_s}\}$. Since the number of realizations are finite, one can come up with all of the possible instances, called also as the scenario tree.

- **Unknown-but-bounded disturbances:** Uncertainties with bounded support can also be affecting the processes, the case for Σ in (2.5) with $w_k \in \mathcal{W} \subset \mathbb{R}^{n_w}$. In literature, these uncertainty sets are mainly considered to be either polytopic or ellipsoidal sets, due to the numerical properties of these set classes, see [54] for an extensive discussion on the set-theoretic methods.
- **Statistical descriptions:** The distributional uncertainties can also be used instead of bounded disturbances. In this case the uncertain state evolution can be modeled as driven by an exogenous signal that is assuming realizations from its pdf. such as the estimated (initial) state that is deviating from the true state with a finite covariance value.

Remark 2.1.1 *The uncertainty models are in general not equivalent to each other, that is the uncertain dynamics describe different behaviours (or different processes) for different uncertainty descriptions. This results in MPC algorithms belonging to different complexity classes. The MPC literature is very scarce with regard to rigorous uncertainty modeling, in general the uncertainties are not modeled at all, instead uncertainty descriptions are taken as a given.*

2.1.4 Reformulating the MPC Problem as the Robust Counterpart Problem

One can distinguish three different sources of uncertain predictions in MPC problems; i) wrong initial condition estimation due to measurement noise or lack of sensors measuring all of the states; ii) perturbations to the dynamics; iii) plant-prediction model mismatch. In any source of uncertainty, the state predictions result in a collection of trajectories (with or without stochastic properties). To incorporate possible prediction errors to the MPC formulation, one needs to apply a risk mapping ($\mathcal{R}^{cost}$ or $\mathcal{R}^{const}$ in Equation (2.3a)) to the functions of uncertain state, i.e., the cost and constraint functions. There are five different common approaches to map an uncertain optimization problem, which are shown in Figure 2.3. The resulting MPC problem, and hence the control law, differs according to the used

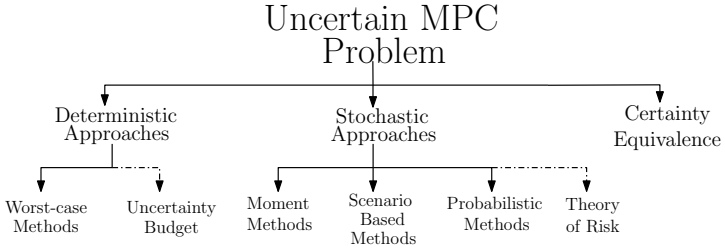


Figure 2.3: Commonly used projection techniques for uncertain optimization problems, in particular RMPC.

risk mapping. These risk mappings re-shape the feasible area of the optimization problem, in general drastically tightening it, see Figure 2.4 for the worst case, the probabilistic and the moment-based cases, but also the selected method effects the computational complexity properties of the robust counterpart problem. We summarize the effects of these risk

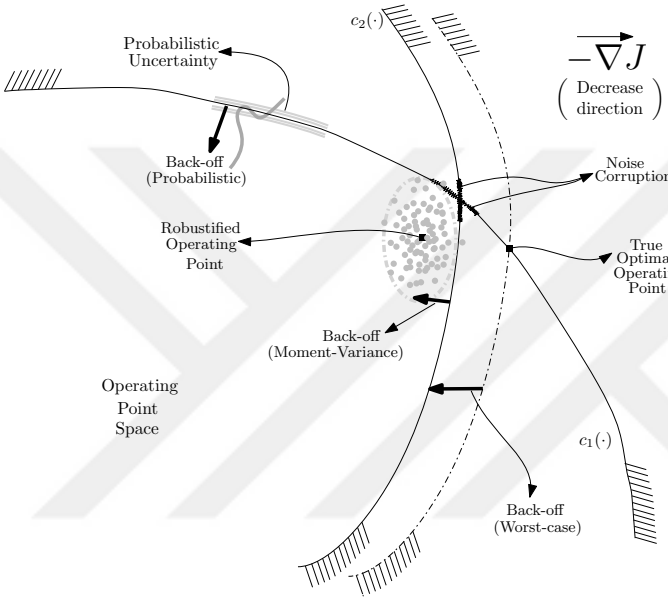


Figure 2.4: The operating point space and various forms of mapping the uncertainty to a deterministic counterpart.

mapping methods in Table 2.1, where the computational complexity of the nominal MPC problem is denoted with $\mathcal{O}(\bar{\mathcal{P}})$, the uncertain constraint (w.l.o.g. the cost) function with $c(x_{j|k}, u_{j|k}, y_{j|k}, \zeta_{j|k}) \leq 0$ or simply $c(\zeta) \leq 0$, the functions mapping the uncertain constraints to robust counterparts with $\mathcal{R}_i$, where i is the indicator of the respective mapping, the probability of satisfying the constraint with $\mathbb{P}\{c(\zeta) \leq 0\}$; 'coherent' corresponds to the risk function $\mathcal{R}_i$ being a coherent risk metric, implying a convex, monotonic and closed mapping, see [10] for the rigorous definition of coherence. All of the techniques result in a trade-off between various aspects, the constraint satisfaction guarantees, the resulting risk aversity, the computational complexity and the effort to model the uncertainty set. In the robust MPC literature, a rigorous discussion encompassing all aspects of the risk mappings in Table 2.1 is missing, even for the linear systems case.

Some of these techniques are already well studied, such as the worst-case or the scenario based approach, leading to many survey articles and books available in literature, see Table 2.2. Furthermore, we mention two topics that are closely related to the RMPC research topic, the contributions from the robust optimization and the (asymmetric) risk metrics research areas.

Approach		Uncertainty and Risk Function	Constraint Guarantees	Risk Aversity	Computational Complexity	Modelling Uncertainty
Scenario Based Approach	Nominal $\mathcal{P}^{CE}$	$\mathcal{R}_{CE} : c(\zeta) \to c(\bar{\zeta})$, $\bar{\zeta} = \mathbb{E}\{\zeta\}$ Coherent	No guarantees	Optimistic Risk-loving	$\mathcal{O}(\mathcal{P}^{CE}) = \mathcal{O}(\bar{\mathcal{P}})$	$Easy$ Calculate $\mathbb{E}\{\zeta\}$
	Randomized $\mathcal{P}^{rand}$	$\mathcal{R}_{rand} : c(\zeta) \to c(\bar{\zeta}^s)$, $s = 1, \ldots, N_s$ Coherent	Probabilistic guarantees $\mathbb{P}\{c(\zeta) \leq 0\} \geq \beta(N_s)$ with prob. $1 - \epsilon(N_s)$	Risk Decision Select N_s As $N_s \gg 1$ Pessimistic	$\mathcal{O}(\mathcal{P}^{rand}) \approx N_i \mathcal{O}(\bar{\mathcal{P}})$ Parallelizable	$Easy$ Generate scenario tree
Worst Case Based Approach	Full Set $\mathcal{P}^{WC}$	$\mathcal{R}_{WC} : c(\zeta) \to \max_{\zeta \in \Delta} c(\zeta)$, Coherent	$\mathbb{P}\{c(\zeta) \leq 0\} = 1$	Highly pessimistic	$\mathcal{O}(\mathcal{P}^{WC}) \gg \mathcal{O}(\bar{\mathcal{P}})$ Reformulations exist	$Difficult$ All possible realizations
	Uncertainty Budget Set $\mathcal{P}^{BWC}$	$\mathcal{R}_{BWC} : c(\zeta) \to \max_{\zeta \in \bar{\Delta}} c(\zeta)$, Coherent	$\mathbb{P}\{c(\zeta) \leq 0\} = \alpha_{BWC}(\bar{\Delta})$ $\lim_{\bar{\Delta} \to \Delta} \alpha_{BWC} \to 1$	Risk Decision Select $\bar{\Delta}$ As $\bar{\Delta} \to \Delta$ Pessimistic	$\mathcal{O}(\mathcal{P}^{BWC}) \gg \mathcal{O}(\bar{\mathcal{P}})$ Reformulations exist	$Easy$ Select realizations
Moment Based Approach	Mean $\mathcal{P}^{M}$	$\mathcal{R}_{M} : c(\zeta) \to \mathbb{E}\{c(\zeta)\}$, Coherent	$\mathbb{P}\{c(\zeta) \leq 0\} = \alpha_M$ α_M depends on $c(\cdot)$	Optimistic Risk-loving	(Online) $\mathcal{O}(\mathcal{P}^{M}) = \mathcal{O}(\bar{\mathcal{P}})$ (Offline) $\mathcal{O}(\mathbb{E}\{c(\cdot)\})$	$Easy$ Convergent with samples
	Mean Variance $\mathcal{P}^{MV}$	$\mathcal{R}_{MV} : c(\zeta) \to f_{MV}(c(\cdot), \zeta)$ $f_{MV}(\cdot) = \mathbb{E}\{c(\zeta)\} + \lambda_v \mathbb{D}\{c(\zeta)\}$, Not coherent Not monotonic	$\mathbb{P}\{c(\zeta) \leq 0\} = \alpha_{MV}$ $\lim_{\lambda_v \to \infty} \alpha_{MV} \to 1$	Risk Decision Select λ_v As $\lambda_v \gg 1$ Pessimistic	(Online) $\mathcal{O}(\mathcal{P}^{MV}) = \mathcal{O}(\bar{\mathcal{P}})$ (Offline) $\mathcal{O}(\mathbb{E}\{c(\cdot)\}, \mathbb{D}\{c(\cdot)\})$	$Easy$ Convergent with samples
Probabilistic Approach	$\mathcal{P}^{CC}$	$\mathcal{R}_{CC} : c(\zeta) \to q_{\alpha_{CC}}(c(\zeta))$, Not coherent Not convex	$\mathbb{P}\{c(\zeta) \leq 0\} = \alpha_{CC}$	Risk Decision Select α_{CC} As $\alpha_{CC} \to 1$ Pessimistic	(Online) $\mathcal{O}(\mathcal{P}^{CC}) \gg \mathcal{O}(\bar{\mathcal{P}})$ (Offline) Calculate $\mathbb{P}\{c(\zeta) \leq 0\}$	$Difficult$ Calculate true pdf of ζ, and propogate ζ

Table 2.1: Common approaches and properties of projecting an uncertain MPC problem to a robust counterpart MPC problem.

	Nominal MPC	Robust MPC	Stochastic MPC	Practical Aspects Industrial Applications	Robust Optimization	Theory of Risk
Surveys	[132, 235, 242, 250, 301] [30, 129, 206]	[207, 236, 242, 250, 301] [30, 129, 132, 133, 235] [206]	[235, 236, 246]	[290], [306], [288], [289]	[128], [43]	[128], [43], [314], [308], [319]
Books	[226], [123], [69], [145], [197], [302], [5, 187], [188], [135]	[226], [123], [69], [145], [197], [302], [5], [188]	[123], [197], [188]	[226], [123], [69], [5], [135]	[33]	[33]

Table 2.2: Surveys and books on MPC with robustness properties.

2.2 Literature Review on Robust Formulations of MPC Problems with Uncertain Elements

2.2.1 Deterministic Approaches to Uncertain Effects in MPC Problems

In this section, we, first, introduce the worst case (also called as the min-max) MPC approach, then we discuss the uncertainty budgets and their relation with MPC. Lastly, we present and classify contributions from MPC with deterministic robustness properties.

The worst-case (WC) optimization approach can be summarized with the following three distinct statements;

- The control action is calculated either in a “here-and-now" fashion, the uncertainty reveals itself after the control decisions are made, or “wait-and-see" fashion, the control action is decided after some of the uncertain variables reveal themselves to the decision maker.
- The decisions are made for, and only for, a known/decided subset of the uncertainty. If the whole (true) uncertainty set, say Δ, is taken, then we call it the worst case MPC (WC-MPC), else, if a strict subset is taken into account $\bar{\Delta} \subset \Delta$, then we name it the budgeted

worst case MPC (BWC-MPC).

- Any realization from the uncertainty set can not violate the constraints or destabilize the closed-loop system.

All of these statements are highly effective on the resulting control action and introduces the high pessimism in closed-loop response. This is due to the fact that the true trajectories of the system and predicted, but highly unlikely, set of trajectories deviate from each other, see Figure 2.5. Allowing the designer to decide on the set of uncertainty realizations, i.e.,

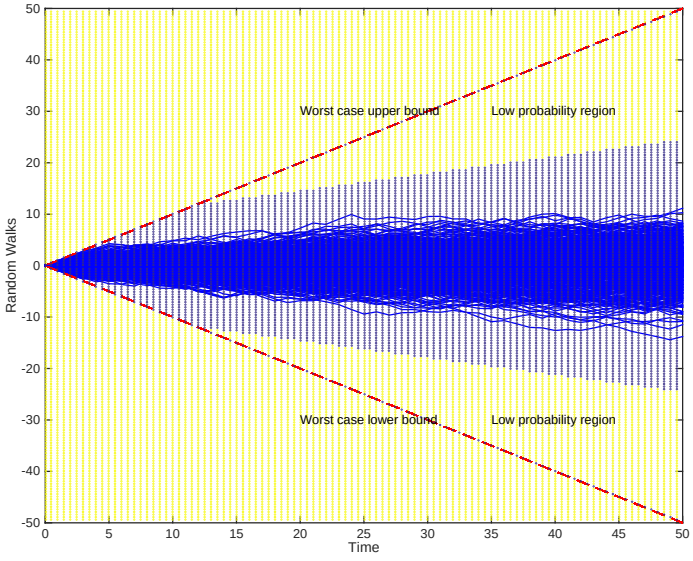


Figure 2.5: Fan of trajectories and the conservative worst-case bounds on the trajectories, for the discrete time integrator system effected by additive uncertainty with bounds $[-1, 1]$, inspired from [46].

$\bar{\Delta}$, might eventually improve the closed-loop performance. By actively selecting the set of uncertainties, one can quantitatively discuss the trade-off of incorporating a larger set of uncertainty versus adjusting the operating point.

In many cases WC-MPC problems are generally presented with here-and-now strategy and full uncertainty set case ([22]). The inherent pessimism within the construction is difficult to avoid, which leads to a tremendous effort in incorporating closed-loop predictions into the RMPC algorithms and computational problems. To reduce the pessimism, one can allow the control actions to be parameterized as function of future uncertainties or cast them as control policies, the wait-and-see formulation. Once the uncertainty set is decided and the control law structure (sequence or policies) is selected, then the distinction between different min-max MPC techniques depends on how the stability and constraint satisfaction

are guaranteed. Various methods are proposed in the literature which we group in Figure 2.6.

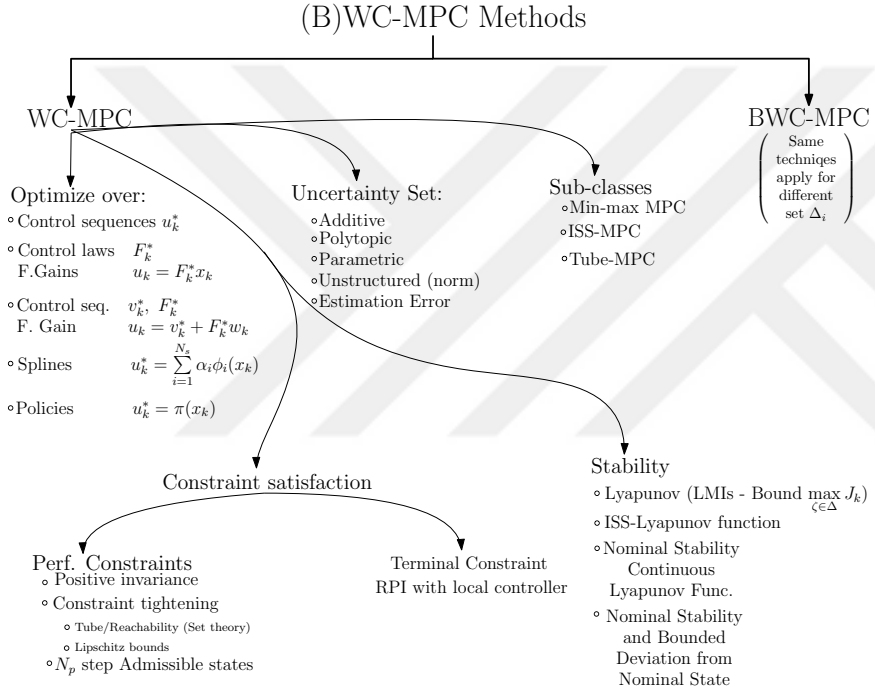


Figure 2.6: Diagram visualizing the commonly used techniques for WC-MPC methods, with the approaches utilized for achieving robust constraint satisfaction or stability.

One example of WC-MPC problems is given as in $\mathcal{P}^{WC}$,

$$\mathcal{P}^{WC}: \begin{cases} \min\limits_{u_{[0,N_p-1]|k}} \max\limits_{\zeta_{[0,N_p-1]|k}} & \sum\limits_{j=0}^{N_p-1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k} + x_{N_p|k}^\top Q_f x_{N_p|k}, \\ \text{s.t.} & x_{j+1|k} = A(\delta_{j|k}) x_{j|k} + B(\delta_{j|k}) u_{j|k} + F w_{j|k}, \\ & y_{j|k} = C(\delta_{j|k}) x_{j|k} + D(\delta_{j|k}) u_{j|k} + v_{j|k}, \\ & c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k}, y_{j|k}) \le 0,\ \forall \zeta_{[0,j]|k} \in \Delta_{[0,j]|k}, \\ & \zeta_{j|k}^\top = \begin{bmatrix} \delta_{j|k}^\top & w_{j|k}^\top & v_{j|k}^\top \end{bmatrix} \in \Delta_{j|k}, \\ & i = 1, \dots, N_c^j,\ \ j = 0, \dots, N_p - 1,\ \ x_{0|k} = x_k, \end{cases} \tag{2.6}$$

hence the risk mappings $\mathcal{R}^{cost}$ and $\mathcal{R}^{const}$ are selected as

$$\mathcal{R}_{WC}^{cost}(J) = \max_{\zeta \in \Delta} J(\zeta), \quad \mathcal{R}_{WC}^{const}(c_{ij}) = \max_{\zeta \in \Delta} c_{ij}(\zeta).$$

Here two important aspects of risk mappings $\mathcal{R}^{(\cdot)}$ are striking, (i) one minimizes the worst case cost and (ii) the constraints $c_{ij}(\cdot)$ should be satisfied for all realizations of $\zeta_{[0,j]|k}$. From these points and regarding the closed-loop performance, the min-max-MPC formulation is;

- leading to excessive backing-off from the nominally optimal operating point, since the constraints are forced to be satisfied for all uncertainty realizations within the prediction horizon. This results in a shrinking feasible set of control actions, see Table 2.1. Furthermore the min-max MPC algorithms are highly fragile to uncertainties with outlier realizations or unstable dynamics;
- deteriorating the average control performance. The worst-case cost is not a good representative of the true performance measure, in general, since

$$u_k^*(\arg\min_u \max_{\zeta\in\Delta} J(\zeta)) \neq u_k^*(\arg\min_u \mathcal{R}(J(\zeta))),$$

where $\mathcal{R}_{(\cdot)}$ is a risk mapping other than worst case method with full set of uncertainty, see Table 2.1.

From the computational point of view, the main difficulty is the inner maximization step (for the cost function) or the satisfaction of constraints for all realizations, leading to semi-infinite constraints for uncertainties with continuous domains. For convex problems, the constraint satisfaction is relatively well discussed in literature ([62]). The tube-based MPC is introduced into the RMPC (and later to SMPC) area by incorporating reachability analysis. For guaranteeing robust stability the stability conditions should be assured while guaranteeing the constraint satisfaction after the prediction horizon and bounded gains from the uncertain effects. A common technique to establish robust stability is to; 1) use a robust positively invariant (RPI) set, i.e., a set $\mathcal{X}_{RPI}$ satisfying the invariance property over time iterations;

$$x \in \mathcal{X}_{RPI} \rightarrow f(x,u,\zeta) \in \mathcal{X}_{RPI}, x \in \mathcal{X}_{RPI}, \exists u \in \mathcal{U}, \forall \zeta \in \Delta,$$

where $f(x,u,\zeta)$ is the dynamics of the considered system[6]; 2) construct a (locally continuous) Lyapunov function (inside the RPI set) that is sufficiently decreasing after each step. This approach is then further investigated within the input-to-state-stability (ISS) reasoning.

Remark 2.2.1 *The invariant sets play a crucial role in MPC problems and here we mention some important contributions on the area of invariant set calculations. In [196], authors present the construction of minimal invariant sets for linear systems with additive external uncertainties. A comprehensive treatment on the invariant sets of dynamical systems is also conducted in [182]. The paper [3] investigates the polytopic invariant set approximations from the ellipsoidal invariant sets for nonlinear systems effected by parametric uncertainties and exogenous inputs*[7]*. The wrapping effect, the phenomena observed when sets are over-approximated with simpler (to parametrize) sets, severely affects the invariant sets' size over*

[6]The set $\mathcal{X}_{RPI}$ is computationally difficult to find for large-scale nonlinear processes.

[7]In here we skip the detailed discussion on the calculation of invariant sets for dynamical systems affected by additive or multiplicative uncertainties. However the interested reader is referred to [181, 293] and also [54] for further information on the set-theoretic methods in control theory applications.

the approximation iterations. One effective treatment of this issue is discussed in [194], through the use of zonotopes.

For the min-max MPC methods there are well-established approaches and numerical methods that exist in the literature to overcome known drawbacks. In the next section, we discuss briefly two approaches that stem from original min-max MPC approach, the input-to-state MPC and tube based MPC techniques.

Input-to-state stability and RMPC

The notion of input-to-state stability ([337]) is frequently used to establish robust stability properties of uncertain (generally nonlinear) systems. One can make use of the ISS-Lyapunov functions to guarantee robust stability under continuity conditions on dynamics and/or the cost function. For the case of constrained problems, the constraints should also be tightened to account for the uncertain effects. Here we present the case by using explicit (but generally theoretical) Lipschitz bounds on the uncertain trajectories, summarized from [232]. Consider a nonlinear system with additive uncertainty, expressed with its nominal counterpart as

$$\Sigma : \left\{ \begin{array}{rl} x_{k+1} = & f(x_k, u_k, \zeta_k), \\ y_k = & h(x_k, u_k, \zeta_k), \end{array} \right. \quad \Sigma_{nom} : \left\{ \begin{array}{rl} \bar{x}_{k+1} = & \bar{f}(\bar{x}_k, u_k), \\ \bar{y}_k = & \bar{h}(\bar{x}_k, u_k), \end{array} \right. \tag{2.7}$$

with $\zeta_k \in \Delta$. Now assume that there exists a Lipschitz gain L_f for the nominal system, i.e.,

$$||\bar{f}(\bar{x}_1, \kappa(\bar{x}_1)) - \bar{f}(\bar{x}_2, \kappa(\bar{x}_2))||_p \leq L_f ||\bar{x}_1 - \bar{x}_2||_p$$

holds for all $\bar{x}_1, \bar{x}_2$ inside the region of interest $\mathcal{X}$ and $\kappa(\bar{x})$ is an admissible control law. If the uncertainties are bounded, i.e., $\zeta_k \in \Delta \subseteq \gamma \mathcal{B}_p(0)$, for all $k \in \mathbb{Z}_{\geq 0}$, then the predictions of the true state $x_{j|k}$ can be encapsulated inside a cone defined through the Lipschitz gain L_f, γ and $\bar{x}_{0|k} = x_k$ which is an overapproximation of the reachable set. We denote the upper bounds as $\gamma^j_{L_f,\gamma}\mathcal{B}_p(0)$, where $\gamma^j_{L_f,\gamma}$ is the calculated upper bound. Then, tighten the constraints with this bound as $\bar{\mathcal{X}}_{j|k} := \mathcal{X} \ominus \gamma^j_{L_f,\gamma}\mathcal{B}_p(0)$. Under some technical properties on the dynamics and cost function, the optimal solution u^*_k of $\mathcal{P}^{ISS}$ (Equation (2.8)) steers the state into the $\mathcal{X}_{Ter}$, guarantees recursive feasibility, and robustly stabilizes the closed-loop system,

$$\mathcal{P}^{ISS} : \left\{ \begin{array}{ll} \min\limits_{u_{[0,N_p-1]|k}} & \sum\limits_{j=0}^{N_p-1} \bar{x}^\top_{j|k} Q \bar{x}_{j|k} + u^\top_{j|k} R u_{j|k} + \bar{x}^\top_{N_p|k} Q_f \bar{x}_{N_p|k}, \\ \text{s.t.} & \bar{x}_{j+1|k} = \bar{f}(\bar{x}_{j|k}, u_{j|k}), \\ & \bar{x}_{j|k} \in \bar{\mathcal{X}}_{j|k}, \;\; u_{j|k} = \kappa(\bar{x}_{j|k}) \in \mathbb{U}, \;\; \bar{x}_{N_p|k} \in \mathcal{X}_{Ter} \\ & j = 0, \ldots, N_p - 1, \;\; \bar{x}_{0|k} = x_k, \end{array} \right. \tag{2.8}$$

Performance wise, by construction, the Lipschitz constant cased bounds, i.e., $\gamma^j_{L_f}$ for $j \in [1, N_p]$, can be overly conservative, causing over-tightening of the constraints, which potentially leads to infeasibilities or poor average performances. Computationally it is not clear whether the computational load is due to the requirements of ISS-MPC (tightening or Lipschitz bound calculation steps) or naturally induced from the nonlinearities of the plant.

Tube Based RMPC

The tube-based MPC is another WC-MPC approach based on the reachability analysis. The interesting observation of tube-based MPC is the separation of the robustness problem from the MPC problem. Consider the nonlinear system Σ and the nominal counterpart as in Equation (2.7). The idea is to keep the true state x_k close to the nominal state $\bar{x}_k$ over time instants $k \in \mathbb{Z}_{\geq 0}$, while controlling the nominal state $\bar{x}_k$ into a RPI set $\mathcal{X}_{Ter}$ at time $k+N_p$. For this purpose (although not necessary) with a "low-level" controller K, the control signal is constructed as $u_k = \bar{u}_k + K(x_k - \bar{x}_k)$, where the $K(x_k - \bar{x}_k)$ term attenuates the effect of the uncertainties. The tube is then all the possible values of the mismatch $x_{j|k} - \bar{x}_{j|k}$. Hence by subtracting the tube from the original constraints, the tightened constraints are found. Controlling the nominal state (the center of the tube) within the tightened constraint sets guarantees that the true state trajectory is inside the (true) constraint sets, while with several technical but common conditions on nominal MPC, the robust stability follows. A tube-MPC problem can be stated as follows;

$$\mathcal{P}^{tube} : \begin{cases} \min\limits_{\bar{u}_{[0,N_p-1]|k}} & \sum\limits_{j=0}^{N_p-1} \bar{x}_{j|k}^\top Q \bar{x}_{j|k} + \bar{u}_{j|k}^\top R \bar{u}_{j|k} + \bar{x}_{N_p|k}^\top Q_f \bar{x}_{N_p|k}, \\ \text{s.t.} & \bar{x}_{j+1|k} = A\bar{x}_{j|k} + B\bar{u}_{j|k}, \\ & \bar{x}_{j|k} \in \bar{\mathbb{X}}_{j|k},\ \bar{u}_{j|k} \in \bar{\mathbb{U}}_{j|k},\ \bar{x}_{N_p|k} \in \bar{\mathbb{X}}_{Ter|k}, \\ & j = 1, \ldots, N_p - 1,\ \bar{x}_{0|k} = x_k \end{cases} \tag{2.9}$$

with the standard stability conditions on the terminal set $\bar{\mathbb{X}}_{Ter|k}$, the terminal cost Q_f and the terminal controller K_f, which is not necessarily equal to the low-level controller K. The tightened constraints $\bar{\mathbb{X}}_{j|k}$, $\bar{\mathbb{X}}_{Ter|k}$, $\bar{\mathbb{U}}_{j|k}$ are calculated by following a three-step procedure;

i) Computation of reachable states from the initial state x_k with the plant dynamics for the prediction horizon as

$$\mathcal{X}_{[1,N_p]|k} := \begin{bmatrix} \mathcal{X}_{1|k} & \ldots & \mathcal{X}_{N_p|k} \end{bmatrix}, \tag{2.10a}$$

$$\mathcal{X}_{i|k} := \{x' | x' = f(x'', u, \zeta), u \in \mathbb{U}, \zeta \in \Delta, x'' \in \mathcal{X}_{i-1|k}\}, \tag{2.10b}$$

with $\mathcal{X}_{0|k} := \{x_k\}$;

ii) Calculation of the tightened constraint sets by

$$\begin{aligned} \bar{\mathbb{X}}_{[1,N_p-1]|k} &:= \mathbb{X} \ominus \mathcal{X}_{[1,N_p-1]|k}, \quad \mathcal{U}_{[1,N_p-1]|k} := K\mathcal{X}_{[1,N_p-1]|k}, \\ \bar{\mathbb{X}}_{Ter|k} &:= \mathbb{X}_{Ter} \ominus \mathcal{X}_{N_p|k}, \quad \bar{\mathbb{U}}_{[1,N_p-1]|k} := \mathbb{U} \ominus \mathcal{U}_{[1,N_p-1]|k}, \end{aligned} \tag{2.10c}$$

where $\mathbb{X}$ and $\mathbb{U}$ are the original constraint sets.

iii) Solution of the MPC problem $\mathcal{P}^{tube}$, from which one applies the first control signal to the nominal system while applying the control signal $u_k = \bar{u}_k + K(x_k - \bar{x}_k)$ to the uncertain plant.

Performance-wise, if no precompensator, K, is used, since the reachable state sets are equivalent, the tube-MPC is equivalent with the min-max MPC. However a well-defined feedback structure with K causes the deviations to be suppressed between the nominal and the true state within the predictions and hence results in a tube diameter which is much smaller in comparison to the $K = 0$ case. This is expected to increase the closed-loop

performance by allowing the MPC controller act on top of the controller K. On the other hand one important drawback of tube based MPC with controller structure is the shrinking input constraint set due to the compensation action $K(x_k - \bar{x}_k)$. If the compensator is aggressive to suppress the deviation $(x_k - \bar{x}_k)$, then the input constraints are tightened substantially, which can be so much that the feasible input set becomes an empty set.

Computationally, if one considers only the MPC problem, then it is simplified since we solve an nominal MPC problem. However tube-MPC approach contains (i) complex set-based operations, Minkowski sums (for reachable set calculations) and Pontryagin differences (for constraint tightening step), and (ii) calculation of a prestabilizing controller K for the process. Two major issues are inherent to set-based operations, (a) the memory requirement for expressing the n-step reachable sets grow quite fast (and approximations induce further conservativeness), (b) the effect of dimensionality, for medium to large sized systems (say more than 100 states), for the mentioned set operations are leading to numerical problems, even in the simplest forms of sets, see [127] for a discussion.

Remark 2.2.2 ***Affine State or Disturbance Feedback Structures:*** *An important development in the RMPC area is using low level controllers K within the optimization step. In parallel to this observation, one might seek to optimize the state feedback controller K at each time step on top of optimizing over control actions $\bar{u}_{[0,N_p-1]|k}$, i.e.,*

$$u_{j|k} = \bar{u}_{j|k} + \sum_{i=0}^{j} K_{j,i|k} x_{i|k}.$$

However the resulting MPC problem in $\bar{u}_{j|k}$ and $K_{j,i}$ is shown to be non-convex, hence leads to suboptimality or computational issues. The remarkable extension comes via adjustable robust optimization (ARO) approach, leading to convex RMPC problems with affine disturbance feedback structures, ([139]),

$$u_{j|k} = \bar{u}_{j|k} + \sum_{i=0}^{j} K_{j,i|k} w_i.$$

Price of Robustness, Uncertainty Budgets and RMPC Problems

In recent years, the robust optimization (RO) techniques, e.g., ([33], [128]), highlight one crucial question relevant also in the robust control area, “what is the corresponding performance deterioration with respect to the uncertainty set?". Here we name this aspect as the uncertainty budgets for robustness. Now consider the approach as in the worst-case MPC reasoning, but replace the uncertainty set Δ with a subset of it, say $\bar{\Delta}$. For an uncertain optimization problem $\mathcal{P}^{WC}(\Delta)$ that is feasible under Δ, then for any subset of the true uncertainty set, i.e., $\bar{\Delta} \subseteq \Delta$, the problem $\mathcal{P}^{BWC}(\bar{\Delta})$ remains feasible and the performance w.r.t. cost function is non-decreasing. Furthermore, since the selected set $\bar{\Delta}$ is directly effecting the resulting control action, i.e.,

$$u^* = \arg\min_u \max_{\zeta \in \bar{\Delta}} (J(x_k, u, \zeta)) = \arg\min_u \mathcal{R}^{cost}_{BWC}(J(x_k, u, \zeta)),$$

one can actively design an uncertainty set to which a RMPC is synthesized with desired performance properties, while providing robustness against $\bar{\Delta}$, see Table 2.1. With this construction one can circumvent the pessimism problems or give quantitative measures on the robustness to the uncertainty with unbounded distributions, [33], hence increasing the degree of freedom of the designer.

A generic RMPC example with uncertainty budgets can be stated as follows,

$$\mathcal{P}^{BWC}: \begin{cases} \min\limits_{\bar{u}_{[0,N_p-1]|k}} \max\limits_{\zeta_{[0,N_p-1]|k}} & \sum\limits_{j=0}^{N_p-1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k} + x_{N_p|k}^\top Q_f x_{N_p|k}, \\ \text{st.} & x_{j+1|k} = A(\delta_{j|k})x_{j|k} + B(\delta_{j|k})u_{j|k} + F w_{j|k}, \\ & y_{j|k} = C(\delta_{j|k})x_{j|k} + D(\delta_{j|k})u_{j|k} + v_{j|k}, \\ & c_{ij}(x_j, u_j, \zeta_j, y_{j|k}) \leq 0,\ \forall \zeta_{j|k} \in \bar{\Delta}_{j|k}, \\ & i = 1, \ldots, N_c^j,\ \ j = 0, \ldots, N_p - 1,\ \ x_{0|k} = x_k, \end{cases} \tag{2.11}$$

where $\bar{\Delta}$ is adjusting the trade-off between guaranteed robustness versus closed-loop performance (risk-aversion), [44]. The associated risk mappings are defined as

$$\mathcal{R}_{BWC}^{cost}(J) = \max_{\zeta \in \bar{\Delta}} J(\zeta), \quad \mathcal{R}_{BWC}^{const}(c_{ij}) = \max_{\zeta \in \bar{\Delta}} c_{ij}(\zeta).$$

The computational complexity of $\mathcal{P}^{BWC}$ problem is still challenging, similar to the $\mathcal{P}^{WC}$ problem, the min-max (saddle) nature of the problem is still effective. For generic convex uncertain optimization problems, the robust counterparts are in general intractable, while for some uncertainty classes the robustified problems remain convex, hence tractable ([218], [43]). For examples from literature, we refer to [34] for the robust convex optimization problems, [116] for the robust least squares problems, [117] for the robust semidefinite programming problems, [35] for the robust linear programming problems, [36] for the robustified quadratic programming or the robust conic programming problems, and lastly [45] for the approximations of the robust conic programming problems.

Contributions from RMPC Literature

In this section we present and classify various contributions in RMPC literature. One group of the earliest works done in close association with WC-MPC problems is the 'pursuit-evasion problems', i.e., [41, 131, 282, 368], closely related to the developments in the optimal control theory. However, the first works conducted strictly in receding horizon fashion was [22, 248, 377] and [209]; where both [248] and [22] elegantly present the WC-MPC problem for nonlinear and DAE systems, respectively, and [209] focuses on the frequency domain robust performance measures. The article [327] constructs the robust stability proof quite similar to the standard stability arguments for the nominal MPC problems. Another milestone contribution for guaranteeing robust stability property is [185] which uses linear matrix inequalities (LMIs) in RMPC problems for polytopic or parametric internal uncertainties through overbounding the cost function and using positive invariance arguments for the constraints on the future trajectory. Later, specifically for nonlinear systems, input-to-state stability (ISS) arguments are introduced for WC-MPC in [232]. A two-level (sometimes, though incorrectly, called as the closed-loop MPC) controller approach is presented in [86] which consists of a MPC and an $\mathcal{H}_\infty$ controller. Constructing a low-level controller

(precompansator) on top of a MPC lead to WC-MPC with disturbance feedback control structure [138] and tube based MPC approach, [199]. Lastly we mention the min-max MPC constructions with cost functions cast as linear norms such as [6, 72, 73] and more recently [29] while [313] provides a nice discussion on linear programming in MPC domain.

In Table 2.3 we distinguish publications in RMPC area with respect to; underlying system class (linearity), type of uncertainties, the high-level approach (min-max, ISS, tube), control parameterization, how the constraints are robustified and lastly the complexity of simulation examples provided in these publications. To comment on some observations on

System Dynamics and Control Implementation	Linear		Nonlinear		
	State Feedback	Output Feedback	State Feedback	Output Feedback	
	The majority of the papers	[240], [241], [209], [220], [221], [222], [367] [363], [311], [361], [264],	[237], [22], [232], [248], [86], [233], [203], [277], [292], [200] [201], [291], [1], [178], [74], [239], [219], [370], [247], [238]	[370], [340], [81],	
Uncertainty types	Additive External	Polytopic or Parametric Internal	Dynamic Internal	State Estimation Measurement Noise	Multi-model and Linearization
	[199], [237], [240], [180], [241], [22], [232], [139], [248], [327], [138], [29], [361], [233], [203], [277], [292], [200], [201], [291], [1], [239], [219], [370], [238], [340], [81], [137], [28], [217], [179], [351], [173], [315], [243], [298], [293], [295], [297], [296], [190], [355], [136],	[22], [248], [185], [29], [367] [363], [311], [233], [203], [292], [200], [201], [291], [1], [219], [370], [340], [28], [111], [364], [365], [279], [278], [208], [186], [107], [265],	[86], [220], [221], [222], [80], [176],	[240], [241], [209], [247], [340],	[263], [264], [74],
High-level Approach	Min-max	Cost overbounds	ISS	Tube	Other
	[180], [327], [248], [22], [29], [367], [292], [200], [201], [178], [28], [217], [173], [111], [208], [107], [265],	[263], [185], [220], [221], [222], [363], [311], [264], [291], [1], [364], [365], [279], [80], [176], [98]	[232], [139], [203], [277], [292], [370], [247], [340], [137], [315], [355], [136], [186],	[199], [237], [240], [241], [361] [74], [238], [243], [298], [293], [295], [297], [296], [190], [341],	[209], [86], [219], [351]
Control Parameterization	Control Sequence	Stabilizing Feedback Gains	Affine Disturbance Feedback Parameterization	Control Policy	
	[199], [237], [22], [232], [248], [327], [86], [29], [361], [292], [74], [219], [370], [247], [238], [340], [28], [351], [173], [315], [243], [298], [293], [186], [341],	[263], [185], [220], [221], [222], [367], [363], [311], [264], [364], [365], [279], [80], [175], [98],	[180], [139], [178], [137], [295], [297], [296], [190], [136],	[138], [200], [201], [291], [111], [341]	
Robust Constraints	Constraint Tightening via Tube	Constraint Tightening via Lipschitz Bound	Positive Invariance and LMIs	Admissible Sets Explicit Solution	
	[199], [237], [240], [241], [361], [74], [238], [243], [298], [293], [295], [297], [296], [190], [341],	[22], [232], [248], [86], [1] [178], [355], [186],	[263], [327], [185], [221], [222], [367], [363], [219], [351], [364], [279], [176], [98],	[180], [138], [29], [233], [203], [277], [200], [201], [291], [370], [247], [340], [137], [28], [173], [111], [107], [265], [315], [136],	
Complexity of simulation example	No example	1-2 State	3-4 State	5 or more states	
	[22], [138], [233], [203], [200] [291], [178], [219], [137], [136], [279],	[199], [240], [180], [241], [263], [232], [209] [327], [185], [86], [29], [367], [363], [361], [264], [277], [292], [201], [1], [74], [370], [247], [238], [28], [173], [315], [243], [293], [295], [297], [296], [190], [355], [111], [279], [186], [107], [265], [341], [294],	[237], [209], [185], [29], [220], [221], [222], [367], [363], [311], [81], [217], [364], [80], [176], [98],	[264], [74], [340], [81], [351],	

Table 2.3: An overview of RMPC approaches, via a group of selected papers.

the Table 2.3, due to the various contributions from output feedback structure and nonlinear systems domains, the current effort is directed towards practical implementations with various complexity reduction formulations. Similarly in many contributions the attention has been directed towards disturbance rejection properties (additive external uncertainties), while the polytopic internal uncertainty case has been treated (mainly) with LMI based techniques. The dynamic uncertainty and the effect of estimator in the loop (the initial condition mismatch case) is still lacking a detailed treatment. The ISS and the tube based RMPC approaches offer various solution strategies by decreasing the complexity back to the nominal MPC problem (with additional computationally complex operations) and the affine disturbance feedback parametrization lead to a huge reduction in the pessimism, while the sub-optimality should be further discussed in comparison to the policy based approaches. Lastly constraint tightening methods operate in coherence with the ISS and tube based ap-

proaches, while calculation of these sets (or approximations) is still to be mastered, which necessitates using highly complex (or realistic) examples to observe the drawbacks of the proposed techniques and convince the practitioners.

Although not in the RMPC area, but min-max based methods are also used in MPC design for i) LPV systems, [223, 268], ii) game-theoretic approaches, [94, 95], iii) decentralized or distributed control, [170], iv) (robust) estimation, [169], v) energy-consumption and smart-communication, [37]. Similarly due to the highly structured construction of MPC problems, multi-parametric programming approaches ([2]) were utilized for RMPC problems also. The article [29] constructs the robustly invariant set of an uncertain linear system with polytopic uncertainties for explicit RMPC purposes which is then extended by [173, 186, 315] and [180]. In [340], one of the first applications of explicit RMPC for a linearized system, solving the estimation and robust control problem for a batch polymerization process, is presented.

2.2.2 Stochastic Approaches to Uncertain Effects in MPC Problems

The second direction in establishing robustness properties of the closed-loop systems is by considering the uncertainties as stochastic variables, see, e.g., [24] or [40] as two main contributions.

In the early days of predictive control, within the GPC approach ([90, 91]) the stochastic optimization problem is transformed into a deterministic optimization problem by the expectation operator on the cost function, $\mathbb{E}\{J(x_k, u, \zeta)\}$. Similarly, we distinguish the approaches to SMPC problems with respect to the treatment of stochastic variables, how the uncertain functions are mapped to deterministic counterparts. There are three inherently different methods to reformulate stochastic cost and constraint functions as deterministic functions;

1. Considering the moments of the functions through the expectation operators.
2. Considering probabilistic (chance) constraints on the functions, utilizing known pdfs,
3. Considering finite number of uncertainty scenarios and casting the RMPC problem w.r.t. these scenarios.

Throughout this section, we assume that each realization of the uncertain variables is independent from another, with known pdf for each variable w_k, v_k, δ_k for $k \in \mathbb{Z}_{\geq 0}$.

Moment-based MPC Problems

The moment-based optimization is possibly the most frequently used methodology in optimization problems with stochastic elements due to its relatively easy modeling and low complexity implementation, see Table 2.1. The (centralized) moments provide inherent statistical information, such as the mean, variance, skewness or kurtosis which indicate the average magnitude, the spread, the asymmetry or the fat-tailedness of the predicted trajectories, all of which are desired to be controlled.

In the moment-based approach, one maps the uncertain optimization problem by considering (linear combinations of) the (centralized) moments of the cost and the constraint functions. The expectation of uncertain functions can be easily and analytically expressed for several classes of random variables, which reduces the computational need, while providing

a realistic performance measure ([250]). In its simplest form of mean MPC (M-MPC), one evaluates the mean of the cost and constraints as,

$$\mathcal{P}^M : \left\{ \begin{array}{ll} \min\limits_{u_{[0,N_p-1]|k}} & \mathbb{E}\left\{ \sum\limits_{j=0}^{N_p-1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k} + x_{N_p|k}^\top Q_f x_{N_{p|k}} \right\}, \\ \text{st.} & x_{j+1|k} = A(\delta_{j|k})x_{j|k} + B(\delta_{j|k})u_{j|k} + Fw_{j|k}, \\ & y_{j|k} = C(\delta_{j|k})x_{j|k} + D(\delta_{j|k})v_{j|k}, \\ & \mathbb{E}\{c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k}, y_{j|k})\} \leq 0, \\ & i = 1, \ldots, N_c^j, \quad j = 0, \ldots, N_p, \;\; x_{0|k} = x_k, \end{array} \right. \tag{2.12}$$

hence

$$\mathcal{R}_M^{cost}(J(\zeta) := \mathbb{E}_\zeta\{J(\zeta)\}, \quad \mathcal{R}_M^{const}(c_{ij}(\zeta)) := \mathbb{E}_\zeta\{c_{ij}(\zeta)\}.$$

Performance-wise one needs to distinguish an aspect of M-MPC, for robust constraint satisfaction, the effect of uncertainty in constraints of $\mathcal{P}^M$ is disregarded by the expectation operator. This does not imply any quantitative robustness guarantees for some realizations of uncertainty, see Table 2.1. To improve the robust constraint satisfaction properties, one can introduce variance terms, which indicates the spread of the uncertainty realizations and effectively backs off the operating point such that the constraints are satisfied with a higher probability. This MPC formulation is called here as the mean-variance MPC (MV-MPC) and an example is given in Equation (2.13),

$$\mathcal{P}^{MV} : \left\{ \begin{array}{ll} \min\limits_{u_{[0,N_p-1]|k}} & \mathbb{E}\left\{J(x_{j|k}, u_{j|k}, \zeta_{j|k})\right\} + \lambda_0 \mathbb{D}\left\{J(x_{j|k}, u_{j|k}, \zeta_{j|k})\right\} \\ \text{st.} & x_{j+1|k} = A(\delta_{j|k})x_{j|k} + B(\delta_{j|k})u_{j|k} + Fw_{j|k}, \\ & y_{j|k} = C(\delta_{j|k})x_{j|k} + D(\delta_{j|k})u_{j|k} + v_{j|k}, \\ & \mathbb{E}\{c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k}) + \lambda_{i,j}\mathbb{D}\{c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k})\} \leq 0, \\ & i = 1, \ldots, N_{c_i^j}, \quad j = 0, \ldots, N_p, \quad x_{0|k} = x_k, \end{array} \right. \tag{2.13}$$

where $\mathbb{D}\{\cdot\}$ is the variance operator and $\lambda_{i,j}$ (also λ_0) are positive weights, related to the risk aversion of the designer. This yields the risk mappings

$$\begin{array}{rl} \mathcal{R}_{MV}^{cost}(J(\zeta) := & \mathbb{E}_\zeta\{J(\zeta)\} + \lambda_0 \mathbb{D}_\zeta\{J(\zeta)\}, \\ \mathcal{R}_{MV}^{const}(c_{ij}(\zeta)) := & \mathbb{E}_\zeta\{c_{ij}(\zeta)\} + \lambda_{i,j}\mathbb{D}_\zeta\{c_{ij}(\zeta)\}. \end{array}$$

Regarding the performance aspects, ([50]), through including the $\mathbb{D}\{\cdot\}$ term into the cost function, one introduces the effect of uncertainty in the control actions, as λ_0 gets larger, the controller acts to reduce the spread of the cost function. For robust constraint satisfaction, which depends on the weights $\lambda_{i,j}$, the resulting operating point is backed off, as the feasible set is tightened by increasing $\lambda_{i,j}$ or $\mathbb{D}\{\cdot\}$. If $\lambda_{i,j}$ are large then the constraints are tightened substantially, hence allowing many realizations of ζ to occur without violating the constraints. However guaranteeing constraint satisfaction with certainty is not possible upto a large value of $\lambda_{i,j}$ for bounded domain ζ, see Table 2.1.

Yet another aspect is the asymmetry in the distribution functions of the constraints or cost. The deviation-measure term on the cost or constraint functions (here taken as the variance) is not necessarily distinguishing the 'good' and 'bad' realizations of the uncertainty, see Figure 2.1. Some realizations are steering the operation out of the feasible region, called

as the 'bad' realizations, while some realizations push the operating point further inside the constrained region, called as the 'good' realizations (the good-bad aspect reverses, in general, from the performance point of view). If the risk-neutral approach i.e. the symmetric evaluation of uncertainty realizations, such as the variance, is taken, then the control performance might reduce for a longer (skewness) or fatter (kurtosis) 'good' realization tail, see also [10, 88].

Computationally, moment-based MPC problems are divided into two distinct phases. During the offline phase the moments of the cost and constraint functions are calculated (or approximated through histograms) and during the online stage an uncertainty-free (computationally equivalent to the nominal) MPC problem is solved. Expressions of the moments $\mathbb{E}\{J(\cdot)\}$ (or $\mathbb{E}\{c_{ij}(\cdot)\}$) or $\mathbb{D}\{J(\cdot)\}$ (or $\mathbb{D}\{c_{ij}(\cdot)\}$) are easily obtained for linear systems, while for nonlinear systems computationally cheap methods exists, see [339, 347] or Table 2.1.

Probabilistic Approaches to MPC Problems

The second approach to transform SMPC problems is using the probabilistic constraints, also called as the chance constraints, leading to chance-constrained MPC (CC-MPC) formulation. In this setting, one finds the control actions to reduce the frequency of constraint violations according to predefined probability levels α_{ij}, see [82] and also Table 2.1. This provide quantitative ways of selecting the reliability of constraint satisfaction, an additional degree of freedom to reduce the conservatism induced by robustness. In the limiting case, satisfying a constraint with certainty is equivalent to min-max formulation. However if one allows a sufficiently small margin of not satisfying the constraints (for the infrequent tail events), this tolerance can improve the closed-loop performance a lot while the system is robustified against frequently observed realizations. One possible CC-MPC problem can be provided as

$$\mathcal{P}^{CC}: \left\{ \begin{array}{lll} \min\limits_{u_{[0,N_p-1]|k},\gamma} & \gamma & \\ \text{st.} & \mathbb{P}\left\{ \sum\limits_{j=0}^{N_p-1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k} + x_{N_p|k}^\top Q_f x_{N_p|k} \leq \gamma \right\} \geq \alpha_0 & \\ & x_{j+1|k} = A(\delta_{j|k})x_{j|k} + B(\delta_{j|k})u_{j|k} + F w_{j|k}, & \\ & y_{j|k} = C(\delta_{j|k})x_{j|k} + D(\delta_{j|k})v_{j|k}, & \\ & \mathbb{P}\{c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k}, y_{j|k}) \leq 0\} \geq \alpha_{ij}, & \\ & i = 1, \ldots, N_{c_i^j}, \quad j = 0, \ldots, N_p, \quad x_{0|k} = x_k. & \end{array} \right. \tag{2.14}$$

In this case the associated risk mappings $\mathcal{R}_{CC}^{cost}$ and $\mathcal{R}_{CC}^{const}$ are defined as

$$\begin{aligned} \mathcal{R}_{CC}^{cost}(J(\zeta) := & \quad \mathbb{P}\left\{ \sum_{j=0}^{N_p-1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k} + x_{N_p|k}^\top Q_f x_{N_p|k} \leq \gamma \right\} \geq \alpha_0, \\ \mathcal{R}_{CC}^{const}(c_{ij}(\zeta)) := & \quad \mathbb{P}\{c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k}, y_{j|k}) \leq 0\} \geq \alpha_{ij}. \end{aligned}$$

Closed-loop performance aspects of CC-MPC is yet to be established rigorously, though improvements in closed-loop performance compared to WC-MPC are already reported in the literature ([8]). Regarding the constraint satisfaction, the violation margins α_{ij} determine exactly the allowed frequency of violations. The correlations between the constraints,

if the constraints are to be satisfied jointly, i.e.,

$$\mathbb{P}\left\{c_{ij}(x_{j|k}, u_{j|k}, \zeta_{j|k}) \leq 0, i = 1, \ldots, N_c^j, j = 0, \ldots, N_p\right\} \geq \alpha,$$

is generally a source of conservatism. For using the individual constraints, the computationally efficient case, the approximations of the joint constraints are required. These approximation methods are heuristic and generally highly conservative. See [255] or [383] for the exactness of these relaxations.

The robust stability aspect is slightly vague in comparison to the WC-MPC formulations. In literature, researchers se;ect, in general, a moment-based cost function with chance constraints. The main problem is establishing the stochastic stability within the MPC context, while the implications of employing a specific stability approach is still open to discussion[8].

The possible reduction in conservatism comes with a price. In general the chance constraints are non-convex functions, see Table 2.1, hence computational problems should be expected. Furthermore it is an inherently difficult task to optimize w.r.t the joint pdfs of uncertain functions, since the probability levels after propagation of pdf through the dynamics and the gradients of constraints over these pdfs are required to be calculated. This induces severe computational problems for nonlinear dynamical systems. Another point is related to the modeling aspect of the uncertainties. Modeling the exact pdf of the uncertain effect is an impossible task, while in $\mathcal{P}^{CC}$ one necessarily assumes this information. To overcome this drawback, some methods are introduced to guarantee robustness with respect to a class of pdfs, [32], instead of only one pdf.

Randomized or Scenario Based MPC

Sampling the uncertainty space and conducting the optimization problem with respect to the selected realizations is a recently appreciated technique in MPC domain for solving uncertain optimization problems ([70]), also see the Monte Carlo sampling methods ([303, 329]). Although providing increasingly good approximating solutions to the WC-MPC problem, due to the requirement of 'excessive' number of scenarios, the scenario based methods were not popular until recently in the MPC community.

Similar to previous methods, the randomized MPC techniques consists of two steps, constructing (or sampling) the uncertain space and establishing robustness properties of MPC towards them. By extracting (or generating) a number of scenarios, one samples the uncertain space. This replaces the uncertainties with deterministic values for which the MPC problem can be robustified systematically and easily. A great effort has been directed to a-priori guarantees (for convex problems) and a-posterior guarantees on the probability levels of constraint satisfaction for all uncertainties through the finitely many generated scenarios, see Table 2.1.

[8] For an elegant discussion on various different formulations of stochastic stability, we refer to [191].

Here we construct the randomized MPC, for the worst case formulation, as

$$\mathcal{P}^{rWC}: \begin{cases} \min\limits_{u_{[0,N_p-1|k},\gamma} & \gamma, \\ \text{st.} & x^s_{j+1|k} = A(\delta^s_{j|k})x^s_{j|k} + B(\delta^s_{j|k})u_{j|k} + Fw^s_{j|k}, \\ & y^s_{j|k} = C(\delta^s_{j|k})x_{j|k} + D(\delta^s_{j|k})v^s_{j|k}, \\ & \sum\limits_{j=0}^{N_p-1} (x^s_{j|k})^\top Q x^s_{j|k} + u^\top_{j|k} R u_{j|k} + (x^s_{N_p|k})^\top Q x^s_{N_p|k} \leq \gamma, \\ & c_{ij}(x^s_{j|k}, u_{j|k}, \zeta^s_{j|k}, y^s_{j|k}) \leq 0, \\ & i = 1, \dots, N^j_c,\ j = 0, \dots, N_p,\ s = 1, \dots, N_s,\ x^s_{0|k} = x_k, \end{cases} \tag{2.15}$$

where $s \in \mathbb{Z}_{[1,N_s]}$ is the index running over the selected samples of uncertainty $\zeta^s_{j|k}$ for $j \in \mathbb{Z}_{[0,N_p]}$ and N_s is the total number of scenarios. Here the risk mappings are taken as

$$\begin{aligned} \mathcal{R}^{cost}_{rand}(J(\zeta)) &:= \sum_{j=0}^{N_p-1} (x^s_{j|k})^\top Q x^s_{j|k} + u^\top_{j|k} R u_{j|k} + (x^s_{N_p|k})^\top Q x^s_{N_p|k}, \\ \mathcal{R}^{const}_{rand}(c_{ij}(\zeta)) &:= c_i(x^s_{j|k}, u_{j|k}, \zeta^s_{j|k}, y^s_{j|k}) \leq 0, s = 1, \dots, N_s. \end{aligned}$$

This problem both guarantees the robustness w.r.t. N_s selected scenarios and provides a probabilistic bound on the violation of constraints for unconsidered scenarios, i.e.,

$$\mathbf{P}(\mathcal{P}^{rWC}, \epsilon) := \mathbb{P}\{\mathbb{P}\{c_i(x_{j|k}, u^*_{j|k}, \zeta_{j|k}) \geq 0, \zeta \in \Delta, i = 0, 1, \dots, N_{c_i}, j = 0, \dots, N_p\} > \epsilon\}.$$

The $\epsilon-$constraint violation probability $\mathbf{P}(\mathcal{P}^{rWC}, \epsilon)$ is bounded from above and below as, [66],

$$(1-\epsilon)^{N_s} \leq \mathbf{P}(\mathcal{P}^{rWC}, \epsilon) \leq \beta(N_s, \epsilon), \tag{2.16}$$

where the upper bound $\beta(N_s, \epsilon)$ is a random variable with the Bernoulli distribution, i.e.;

$$\beta(N_s, \epsilon) := \sum_{j=0}^{N_s-1} \binom{N_s}{j} \epsilon^j (1-\epsilon)^{N_s-j}.$$

In this formulation $\beta(N_s, \epsilon)$ is the design parameter to set the probability of constraint violation, ϵ. For small values of $\beta(N_s, \epsilon)$, N_s has a logarithmic growth and as N_s gets larger $\beta(\epsilon)$ tends to zero.

Relatively recently in [71], the scenario based optimization formulation is also discussed in the context of non-convex problems and also reducing the conservatism in the provided violation probability bounds β, with similar discussion also reported in [141, 148, 379]. One remarkable extension that scenario based approach provides is the relatively easy incorporation of various type of uncertainties effecting the dynamical system. Once the scenarios are constructed, the problem $\mathcal{P}^{rWC}$ in Equation (2.15) becomes a nominal MPC problem, with a large number of constraints, increasing the applicability with the current solvers, see Table 2.1. Furthermore, similar to the randomized worst case approach ([65]), one can cast the chance constrained MPC problems in randomized fashion ([323]), where the randomized MPC problem has a similar structure.

Risk and Deviation Metrics

The so-called coherent risk measures are raising considerable attention in operations research and in economics communities, while incorporation into SMPC is already reported in [88]. Following the previous SMPC constructions, one can observe that; (i) in the moment MPC, the constraints in $\mathcal{P}^M$ or $\mathcal{P}^{MV}$ are formulated from risk-natural metrics, not considering asymmetries of the pdfs of constraint functions, hence leading to a possible performance decrease; (ii) the chance constraints are computationally difficult to calculate, due to the non-convexity properties. The use of various other reformulations being developed in the risk theory area, such as semi-deviations or conditional expectations, which can improve both the closed-loop performance and the computational aspects of the resulting SMPC problem.

In [309], the equivalence between chance constraints and a risk indicator, the notion of value at risk (VaR), is established. The $\text{VaR}_\alpha(c_{ij}(\zeta))$ is defined, for a risk level of α as

$$VaR_\alpha(c_{ij}(\zeta)) := \min\{\gamma \in \mathbb{R} | \mathbb{P}\{c_{ij}(\zeta) > \gamma\} \leq 1-\alpha\}.$$

The VaR values consider one side of the (cost or constraint) pdf, so they are not risk-neutral. Same authors have shown that the conditional VaR (CVaR), also called as integrated chance constraints, are behaving far better than the VaR constraints in the optimization problems since; (i) the feasible set is convex, see [39], and (ii) the CVaR values are relatively easy to calculate through scenarios. The α-level CVaR of an uncertain function $c_{ij}(\zeta)$ is defined as,

$$\text{CVaR}_\alpha(c_{ij}(\zeta)) := \frac{1}{1-\alpha} \int_{1-\alpha}^{1} \text{VaR}_\beta(c_{ij}(\zeta)) d\beta.$$

In [309], the authors provide different algorithms for expressing CVaR. One drawback is that integrated chance constraints are more conservative than chance constraint formulations, since $\text{CVaR}_\alpha(c_{ij}(\zeta)) \geq \text{VaR}_\alpha(c_{ij}(\zeta))$ for the positive tail of the pdf, see Figure 2.7. However this implies that the constraint satisfaction guarantees for $\text{CVaR}_\alpha(c_{ij}(\zeta))$ are also valid for the chance constraints (VaR constraints) with the violation level of α. Similar to replacing the chance constraints (or moment formulations) with integrated chance (CVaR) constrains, one can make use of other risk or deviation measures such as a re-formulation of $\mathcal{P}^{MV}$,

$$\mathcal{P}^{Risk}: \begin{cases} \min\limits_{u_{[0,N_p-1]|k}} & \tilde{\mathbb{E}}\left\{J(x_{j|k}, u_{j|k}, \zeta_{j|k})\right\} + \lambda_0 \tilde{\mathbb{D}}\left\{J(x_{j|k}, u_{j|k}, \zeta_{j|k})\right\} \\ \text{st.} & x_{j+1|k} = A(\delta_{j|k})x_{j|k} + B(\delta_{j|k})u_{j|k} + Fw_{j|k}, \\ & y_{j|k} = C(\delta_{j|k})x_{j|k} + D(\delta_{j|k})v_{j|k}, \\ & \tilde{\mathbb{E}}\{c_i(x_{j|k}, u_{j|k}, \zeta_{j|k}, y_{j|k}) + \lambda_{i,j}\tilde{\mathbb{D}}\{c_i(x_{j|k}, u_{j|k}, \zeta_{j|k}, y_{j|k})\} \leq 0, \\ & i = 1, \ldots, N_{c_i}, \qquad j = 0, \ldots, N_p, \end{cases} \tag{2.17}$$

where $\tilde{\mathbb{E}}\{\cdot\}$ and $\tilde{\mathbb{D}}\{\cdot\}$ are generalized risk metric and generalized deviation metric, respectively, instead of expectation or variance operators. Rigorous definitions of generalized risk and deviation metrics are provided in [308, 310], while some examples of these metrics

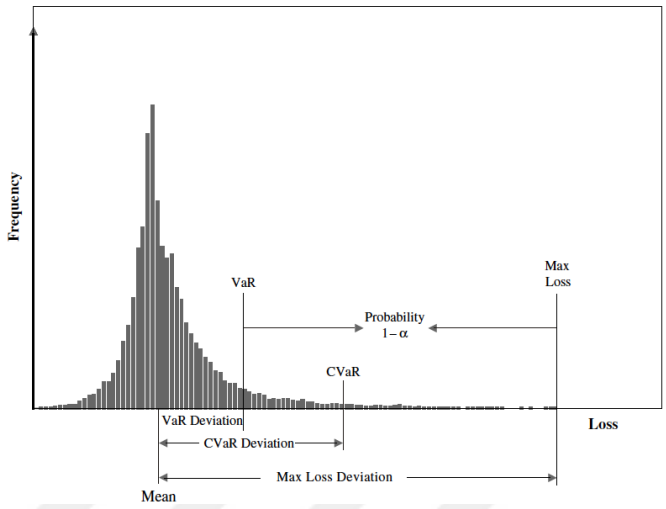


Figure 2.7: Various risk and deviation metrics, taken from [319].

from the literature can be given as,

$$
\begin{aligned}
&\text{Entropic VaR:} && EVaR_\alpha := \min_{\gamma} \left\{ \gamma \in \mathbb{R}_{>0} | \frac{\ln(\frac{M_\zeta(\gamma)}{1-\alpha})}{\gamma} \right\}, \\
&\text{Semideviations:} && \sigma_+^2 := \mathbb{E}\left\{ \max\{\zeta - \mathbb{E}\{\zeta\}, 0\}^2 \right\}, \\
& && \sigma_-^2 := \mathbb{E}\left\{ \max\{\mathbb{E}\{\zeta\} - \zeta, 0\}^2 \right\}, \\
&\text{Maximum Deviations:} && \rho_+ := \max_{\zeta} \zeta - \mathbb{E}\{\zeta\}, \\
& && \rho_- := \max_{\zeta} \mathbb{E}\{\zeta\} - \zeta, \\
&\text{Mean Absolute Deviation:} && \mathrm{MAD}(\zeta) := \int_{\Xi} |\zeta - \mathbb{E}\{\zeta\}| d\zeta.
\end{aligned}
\tag{2.18}
$$

Contributions From SMPC Literature

In this section, we present and classify various relevant entries from the SMPC literature. Similar to the WC-MPC case and following a similar classification with [246], in Table 2.4 we distinguish the publications with respect to the underlying system class (linearity), type of uncertainties effecting the dynamics, how uncertainties are treated or integrated, control parameterization, how the robust counterparts of constraints are expressed and lastly the complexity of simulation examples provided in these papers.

As mentioned above, the extensions of SMPC methods to nonlinear systems are difficult to achieve, since moments (for $\mathcal{P}^M$ or $\mathcal{P}^{MV}$) and pdf calculations (for $\mathcal{P}^{CC}$) are quite difficult to evaluate. Scenario based methods dominate the cases where it is difficult to obtain the pdfs to treat the uncertainty in the optimization problems. Lastly how constraints should be treated and, similar to the RMPC case, the closed-loop performance aspects to practical examples are still yet to be investigated further.

Next we point out some specific contributions outside of the SMPC research area. The

System Dynamics and Control Implementation	Linear		Nonlinear		
	State Feedback	Output Feedback	State Feedback	Output Feedback	
	The majority of the papers	[355], [275], [150], [326], [211], [356], [18], [163]	[369], [53], [93], [134], [31], [371], [228], [85], [52], [177]	[357],	
Uncertainty types	Additive External	Multiplicative External	Multiplicative Time invariant		
	[50], [323], [275], [150], [356], [18], [163], [93], [371], [228], [52], [177], [357], [184], [162], [84], [284], [20], [51], [262], [89], [227], [119], [121], [144], [76], [183], [380], [378], [324], [110], [27], [260], [360], [259], [155], [75]	[88], [369], [93], [134], [31], [85], [52], [177], [121], [27], [122], [153], [68], [67], [124], [287], [38], [108]	[326], [210], [92]		
Uncertainty Treatment	Integration/Expectation	Scenario	Probabilistic	Other	
	[275], [150], [326], [211], [356], [18], [369], [85], [184], [162], [84], [284], [51], [262], [89], [227], [119], [121], [144], [76], [183], [360], [259], [155], [122], [287], [92], [120],	[50], [88], [71], [141], [148], [323], [18], [53], [93], [134], [31], [52], [177], [357], [380], [378], [324], [27], [153], [68], [67], [38], [108],	[228]	[371], [20], [260], [19],	
Control Parameterization	Control Sequence	Affine Disturbance Feedback Parameterization	Control Policy		
	[50], [211], [18], [369], [53], [93], [31], [371], [228], [52], [177], [51], [262], [144], [380], [378], [110], [155], [153], [68], [67], [38], [108],	[323], [356], [163], [357], [184], [84], [284], [227], [119], [121], [76], [183], [380], [378], [27], [260], [360], [259], [75], [122], [287], [376], [189], [120], [124],	[275], [150], [18], [163], [134], [85], [20], [324], [19],		
Constraints	Chance Constraints via Relaxations	Chance Constraints via Moments	Worst Case Constraints via Samples	Chance Constraints via Samples	Other
	[323], [184], [89], [227], [119], [144], [260], [360], [155], [122], [124], [287], [120],	[326], [211], [369], [53], [357], [262], [89], [360], [75], [189],	[93], [284], [380], [110], [153], [67],	[50], [51], [378], [324], [68] [108],	[88], [275], [150], [356], [18], [163], [134], [31], [371], [228], [85], [162], [84], [20], [183], [27], [38], [92], [19],
Complexity of simulation example	No example	1-2 State	3-4 State	5 or more states	
	[356], [53], [85], [155]	[50], [88], [323], [275], [150], [369], [228], [177], [184], [84], [20], [51], [227], [121], [76], [183], [110], [27], [259], [75], [122], [153], [68], [67], [124], [287], [38], [108], [120], [19], [376], [189],	[163], [93], [371], [262], [119], [380], [378], [360], [92],	[326], [211], [52], [357], [284], [89], [144], [324],		

Table 2.4: An overview of SMPC approaches, via a group of selected papers.

expectation and/or variance of the cost functions are discussed in [158, 253, 286, 314] with or without chance constraints. The chance constrained programming (CCP) approach is primarily discussed in [83] and [285]. The CCP algorithms gather attention for optimal design or operation problems also, see [8, 9, 17] for an elegant discussion and implementation considering nonlinear chemical systems. One of the large-scale implementations for SMPC is presented in [144] which, uniquely, considers a linear DAE system with chance constraints.

The randomized algorithms are discussed in depth in [348] or [71, 148]. In order to further improve the bounds or decrease the number N_s of scenarios, one can refine the selection method to sample scenarios, which is not discussed here. One such method is presented in [281], other than the well known techniques such as Latin hypercube or Metropolis-Hastings sampling methods ([8]). Furthermore, scenario based MPC is used for linear parameter varying systems in [324], for a vehicle scheduling study in [134], for finance and econometrics areas in [27, 31, 159], for electricity grids in [152], and lastly for the automotive industry in [110].

2.2.3 Simulation Examples Comparing Different Robust MPC Formulations

In this section, we implement the robust and stochastic MPC algorithms detailed in previous sections to compare their closed-loop and computational performances. The examples are selected mainly from literature, to show the effect of different risk mappings. The first example considers a batch process, where the moment and worst case based methods show opposite behaviour. In the second example we consider an operating point change for an CSTR system. The effect of this change is discussed for different reformulations. Last two examples, mass-spring system and a MIMO debutanizer system, demonstrate the effect of

tuning variables in different risk mappings.

For all of the simulation examples, the optimization routines are implemented via YALMIP ([218]), in Matlab 2015b environment on a computer with 32GB of RAM. The (averaged) computational time, in seconds, are provided in Table 2.5 for the simulation examples. The

	Mean MPC	Mean-Variance MPC	Chance Constrained MPC	Worst-case MPC	Budgetted Worst-case MPC
Semi-batch Reaction	9.3904(s)	14.5054(s)	-	4345.1(s)	3682.6(s)
van der Vusse Reaction	10.6251(s)	12.7974(s)	9.6846(s)	2134.2(s)	1482.5(s)
MSD System	7.3677(s)	7.5501(s)	7.9003(s)	20.3868(s)	19.9146(s)
Debutanizer System	12.3937(s)	12.9231(s)	1266.6(s)	3950.8(s)	3973.6(s)

Table 2.5: Computational times for different examples

results in Table 2.5 are consistent with the theoretical expectations, see Table 2.1. The moment-based robustness evaluations ($\mathcal{P}^M$ and $\mathcal{P}^{MV}$) yield much simpler optimization problems that are solved at each iteration in comparison to the chance constrained (due to non-convexity) and set based techniques, i.e., (budgeted) worst case (due to computational requirements, caused either by excessive randomization or explicit maximization of cost or constraints).

Predictive Control of a Simple Batch Reaction

In this example, we consider a simple exothermic chemical reaction, $A \rightarrow P$. The reaction is controlled through a cooling action. Cooling the reaction temperature is slowing down the reaction rate (exothermic reaction) hence increasing the final time of the batch. Therefore, extensive use of control is not desired. We assume that the reaction dynamics are not known exactly and the uncertain effects are modelled as additive disturbances, with i.i.d Gaussian characteristics ($\mathcal{N}(0, I_2)$). Since the purpose of this process is to finalize the reaction as fast as possible, a *batch-time minimization* problem, while operating within the allowable temperature levels, the pessimistic nature of deterministic MPC constructions can be easily observed. We implement the mean (M), mean-variance (MV), worst-case (WC), and budgeted-robust (BWC) MPC controllers for this example. The reaction dynamics are taken from [256] as

$$\begin{aligned} \dot{c}_A = -r, \quad \dot{c}_P = r, \\ \varrho \bar{c} \dot{T} = \quad -\Delta h_R r + \tfrac{U_W A_W}{V_R}(T_C - T) + f_2 w^2, \\ r = \quad c_A k_{300} \exp(\tfrac{E}{R}\left(\tfrac{1}{300} - \tfrac{1}{T}\right) + f_1 w^1), \end{aligned} \tag{2.19}$$

where c_A is the reactant concentration, c_P is the product concentration, T is the reactor temperature and the T_C is the control input, the cooling temperature. The initial values of states are taken as $x_0^\top = \begin{bmatrix} 5000 & 0 & 40 \end{bmatrix}$. The parameter values are given in Table 2.6. For the set-based methods, we consider bounded support for the uncertainty $w_k \subset \mathcal{W}^{(\cdot)}$, $k \in \mathbb{Z}_{\geq 0}$. For WC-MPC case, the disturbances assume values from the interval

$V_R =$	6.28 $[m^3]$	$A_W =$	10 $[m^2]$
$U_W =$	500 $[W/m^2/K]$	$\varrho =$	1000 $[kg/m^3]$
$\bar{c} =$	2 $[kJ/kgK]$	$\Delta h_R =$	-10^5 $[kJ/kmol]$
$E =$	40000 $[kJ/kmol]$	$R =$	8.314 $kJ/kmolK$
$k_{300} =$	0.125 [1/h]		

Table 2.6: Semi-batch Reaction Parameters

$\mathcal{W}^{WC} = [-3, 3] \times [-3, 3]$, which is equal to 6 standard deviations. For the BWC-MPC case the disturbance set is shrunk to $\mathcal{W}^{BWC} = [-1, 1] \times [-1, 1]$, since this set incorporates many of the observed realizations already. Furthermore for the MV-MPC we set λ_0 to 2.5, which is tuned to improve the distinction from mean case, while $\lambda_{ij},\ i = 1, \ldots, N_c^j,\ j = 1, \ldots N_p$, values are set to zero, hence no constraint tightening is imposed for MV-MPC formulation. The prediction model is set as the linearized and discretized (with sampling time of 5 seconds) dynamics over an operating trajectory, i.e., a high level deterministic dynamic optimization problem with the nonlinear model is solved to provide the operating conditions which minimize the cost function

$$J^1 = \int_0^{0.75} \alpha_a c_a(t) + \alpha_T T_C^2(t) dt,$$

where c_a and T_C are state variables, $\alpha_a = 1$ and $\alpha_T = 0.001$ are the state and input weighting terms in the optimization problem, respectively. The result of the high-level optimization problem is used as the operating points, hence the linearized dynamics are known to the MPC controllers. The prediction horizon for the MPC problem is taken as 5 time steps to discard the possible effects of linearization and mismatch in the operating points. The optimization problem for the MPC algorithm is taken as,

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} \quad & \mathcal{R}_i\{\textstyle\sum_{j=0}^{N_p-1} x_{j|k}^T Q x_{j|k} + u_{j|k}^T I u_{j|k}\}, \\ \text{s.t.} \quad & \Delta x_{j+1|k} = A_{j|k}\Delta x_{j|k} + B_{j|k}\Delta u_{j|k} + F_{j|k} w_{j|k}, \\ & \begin{bmatrix} x_{k+j} \\ u_{k+j} \end{bmatrix} \in \begin{bmatrix} \mathbb{X} \\ \mathbb{U} \end{bmatrix}, \end{aligned}$$

where the state constraints are taken as

$$\mathbb{X} := \{(x_1,\ x_2,\ x_3) |\ 0 \le x_1 \le 5000, 0 \le x_2 \le 5000, -10 \le x_3 \le 350\},$$

the input constraints are taken as $\mathbb{U} := \{u |\ 0 \le u \le 100\}$,. Lastly the weighting matrix Q is taken as $\text{diag}(1, 0, 10^4)$.

In Figure 2.8, we visualize the performance for the RMPC solutions without any disturbance effects, i.e., $w^i(t) = 0$ for $i = 1, 2$ and $t \in [0, t_f]$ where t_f is the final time of the reaction. The results indicate the structural differences between the RMPC techniques; the min-max based techniques resulting in longer batch final times, compared to Mean-MPC case, due to the control action cooling down the reaction not to exceed the temperature constraints. Totally opposite behaviour is observed in the MV-MPC case, which greatly speeds

up the reaction. This is due to the instability of the system, the variance term dominates the cost function, hence the controller acts aggressively to compensate the actions induced from uncertain perturbations in prediction stages. We also provide the results for the case where disturbances are not set to zero in Figure 2.9, which demonstrates that the reactor temperature is kept in allowable bounds, in almost all cases. Similar to the uncertainty-free case, WC-MPC and BWC-MPC react similar to Mean-MPC case at the start of the simulation, but as the state constraints start to be effective, the controllers slow down the reaction, which leads to longer total batch time compared to M-MPC case, while the MV-MPC, again, speeds up the reaction. This behaviour leads to higher temperatures over the reaction trajectory, in one case a violation of temperature constraint is observed. This is an expected result since for the moment MPC formulations, the constraints are allowed to be violated.

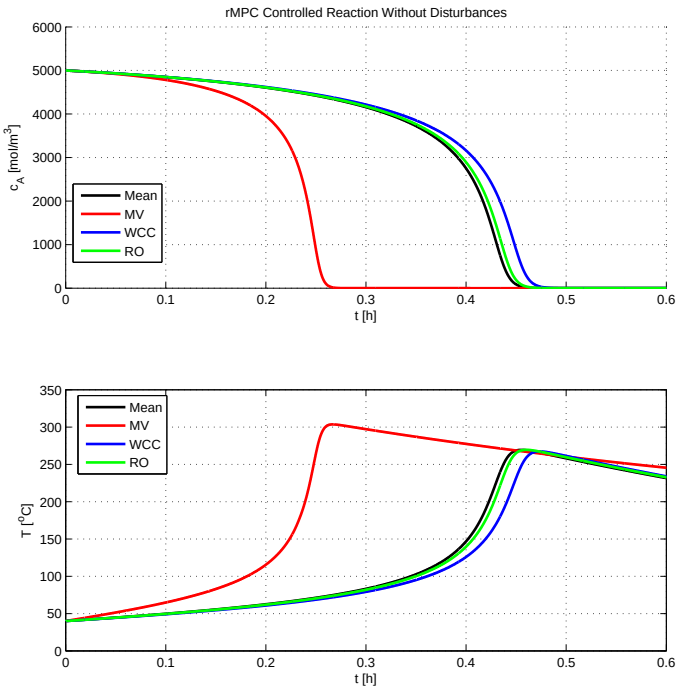


Figure 2.8: The concentration and temperature profiles with RMPC controllers for disturbance free case.

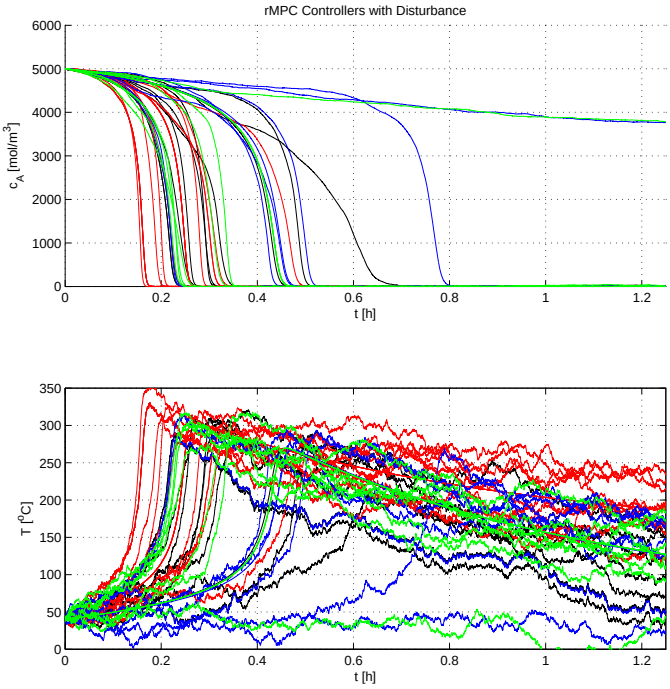


Figure 2.9: The concentration and temperature profiles with RMPC controllers with the effect of disturbances.

Control of CSTR with Changing Operation Point

As a second example, we consider the van der Vusse reaction, taken from [202],

$$\begin{aligned} \dot{C}_a &= -k_1(T)C_a - k_2(T)C_a^2 + (C_{in} - C_a)u_1, \\ \dot{C}_b &= k_1(T)(C_a - C_b) - C_b u_1, \\ \dot{T} &= -f_{react} + \alpha(T_c - T) + (T_{in} - T)u_1, \\ \dot{T}_c &= \beta(T - T_c) + \gamma u_2, \\ f_{react} &= \delta(k_1(T)(C_a H_{ab} + C_b H_{bc}) + k_2(T)C_a^2 H_{ad}), \end{aligned} \tag{2.20}$$

where the states are the concentration of substance A (C_a), the concentration of substance B (C_b), the temperature of reactor (T) and temperature of cooling jacket (T_c). The inputs are the flow rate to the reactor (u_1) and the cooling power (u_2). The parameter values are given in Table 2.7. We construct a scenario where the known disturbances, the inflow rate C_{in} and the inlet temperature T_{in}, are assumed to change over to a different operating point after the first hour of the reaction. Before and after the switching we regulate the state trajectories to

$\alpha =$	30.8285 [1/h]	$\beta =$	86.688 [1/h]
$\gamma =$	0.1 [K/kJ]	$\delta =$	0.3556 [Kl/kJ]
$E_1 =$	9758.3 [ ]	$E_2 =$	8560.03[ ]
$k_{10} =$	$1.287 * 10^{12}$ [l/h]	$k_{20} =$	$9.043 * 10^{9}$ l/(mol*h)
$h_{AB} =$	4.2 [kJ/mol]	$h_{BC} =$	-11 [kJ/mol]
$h_{AD} =$	-41.85 [kJ/mol]		

Table 2.7: Van der Vusse Reaction Parameters

known operating points, ρ_1 and ρ_2;

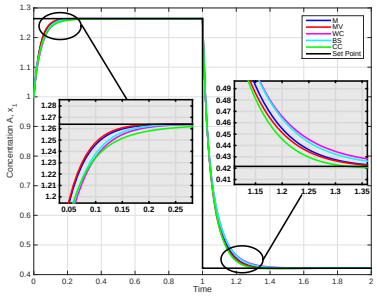
$$\begin{aligned} \rho_1^\top &= \begin{bmatrix} C_{in}(1) & T_{in}(1) & u_1^{eq}(1) & u_2^{eq}(1) & x_1^{eq}(1) & x_2^{eq}(1) & x_3^{eq}(1) & x_4^{eq}(1) \end{bmatrix}, \\ &= \begin{bmatrix} 5.1 & 104.9 & 4.19 & -1113.5 & 1.2639 & 0.9049 & 109.2881 & 108.0037 \end{bmatrix}, \\ \rho_2^\top &= \begin{bmatrix} 1.1 & 109.9 & 4.19 & -1113.5 & 0.4216 & 0.2530 & 101.6720 & 100.3875 \end{bmatrix}, \end{aligned} \tag{2.21}$$

We assume i.i.d. Gaussian additive uncertainty effecting the dynamics with covariance matrix being equal to $\Sigma := \text{diag}(.1,\ .1,\ 5,\ 5)$, 10% error rate for the concentration variables and 2.5% for the temperature measurements. The cost functional is selected as quadratic cost function with weighting terms

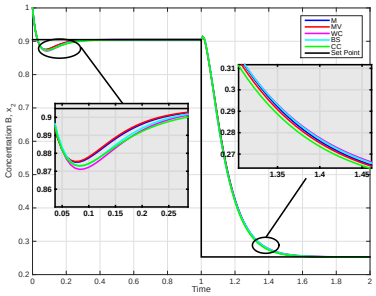
$$Q = 10\ I_4, \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 10^{-6} \end{bmatrix}, \quad Q_f = 0.$$

The state and input constraints are selected as polytopes and defined by $0 \leq x_1 \leq 1.6,\ 0 \leq x_2 \leq 1.6,\ 80 \leq x_3 \leq 130,\ 80 \leq x_4 \leq 130$, and $3 \leq u_1 \leq 35,\ -9000 \leq u_2 \leq 0$.

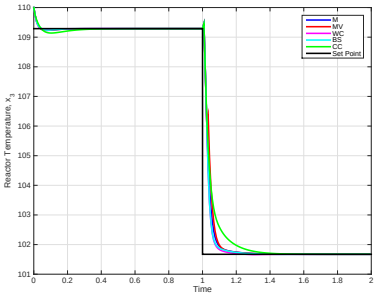
With these information, first we demonstrate the results obtained from the disturbance free case, where we set the uncertainty equal to zero for all time instants, see Figure 2.10. All robustness methods yield similar performances, since the constraints allow for a sufficiently large operating window for the MPC controllers. If one shrinks the constraint set, first the WC-MPC, then BWC-MPC and CC-MPC turn infeasible. Observe the change of chance constrained MPC behaviour (green trajectories in Figure 2.10) in unstable equilibrium, operating point ρ_1, which leads to slow convergence and in the stable equilibrium, operating point ρ_2, which leads to fast convergence. Similar to the previous case, MV-MPC controlled system acts slightly faster than the Mean-MPC case, while budgeting (BWC-MPC) leads to a higher performance in comparison to the WC-MPC. Secondly we present 100 different realizations of the trajectories of the states with different disturbance sequences, for the RMPC controllers in Figure 2.11. Lastly we visualize histograms of state trajectories for two time instants, the 59^{th} minute of reaction, steady state responses before the switching occur, and the 75^{th} minute of the reaction demonstrating the transient behaviour after the switching. From these histograms we observe that two distinct characteristics of different robustness techniques. Firstly, both before and after the operating point change, the set based robust MPCs (WC and BWC) cause unnecessarily pessimistic trajectories, while the chance constrained MPC allows for aggressive control actions, hence drives the trajectories towards the constraints. Secondly, the moment-based techniques yield comparable trajec-



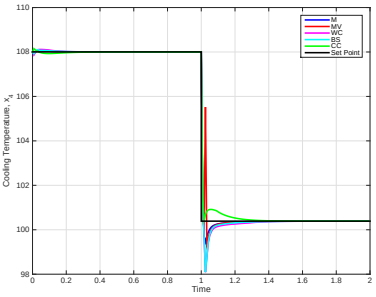
(a) Concentration of component A trajectories.



(b) Concentration of component B trajectories.



(c) Reaction temperature trajectories.



(d) Cooling temperature trajectories.

Figure 2.10: The state trajectories for CSTR simulation without disturbances

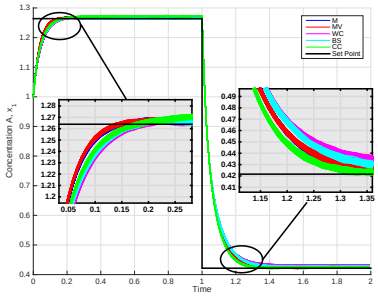
tories in both operating points, indicating necessity for further research on the tuning of λ_0 variable. Rigorous analysis for tuning the λ_0 variable for MV MPC is left as a future study.

A Two Mass and Spring System

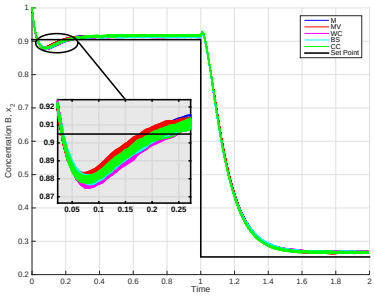
In this example we make use of a mass-spring system, similar to the case in [185], which consists of two masses connected with a spring. The disturbance corrupted control action is effective on one of the masses and the control goal is to effectively stabilize the system. The dynamics of the process is taken as follows;

$$x_{k+1} = Ax_k + B(u_k + w_k),$$
$$A = \begin{bmatrix} 1 & 0 & 0.1 & 0 \\ 0 & 1 & 0 & 0.1 \\ -\frac{K}{m_1} & 0.1\frac{K}{m_1} & 1 & 0 \\ \frac{K}{m_2} & -0.1\frac{K}{m_2} & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ 0.1\frac{1}{m_1} \\ 0 \end{bmatrix}, \tag{2.22}$$

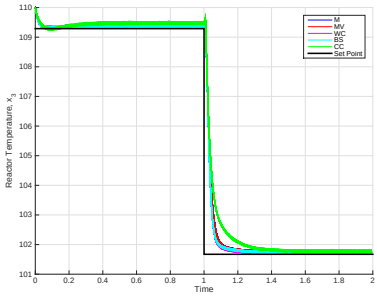
where the parameters, the spring constant K and the masses m_1 and m_2 are taken as; $K = 1$, $m_1 = 0.5$, $m_2 = 2$. We present two sets of solutions, the first case showing



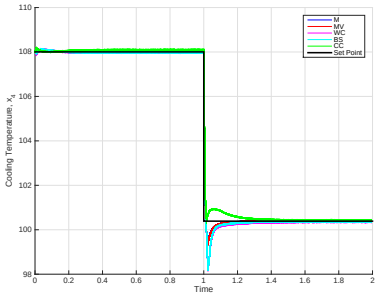
(a) Concentration of component A trajectories.



(b) Concentration of component B trajectories.

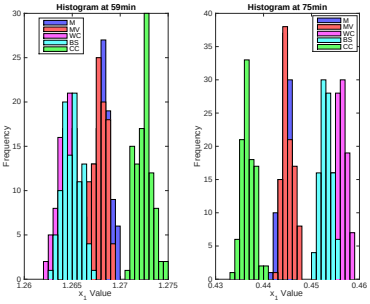


(c) Reaction temperature trajectories.

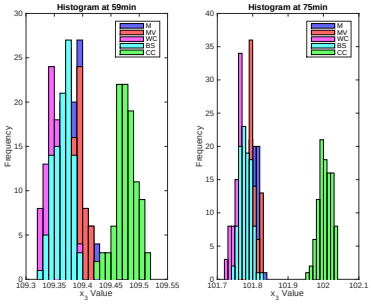


(d) Cooling temperature trajectories.

Figure 2.11: The state trajectories for CSTR simulation with disturbances



(a) Concentration of component A histograms.



(b) Reaction temperature histograms.

Figure 2.12: The state (x_1 and x_3) histograms for CSTR simulation at two different time instants.

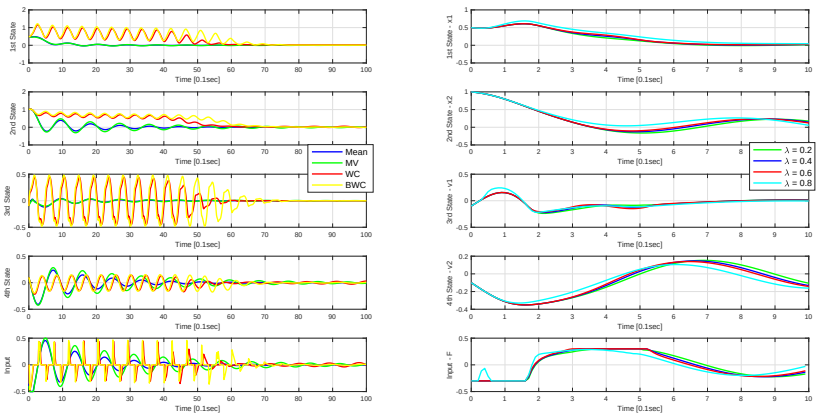
the differences between the closed-loop responses of different predictive controllers (Mean, MV, WC and BWC), in which we skipped the CC-MPC, since the results are almost comparable with the other moment-based formulations. The second study demonstrates the aggressiveness of MV-MPC for different λ_0 values. In both of the simulations we have used the same state and input weighting matrices, such as

$$Q = 5I_4, \quad R = I_1, \quad Q_f = 0.$$

For the first simulation study, we make use of following state and input constraints,

$$\mathbb{X}_1 = \{x | -1 \leq x_1 \leq 1.5, -1 \leq x_2 \leq 1.5, -0.5 \leq x_3 \leq 0.5, -0.5 \leq x_4 \leq 0.5.\},$$
$$\mathbb{U}_1 = \{u | -0.5 \leq u \leq 0.5\},$$

while we initialized the state at $x_0^\top = \begin{bmatrix} 0.5 & 1 & -0.1 & 0.1 \end{bmatrix}$. For comparison purposes, we force the disturbance variables to zero, while the MV-MPC controller assumes that the input disturbance has a standard deviation of 0.15, the WC-MPC and BWC-MPC robustifies the closed-loop system for all of the disturbance realizations between $w_k^{WC} \in \begin{bmatrix} -0.45 & 0.45 \end{bmatrix}$, and $w_k^{BWC} \in \begin{bmatrix} -0.15 & 0.15 \end{bmatrix}$, respectively. The simulated trajectories are visualized in Figure 2.13a. The deterministically robust MPC controllers (WC-MPC and BWC-MPC) are almost inactive during the whole simulation, hence leading to much larger settling times in comparison to the Mean-MPC or the MV-MPC. The second simulation study demonstrates



(a) closed-loop responses for different MPC approaches.

(b) Differences within the closed-loop responses of MV-MPC for different λ_0 values.

Figure 2.13: Simulation results for a two mass-spring system.

the effect of tuning parameter λ_0 for the MV-MPC construction. For this case we reshape the constraints as;

$$\mathbb{X}_2 = \{x | -1.1 \leq x_1 \leq 1.1, -1.1 \leq x_2 \leq 1.1, -0.4 \leq x_3 \leq 0.4, -0.4 \leq x_4 \leq 0.4\},$$

$$\mathbb{U}_2 = \{u | -0.3 \leq u \leq 0.3\},$$

while the λ_0 value assumes values from the set $\{0.2,\ 0.4,\ 0.6,\ 0.8\}$. The constraints are tightened, via the parameter λ_{ij} which is set to 1 for all of the constraints. The simulated trajectories are visualized in Figure 2.13b. For the first three values of λ_0, we observe that the responses are similar to each other, although as the parameter λ_0 increases, the control action is growing in magnitude, and hence causing slight improvements in the state trajectories. The critical change occurs for the case of $\lambda_0 = 0.8$. In this case the optimal input signal differs from the other cases at the initial phase of the simulation, leading to slightly faster response due to the larger effect of variance term in the MPC problem.

Debutanizer System and Budget for Uncertainty

Lastly we provide a comparison study, comparing the achievable closed-loop performances for different uncertainty sets, on a debutanizer system. The simulation example, taken from [311], considers a 2-input 2-output MIMO system, with the transfer function given as

$$G(s) = \begin{bmatrix} \frac{-0.2949}{64.02s^2+61.66s+1} & \frac{0.1310}{854.75s^2+88.03s+1} \\ \frac{0.1287}{168.25s^2+15s+1} & \frac{-0.1434}{32s^2+17.76s+1} \end{bmatrix}, \qquad Y(s) = G(s)(U(s) + W(s)).$$

We first represent this transfer function in state-space from and then apply discretization with a sampling time of 1.8 minutes. We pose hard input and output constraints on a zero initial condition such as

$$\mathbb{Y} = \{y | -1 \leq y_1 \leq 1, -1 \leq y_2 \leq 1, \},$$

$$\mathbb{U} = \{u | -15 \leq u_1 \leq 15, -15 \leq u_2 \leq 15\}.$$

The goal of the process is to reach set point $y_i = 1,\ i = 1, 2$, while due to the constraint, the RMPC controllers back-off the transient trajectories and the final operating point. The cost function is taken as a quadratic form of shifted nominal outputs (towards the desired operating point) and inputs, with the weighting matrices taken as $Q = 10^7 I_2, \quad R = 0.1 I_2$ and the prediction horizon set to $N_p = 10$. We compare the responses for different selection of uncertainty sets, such as the disturbance set is set to $\mathcal{W}^{WC} = \{-50, 50\}$, and for the BWC cases we scale this set with ten different values, i.e., $\mathcal{W}^{\alpha BWC} = \alpha \cdot \{-50, 50\}, \alpha = 0, 1, \ldots, 9$, as visualized in Figure 2.14. As can be seen from the second output trajectories, the WC-MPC and BWC-MPC with budget level of $0.9\mathcal{W}$ hit the tightened constraints during the transition between $t \in [75, 90]$, while for the other uncertainty sets, the MPC controllers never reach to the tightened constraint levels for the second output. Regarding the first output, we can observe that 0.1-BWC MPC and nominal MPC are acting exactly same with with each other, hence meaning that Mean MPC is already providing robust operation for additive uncertainties belonging to the set $\mathcal{W}^{0.1BWC} = [-5, 5]$. Furthermore, one can clearly observe the pessimistic results of WC or BWC MPCs with a large assumed set of uncertain effects, since the settling value for the first output deteriorates as the α increases in α-BWC MPC constructions.

The simulation examples studying the different robust MPC constructions introduced in previous sections for the debutanizer system have confirmed the fact that the performance of deterministically robust MPC approaches (WC-MPC and BWC-MPC) is overly conservative and responds cautiously resulting in larger settling times in comparison to SMPC

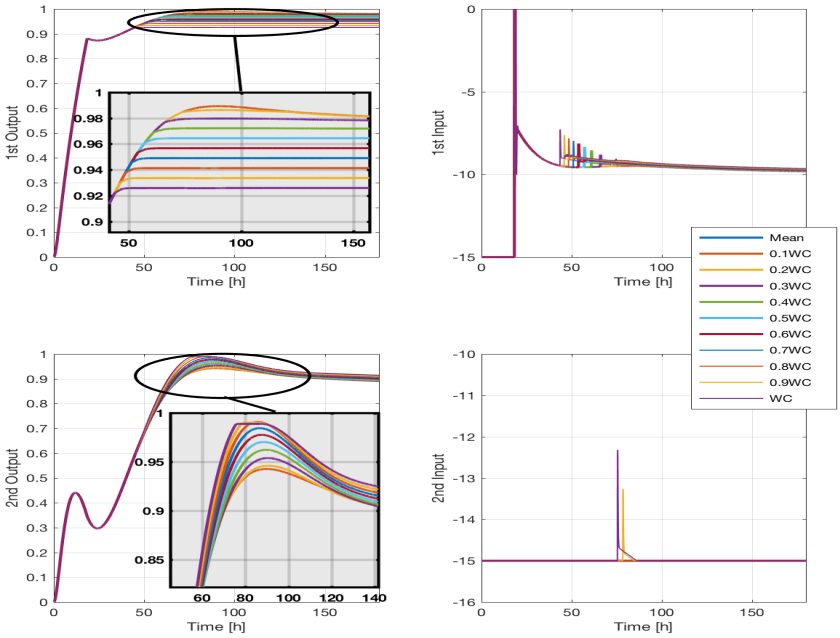


Figure 2.14: Differences within the closed-loop responses of Mean, WC and different BWC-MPC controllers for different $\bar{\Delta}$ sets.

approaches. SPMC methods, on the other hand, have a tendency of providing aggressive control actions leading to constraint violations in some cases.

2.3 Conclusions on the Robust Predictive Control Problems

In this chapter, we present and compare the methods used in the robust and stochastic MPC literature. The vast literature and terminological differences between the branches of robust and stochastic MPC necessitates extensive time and effort for an overall understanding of the topic. This induces a divergence between industrial and academic perspectives on guaranteeing robust operation with MPC controllers. Here, we have discussed various MPC paradigms from the closed-loop pessimism and required computational power perspectives. We present the connections between the robust or risk-aware optimization and predictive control problems, by classifying the contributions from the RMPC and SMPC literature. Furthermore, we apply some these methods to simulation examples, to compare the computational requirements and resulting closed-loop performance.

We highlight some concluding observations as;

- The nominal prediction model is the dominant factor in the induced computational complexity and also the eventual closed-loop performance.

- The price paid for robustness depends on how the uncertainty is evolving in the dynamics. Realistic uncertainty set, distribution or scenario generation is another major factor for obtaining less pessimistic closed-loop operation.
- To improve the response, one should re-evaluate the allocated uncertainty budget for each constraint. If constraint violations are practically allowable, one should prefer chance, moment or scenario constraints rather than the deterministic guarantees.
- Cost functions in MPC problems should incorporate a prediction of operating point/trajectory and also a metric to evaluate the deviations from this point/trajectory.

Chapter 3

Moment-based Model Predictive Control Problem for Linear Systems with Additive Perturbations

Between equal rights, force decides.

Karl Marx - Capital: A Critique of Political Economy

In this chapter, we investigate predictive control of dynamical systems using the statistics of the state predictions and cost functions. We calculate the (centralized) moments of the states based on the model dynamics, which are then used in the MPC problem. We make explicit use of the statistical knowledge of the uncertain state evolution, namely the first, second and third central moments, to project the uncertain cost to a deterministic function standing for the robust counterpart problem[1]. We observe that the resulting cost functions are defined through the nominal system, hence one needs to solve MPC problems which are (theoretically) as complex as the nominal MPC problems. Since the central moments of the disturbances are leading to extra terms in the robust counterpart cost function, the current

[0]Substantial content of this chapter is also published or presented in:

- M.B. Saltik, L. Özkan, S. Weiland. Moment-based Model Predictive Control for Linear Systems Part 1: Additive Perturbations Case;
- M.B. Saltik, L. Özkan, J.H.A. Ludlage, S. Weiland, P.M.J. Van den Hof. On the Moment-based Robust MPC Formulations.
- M.B. Saltik, L. Özkan, S. Weiland, J.H.A. Ludlage. Moment-based Model Predictive Control for Systems with Additive Uncertainty.
- R. Zhang, M.B. Saltik, L. Özkan, Study of Moment-based MPC formulations and their Connection to Classical Control.

[1]The high order moments of the cost function is in many instances much more meaningful indicator of the resulting (true) MPC cost than the commonly used worst-case cost ([250].

control action incorporates of the future uncertainties.

3.1 Effect of Disturbances on Moment-based MPC Formulations

We define a nominal system, Σ^{nom} as in Equation (3.1a) and a system with additive disturbances, Σ^{add} as in Equation (3.1b);

$$\Sigma^{nom} : \begin{cases} \bar{x}_{k+1} = & A\bar{x}_k + Bu_k, \\ \bar{y}_k = & C\bar{x}_k,\ \bar{x}_0 = \bar{x}(0), \end{cases} \tag{3.1a}$$

$$\Sigma^{add} : \begin{cases} x_{k+1} = & Ax_k + Bu_k + Fw_k, \\ y_k = & Cx_k + v_k,\ x_0 = x(0), \end{cases} \tag{3.1b}$$

where x_k (or $\bar{x}_k$) $\in \mathbb{R}^n$, and $u_k \in \mathbb{R}^{n_u}$, and y_k (or $\bar{y}_k$) $\in \mathbb{R}^{n_y}$ are the disturbance corrupted (or nominal) state, the control input and the corrupted (or nominal) output at discrete time instant $k \in \mathbb{Z}_{\geq 0}$, respectively. The uncertain plant is subject to additive disturbances $w_k \in \mathbb{R}^{n_w}$ in the state evolution equation and $v_k \in \mathbb{R}^{n_v}$ in the output equation for $k \in \mathbb{Z}_{\geq 0}$. Here w_k, v_k are assumed to be independent and identically distributed (i.i.d.) over time k, Gaussian vector sequences with zero mean, known and time invariant covariance matrices $\Sigma_w \geq 0$ and $\Sigma_v \geq 0$, respectively.

For the states $x_{[0,N_p-1]|k} := x_{[k,k+N_p-1]}$, the prediction matrices are defined as:

$$\begin{aligned} \Theta_{N_p}^\top &= \begin{bmatrix} I & A^\top & A^{2^\top} & \dots & A^{N_p-1^\top} \end{bmatrix} \in \mathbb{R}^{n_x \times N_p n_x}, \\ T_B^{N_p+1,N_p} &= \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ B & 0 & \dots & 0 & 0 \\ AB & B & \dots & 0 & 0 \\ A^2B & AB & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A^{N_p-1}B & A^{N_p-2}B & \dots & AB & B \end{bmatrix}, \end{aligned}$$

which yield

$$\begin{aligned} \bar{x}_{[0,N_p-1]|k} = &\ \Theta_{N_p}\bar{x}_{0|k} + T_B^{N_p} u_{[0,N_p-1]|k}, \\ x_{[0,N_p-1]|k} = &\ \Theta_{N_p}\bar{x}_{0|k} + T_B^{N_p} u_{[0,N_p-1]|k} + T_F^{N_p} w_{[0,N_p-1]|k}, \end{aligned}$$

where N_p is the prediction horizon and the Toeplitz matrix $T_B^{N_p,N_p}$ is expressed as $T_B^{N_p}$ for brevity. The identity matrix (of dimension N_p) is denoted as I_{N_p}. For a matrix A, $\rho(A)$ denotes the spectral radius of the matrix A.

3.1.1 Problem Formulation for Regulation Problem via Moment-based MPC

In this section we consider a quadratic cost function to be minimized for the MPC problem, which at time instant k is expressed as

$$J(x_k, u_{[0,N_p-1]|k}) = \sum_{j=0}^{N_p-1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k} + x_{N_p|k}^\top Q_f x_{N_p|k}, \tag{3.2a}$$

where $N_p \in \mathbb{Z}_{>0}$ is the prediction horizon, $x_{j|k} \in \mathbb{R}^{n_x}$ is the prediction of state at time instant $(k+j)$ generated with the dynamics Σ^{add}, the initial condition $x_{0|k} = x_k$ and the exogenous inputs $u_{[0,N_p-1]|k}, w_{[0,N_p-1]|k}$. The optimization variables are $u_{j|k}$, $j = 0, 1, \ldots, N_p - 1$[2]. The cost function $J(x_k)$ can be expressed in matrix form as,

$$J(\tilde{\xi}_k) = \tilde{\xi}_k^\top H^{N_p+1} \tilde{\xi}_k \tag{3.2b}$$

where

$$H^{N_p+1} := \begin{bmatrix} \Theta_{(\cdot)}^\top \bar{Q} \Theta_{(\cdot)} & \Theta_{(\cdot)}^\top \bar{Q} T_B^{(\cdot)} & \Theta_{(\cdot)}^\top \bar{Q} T_F^{(\cdot)} \\ {T_B^{(\cdot)}}^\top \bar{Q} \Theta_{(\cdot)} & {T_B^{(\cdot)}}^\top \bar{Q} T_B^{(\cdot)} + \bar{R} & {T_B^{(\cdot)}}^\top \bar{Q} T_F^{(\cdot)} \\ {T_F^{(\cdot)}}^\top \bar{Q} \Theta_{(\cdot)} & {T_F^{(\cdot)}}^\top \bar{Q} T_B^{(\cdot)} & {T_F^{(\cdot)}}^\top \bar{Q} T_F^{(\cdot)} \end{bmatrix},$$

$$\tilde{\xi}_k := \begin{bmatrix} x_{0|k} \\ u_{[0,N_p-1]|k} \\ w_{[0,N_p-1]|k} \end{bmatrix}, \bar{Q} := \begin{bmatrix} I_{N_p} \otimes Q & 0 \\ 0 & Q_f \end{bmatrix}, \bar{R} := I_{N_p} \otimes R,$$

in which Θ_{N_p+1}, $T_B^{N_p+1,N_p}$ and $T_F^{N_p+1,N_p}$ are denoted with $\Theta_{(\cdot)}, T_B^{(\cdot)}$ and $T_F^{(\cdot)}$, respectively.

The cost function in Equation (3.2a) is inherently stochastic, hence it needs to be projected to a deterministic function before the optimization step. Also a robustification, in fact a risk allocation, step is needed for the predicted effects of future uncertainty. For this purpose we use the (centralized) moments of the uncertain cost function $J(x_k)$;

$$J^M(\xi_k) := \mathbb{E}\{J(\tilde{\xi}_k)\}, \tag{3.3a}$$

$$J^{MV}(\xi_k) := \mathbb{E}\{J(\tilde{\xi}_k)\} + \lambda_v \mathbb{D}\{J(\tilde{\xi}_k)\}, \tag{3.3b}$$

$$J^{MVS}(\xi_k) := \mathbb{E}\{J(\tilde{\xi}_k)\} + \lambda_v \mathbb{D}\{J(\tilde{\xi}_k)\} + \lambda_s \mathbb{S}\{J(\tilde{\xi}_k)\}, \tag{3.3c}$$

where $\lambda_v, \lambda_s \in \mathbb{R}_{\geq 0}$ are the risk aversion factors ([229]), $J(\tilde{\xi}_k)$ is the cost function given in Equation (3.2a) and

$$\xi_k := \begin{bmatrix} x_{0|k} \\ u_{[0,N_p-1]|k} \end{bmatrix}.$$

We use these three cost functions to establish mean-square stability (MS-stability)[3] of closed loop stochastic systems. MS-stability is defined as follows.

Definition 3.1.1 *([191]) A system Σ^{add} is said to be Lyapunov stable in the 2^{nd} moment, or Mean Square stable (MS-stable), if the mean and the covariance of x_k exists for all $k \in \mathbb{Z}_{\geq 0}$, and if for all $\epsilon > 0$, there exists δ such that*

$$||x_0||_2^2 < \delta \implies \mathbb{E}\{\sup_{k \geq k_0} ||x(k; x_0)||_2^2\} < \epsilon.$$

[2] One can also employ convexity preserving disturbance feedback structures, i.e., $u_{j|k} = \tilde{u}_{j|k} + \sum_{i=0}^{j-1} M_{j,i} w_{i|k}$, if $w_{i|k}$ are available.

[3] The stability of stochastic systems can be defined from various aspects, such as stability in probabilities or in moments or in all sample behavior. A discussion on the stability of stochastic systems is given in [191].

This notion of MS-stability is closely related to Lyapunov theory of stability, see also [38], as demonstrated next.

Proposition 3.1.1 *[38] For a dynamical system $\Sigma_{CL}^{add} : x_{k+1} = \tilde{A}x_k + Fw_k$ and a quadratic positive definite function $V(x_k) = x_k^\top P x_k$, if $\mathbb{E}_k\{V(x_{k+1})\} - V(x_k) \leq 0$ then the dynamical system is MS-stable.*

Proof. By hypothesis and the definition of conditional expectation,

$$\begin{aligned}\mathbb{E}\{V(x_{k+1})\} = &\quad \mathbb{E}_{[0,k-1]}\{V(x_k) + \mathbb{E}_k\{V(x_{k+1})\} - V(x_k)\},\\ \leq &\quad \mathbb{E}_{k-1}\{V(x_k)\},\end{aligned}$$

and by induction over the time index k

$$\mathbb{E}_k\{V(x_{k+1})\} \leq \quad V(x_0).$$

Then we make use of the properties of weighted-2 norm

$$\lambda_{min}(P)\mathbb{E}\{x_k^\top x_k\} \leq \mathbb{E}\{V(x_k)\} \leq V(x_0) \leq \lambda_{max}(P)x_0^\top x_0$$

for all $k \in \mathbb{Z}_{\geq 0}$. This guarantees MS-stable operation. □

The contribution of this section is twofold. First, for the three cost functions in Equation (3.3), we demonstrate that moment-based MPC problems result in cost functions that can be expressed by nominal system states. This leads to a tremendous reduction in the required computational power, since one needs to solve an MPC problem equivalent with the complexity of the nominal MPC problem. Secondly, high order moment MPC problems provide a natural way of achieving stability without the artificial terminal cost terms, i.e., Q_f in Equation (3.2a) set to zero.

Mean-MPC (M-MPC)

The mean of a stochastic cost function provides an indication of the cost on average, hence the spread (for quadratic $J(x_k)$) and the magnitude of the states on average. Here we start with evaluating the cost function given in Equation (3.3a).

Lemma 3.1.1 *For the stochastic dynamical system Σ^{add} in Equation* (3.1b)*, the cost function for M-MPC problem (Equation* (3.3a)*) is given by*

$$J^M(\xi_k) = \xi_k^\top H_M^{N_p+1}\xi_k + f^M \tag{3.4}$$

where

$$\begin{aligned}\xi_k :=&\quad \begin{bmatrix} x_{0|k} \\ u_{[0,N_p-1]|k}\end{bmatrix},\ f^M := Tr({T_F^{(\cdot)}}^\top \bar{Q} T_F^{(\cdot)} \bar{\Sigma}_w^{N_p+1}),\\ H_M^{N_p+1} :=&\quad \begin{bmatrix} \Theta_{N_p+1}^\top \bar{Q}\Theta_{N_p+1} & \Theta_{N_p+1}^\top \bar{Q} T_B^{N_p+1,N_p} \\ T_B^{N_p+1,N_p^\top}\bar{Q}\Theta_{N_p+1} & T_B^{N_p+1,N_p^\top}\bar{Q}T_B^{N_p+1,N_p} + \bar{R}\end{bmatrix},\end{aligned} \tag{3.5}$$

where $T_F^{N_p+1,N_p}$ is denoted with $T_F^{(\cdot)}$.

Furthermore,

$$\bar{J}^M(\xi_k) := J^M(\tilde{\xi}_k) - Tr(T_F^{N_p+1,N_p\top}\bar{Q}T_F^{N_p+1,N_p}\bar{\Sigma}_w^{N_p}), \tag{3.6a}$$

$$\begin{aligned}\bar{J}^M(\xi_k) = &\quad \xi_k^\top H_M^{N_p+1}\xi_k,\\ \bar{J}^M(\xi_k) = &\sum_{j=0}^{N_p-1} \bar{x}_{j|k}^\top Q\bar{x}_{j|k} + u_{j|k}^\top Ru_{j|k} + \bar{x}_{N_p|k}^\top Q_f\bar{x}_{N_p|k},\end{aligned}$$

where $\bar{x}_{j|k}$ is the prediction at time step j obtained from the nominal system Σ^{nom} in Equation (3.1a)*.*

Proof. The proof directly follows from Equations (3.2a) and (3.3a), and the following properties of the expectation operator

$$\mathbb{E}\{x_0^\top P_1 x_0\} = x_0^\top P_1 x_0,\ \ \mathbb{E}\{x_0^\top P_2 w_{[0,N_p-1]|k}\} = 0,$$
$$\mathbb{E}\{w_{[0,N_p-1]|k}^\top P_3 w_{[0,N_p-1]|k}\} = Tr(P_3\bar{\Sigma}_w^{N_p}),\ \ \bar{\Sigma}_w^{N_P} := I_{N_p}\otimes\Sigma_w.$$

□

So the minimization of $J^M(\xi_k)$ in Equation (3.4) over $u_{j|k}$ is equivalent to the minimization of the cost function in Equation (3.6) in the sense that

$$\begin{aligned}u^*_{[0,N_p-1]|k} = &\quad \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^M(\xi_k),\\ \text{subject to} &\quad \bar{x}_{i+1|k} = A\bar{x}_{i|k} + Bu_{i|k},\\ &\quad \bar{x}_{0|k} = x_k, i\in\mathbb{Z}_{[0,N_p-1]},\end{aligned} \tag{3.7}$$

defines an optimal solution for both problems. Furthermore, this result indicates that the optimal control that solves the nominal-MPC problem, where uncertainties are not taken into account, equals the optimal solution that solves M-MPC. In particular the closed loop trajectories are equivalent for these cases, thus the system Σ^{add} controlled with a nominal-MPC is robust (performance-wise) with respect to expected cost of a system with additive disturbances.

We use of the cost function $\bar{J}^M(\xi_k)$ as the candidate Lyapunov function $V(x_k)$ and force a sufficient decrease of $V(x_k)$ by satisfying an algebraic condition. For this purpose we assume specific characteristics on the nature of uncertainty.

Assumption 3.1.1 *There exists a scalar $\delta_w\in\mathbb{R}$ such that, for all $k\in\mathbb{Z}_{\geq 0}$, $\mathbb{E}\{w_k^\top w_k\} = Tr(\Sigma_w)\leq\delta_w||\bar{x}_k||_2^2$.*

Theorem 3.1.1 *For a stochastic dynamical system Σ^{add} (Equation* (3.1b)*), if the M-MPC problem (Equation* (3.35)*) is feasible at x_0 and there exists; i) a stabilizing controller for the nominal system, K_f, such that $\rho(A+BK_f)<1$; ii) the terminal weight Q_f is the positive definite solution of the matrix equation*

$$\phi^\top Q_f\phi - Q_f + \left(Q + K_f^\top RK_f + \delta_w\bar{g}_w^M\right) = 0, \tag{3.8}$$

where $\phi = A+BK_f$ *and* $\bar{g}_w = \sum_{i=0}^{N_p} Tr(F^\top {A^i}^\top Q_i A^i F), Q_i = Q_f$, *for* $i = N_p$, *and* $Q_i = Q$ *for* $i \neq N_p, i \in \mathbb{Z}_{[0,N_p]}$ *then the closed loop system is MS-stable.*

Proof. We start with

$$\begin{aligned}\mathbb{E}_k\{\hat{V}(x_{k+1})\} - V(x_k) = &-(\|\bar{x}_{0|k}\|_Q^2 + \|u_{0|k}\|_R^2) + g_w^M \\ &+ \bar{x}_{N_p|k}^\top[-Q_f + \phi^\top Q_f \phi + Q + K_f^\top R K_f]\bar{x}_{N_p|k}\end{aligned}$$

where $\hat{V}(x_{k+1})$ is the cost function $\bar{J}^M(\xi_{k+1})$ evaluated with the suboptimal control $u_{[0,N_p-1]|k+1} = \begin{bmatrix} u^*_{[1,N_p-1|k]} & K_f \bar{x}_{N_p|k} \end{bmatrix}$, and $g_w^M = \sum_{i=0}^{N_p} Tr(F^\top {A^i}^\top Q_i A^i F \bar{\Sigma}_w)$. The right hand side of $\mathbb{E}_k\{\hat{V}(x_{k+1})\} - V(x_k)$ is negative if

$$\bar{x}_{N_p|k}^\top[-Q_f + \phi^\top Q_f \phi + Q + K_f^\top R K_f]\bar{x}_{N_p|k} + g_w^M = 0$$

This is a stability condition to check at each time instant $k \in \mathbb{Z}_{\geq 0}$. To transform this into an algebraic condition, δ_w in Assumption 3.1.1 is used to express the stability condition as in Equation (3.8) via

$$Tr(F^\top {A^i}^\top Q_i A^i F \Sigma_w) \leq \delta_w Tr(F^\top {A^i}^\top Q_i A^i F)||\bar{x}_k||_2^2, \forall k \in \mathbb{Z}_{\geq 0}.$$

This guarantees $\mathbb{E}_k\{\hat{V}(x_{k+1})\} - V(x_k) \leq 0$, thus the system is MS-stable. □

Remark 3.1.1 *The Assumption 3.1.1 is a conservative assumption on the nature of uncertainty, since as the state trajectory approaches to the origin, the effect of disturbance* w_k *vanishes. To incorporate less conservative disturbances, one should employ ISS techniques or use a sufficiently large terminal constraint set for which an MS-stabilizing controller is known a-priori, see Section III-C in [119] for further details.*

Mean-Variance MPC (MV-MPC)

In this case we reformulate the MPC problem as in Equation (3.3b) hence incorporating the second moment of the cost function. We consider the case without the terminal cost term, $Q_f = 0$, namely $J(\xi_k)$ consists of only running terms. The terms $\mathbb{E}\{J(\tilde{\xi}_k)\}$ and $\mathbb{D}\{J(\tilde{\xi}_k)\}$ in Equation (3.3b) are found as in Lemma 3.1.2.

Lemma 3.1.2 *For a stochastic dynamical system* Σ^{add} *(Equation* (3.1b)*), the MV-MPC cost function (Equation* (3.3b)*) is given by*

$$J^{MV}(\xi_k) = \quad \xi_k^\top H_{MV}^{N_p} \xi_k + f^M + \lambda_v f^V, \tag{3.9}$$

where f^M *is given in Equation* (3.5)*,*

$$\begin{aligned} f^V := &\quad 2Tr((T_F^{N_p^\top} \bar{Q} T_F^{N_p} \bar{\Sigma}_w)^2), \\ H_{MV}^{N_p} = &\quad \begin{bmatrix} \Theta_{N_p}^\top \bar{Q}^{MV} \Theta_{N_p} & \Theta_{N_p}^\top \bar{Q}^{MV} T_B^{N_p} \\ T_B^{N_p^\top} \bar{Q}^{MV} \Theta_{N_p} & T_B^{N_p^\top} \bar{Q}^{MV} T_B^{N_p} + \bar{R} \end{bmatrix}. \end{aligned} \tag{3.10}$$

Here $\bar{Q}^{MV} = \bar{Q} + 4\lambda_v \bar{Q}\bar{\Upsilon}\bar{Q}$ *and* $\bar{\Upsilon} = T_F^{N_p}\bar{\Sigma}_w^{N_p}T_F^{N_p^\top}$. *Furthermore,*

$$\begin{aligned}
\bar{J}^{MV}(\xi_k) = &\quad J^{MV}(\xi_k) - f^M - \lambda_v f^V, \\
\bar{J}^{MV}(\xi_k) = &\quad \xi_k^\top H_{MV}^{N_p}\xi_k, \\
\bar{J}^{MV}(\xi_k) = &\quad \sum_{j=0}^{N_p-1} \bar{x}_{j|k}^\top Q_j^{MV}\bar{x}_{j|k} + u_{j|k}^\top R u_{j|k} \\
&+2\sum_{l=1}^{N_p-2}\sum_{i=l+1}^{N_p-1} \bar{x}_{l|k}^\top Q_{l,i}^{MV}\bar{x}_{i|k},
\end{aligned} \tag{3.11}$$

where

$$\begin{aligned}
Q_j^{MV} &:= Q + 4\lambda_v Q\sum_{i=0}^{j-1} A^{j-1-i}F\Sigma_w F^\top A^{j-1-i^\top}Q, \\
Q_0^{MV} &:= Q,\ \ j \in \mathbb{Z}_{[1,N_p-1]}, \\
Q_{i,l}^{MV^\top} = Q_{l,i}^{MV} &:= 4\lambda_v Q\sum_{m=0}^{l-1} A^{l-1-m}F\Sigma_w F^\top A^{i-1-m^\top}Q, \\
Q_{0,m}^{MV} = Q_{m,0}^{MV} &:= 0_{n\times n},\ \ l \in \mathbb{Z}_{[1,N_p-2]}, i \in \mathbb{Z}_{[l+1,N_p-1]}
\end{aligned}$$

Proof. We start with

$$\begin{aligned}
&J(\tilde{\xi}_k) - \mathbb{E}\{J(\tilde{\xi}_k)\} = 2\Delta J_1(x_{0|k}) + 2\Delta J_2(u_{(\cdot)}) + \Delta J_3(w_{(\cdot)}) - \Delta J_4, \\
&\Delta J_1(x_{0|k}) = w_{(\cdot)}^\top T_F^{N_p^\top}\bar{Q}\Theta_{N_p}x_0,\ \ \Delta J_2(u_{(\cdot)}) = w_{(\cdot)}^\top T_F^{N_p^\top}\bar{Q}T_B^{N_p}u_{(\cdot)}, \\
&\Delta J_3(w_{(\cdot)}) = w_{(\cdot)}^\top T_F^{N_p^\top}\bar{Q}T_F^{N_p}w_{(\cdot)}\ \ \Delta J_4 = Tr(T_F^{N_p^\top}\bar{Q}T_F^{N_p}\bar{\Sigma}_w).
\end{aligned}$$

Then we use the fact that each term ΔJ_i, $i = 1,2,3,4$ are scalars, $\Delta J_i = \Delta J_i^\top \in \mathbb{R}$, to rearrange the expression to derive the MV-MPC cost function $J^{MV}(\xi_k)$, [276];

$$\begin{aligned}
\mathbb{D}\{J(\tilde{\xi}_k)\} := &\quad \mathbb{E}\{(J(\tilde{\xi}_k) - \mathbb{E}\{J(\tilde{\xi}_k)\})(J(\tilde{\xi}_k) - \mathbb{E}\{J(\tilde{\xi}_k)\})^\top\}, \\
= &\quad 4x_0^\top\Theta_{N_p}^\top\bar{Q}\bar{\Upsilon}\bar{Q}\Theta_{N_p}x_0 + 4x_0^\top\Theta_{N_p}^\top\bar{Q}\bar{\Upsilon}\bar{Q}T_B^{N_p}u_{(\cdot)} + \dots \\
&\quad 4u_{(\cdot)}^\top T_B^{N_p^\top}\bar{Q}\bar{\Upsilon}\bar{Q}\Theta_{N_p}x_0 + 4u_{(\cdot)}^\top T_B^{N_p^\top}\bar{Q}\bar{\Upsilon}\bar{Q}T_B^{N_p}u_{(\cdot)} + \dots \\
&\quad \mathbb{E}\{w_{(\cdot)}^\top T_F^{N_p^\top}\bar{Q}T_F^{N_p}w_{(\cdot)}w_{(\cdot)}^\top T_F^{N_p^\top}\bar{Q}T_F^{N_p}w_{(\cdot)}\} - \dots \\
&\quad Tr(T_F^{N_p^\top}\bar{Q}T_F^{N_p}\bar{\Sigma}_w)^2, \\
= &\quad \xi_k^\top H_V^{N_p}\xi_k + f^V
\end{aligned}$$

□

The MV-MPC results in a shifted cost $\bar{J}^{MV}(\xi_k)$ that only consists of predictions of nominal system trajectories. This has inherent advantages with regard to computational complexity, since it has equivalent computational requirements with the nominal-MPC problem. Moreover, the state weighting matrix 'Q-matrix' is changing over time, which increases over the prediction stages (the diagonal terms), i.e., $Q_j^{MV} \geq Q_{j-1}^{MV}$ and couples the future state predictions to each other via $Q_{l,i}^{MV}$. Furthermore, similar to the M-MPC case, the optimal

solution of unconstrained Mean-Variance MPC problem for system Σ^{add} is given as,

$$\begin{aligned} u^*_{[0,N_p-1]|k} = \quad & \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^{MV}(\xi_k), \\ \text{subject to} \quad & \bar{x}_{i+1|k} = A\bar{x}_{i|k} + Bu_{i|k}, \\ & \bar{x}_{0|k} = x_k, i \in \mathbb{Z}_{[0,N_p-1]}, \end{aligned} \tag{3.12}$$

is also the minimizer of the MV-MPC cost in Equation (3.3b).

We use deductions similar with the M-MPC to show the closed-loop stability. The following variables are used in the derivation of Theorem 3.1.2,

$$\begin{aligned} \mathbb{E}\{\hat{V}(x_{k+1})\} - V(x_k) = \quad & \xi_k M^{MV}(\lambda_v)\xi_k + g_w^{MV}, \\ \bar{M}^{MV}(\lambda_v, \delta_w) = \quad & M^{MV}(\lambda_v) + \begin{bmatrix} \delta_w \bar{g}_w^{MV} & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned} \tag{3.13}$$

while the candidate Lyapunov function $V(x_k)$ is taken as the shifted MV-MPC cost $\bar{J}^{MV}(x_k)$.

Theorem 3.1.2 *For a stochastic dynamical system Σ^{add} in Equation* (3.1b)*, if the MV-MPC problem is feasible for the initial condition x_0 and if (λ_v, δ_w) satisfy $\bar{M}^{MV}(\lambda_v, \delta_w) \leq 0$, then the controlled system is MS-stable.*

Proof. We express the Lyapunov function's decrease at time step k as

$$\begin{aligned} &\mathbb{E}_k\{\hat{V}(x_{k+1})\} - V(x_k) = -(||\bar{x}_{0|k}||_Q^2 + ||u_{0|k}||_R^2) + g_w^{MV} \\ &+\bar{x}_{N_p|k}^\top (Q_{N_p-1}^{MV} + K_f^\top R K_f)\bar{x}_{N_p|k} - \sum_{i=1}^{N_p-1} \bar{x}_{i|k}^\top Q_{ex}^{v,i} \bar{x}_{i|k} \\ &+ \sum_{j=1}^{N_p-2} \left(2\bar{x}_{j+1|k}^\top Q_{j,N_p-1}^{MV} \bar{x}_{N_p-1|k} - 2\bar{x}_{1|k}^\top Q_{1,j+1}^{MV} \bar{x}_{j+1|k} \right) \\ &- \sum_{j=2}^{N_p-2} \sum_{i=j+1}^{N_p-1} 2\bar{x}_{j|k}^\top Q_{ex}^{v,j,i} \bar{x}_{i|k}, \end{aligned}$$

where

$$\begin{aligned} Q_{ex}^{v,i} := \quad & 4\lambda_v Q A^{i-1} F \Sigma_w F^\top {A^{i-1}}^\top Q, \\ Q_{ex}^{v,j,i} := \quad & 4\lambda_v Q A^{j-1} F \Sigma_w F^\top {A^{i-1}}^\top Q, \\ g_w^{MV} := \quad & \sum_{j=0}^{N_p-1} Tr(F^\top {A^j}^\top Q_j^{MV} A^j F \Sigma_w) \\ & +2 \sum_{j=1}^{N_p-2} Tr(F^\top {A^j}^\top Q_{j,N_p-1}^{MV} A^{N_p-1} F \Sigma_w) \\ & +2 \sum_{j=1}^{N_p-3} \sum_{i=j+1}^{N_p-2} Tr(F^\top {A^j}^\top Q_{j,i}^{MV} A^i F \Sigma_w). \end{aligned}$$

One can make use of the Σ^{nom} dynamics to express

$$\mathbb{E}\{\hat{V}(x_{k+1})\} - V(x_k) = \xi_k^\top M^{MV}(\lambda_v)\xi_k + g_w^{MV},$$

where

$$M^{MV}(\lambda_v) = \begin{bmatrix} \Theta_{N_p+1}^\top \bar{Q}_\Delta^v \Theta_{N_p+1} & \Theta_{N_p+1}^\top \bar{Q}_\Delta^v T_B^{(\cdot)} \\ {T_B^{(\cdot)}}^\top \bar{Q}_\Delta^v \Theta_{N_p+1} & {T_B^{(\cdot)}}^\top \bar{Q}_\Delta^v T_B^{(\cdot)} + \bar{R}_\Delta \end{bmatrix},$$

in which $T_B^{(\cdot)} = T_B^{N_p+1,N_p}$,

$$R_\Delta := \begin{bmatrix} -R & 0 \\ 0 & 0 \end{bmatrix}, Q_{ex}^{v,N_p} := Q_{N_p-1}^{MV} + K_f^\top R K_f, \bar{N} = N_p - 1, \tag{3.14a}$$

$$\bar{Q}_\Delta^v := -\begin{bmatrix} Q & 0 & \dots & \dots & \dots & 0 & 0 \\ \star & Q_{ex}^{v,1} & Q_{1,2}^{MV} & \dots & \dots & Q_{1,\bar{N}}^{MV} & 0 \\ \vdots & \ddots & Q_{ex}^{v,2} & Q_{ex}^{v,2,3} & \dots & Q_{ex}^{v,2,\bar{N}} & -Q_{2,\bar{N}}^{MV} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \star & \ddots & Q_{ex}^{v,\bar{N}-1,\bar{N}} & -Q_{\bar{N}-2,\bar{N}}^{MV} \\ \vdots & \ddots & \ddots & \ddots & \ddots & Q_{ex}^{v,\bar{N}} & -Q_{\bar{N}-1,\bar{N}}^{MV} \\ \star & \dots & \dots & \dots & \dots & \star & -Q_{ex}^{v,N_p} \end{bmatrix}. \tag{3.14b}$$

Lastly, we use the bound δ_w in Assumption 3.1.1 to express the Lyapunov function condition as an LMI, i.e.,

$$\begin{aligned} \xi_k \bar{M}^{MV}(\lambda_v, \delta_w)\xi_k = \xi_k M^{MV}\xi_k + \delta_w \bar{x}_{0|k}^\top \bar{g}_w^{MV} \bar{x}_{0|k} \leq & \quad 0, \\ \bar{M}^{MV}(\lambda_v, \delta_w) = M^{MV}(\lambda_v) + \begin{bmatrix} \delta_w \bar{g}_w^{MV} & 0 \\ 0 & 0 \end{bmatrix} \leq & \quad 0, \end{aligned}$$

as in Equation (3.13). If $\bar{M}^{MV}(\lambda_v, \delta_w) \leq 0$ for a λ_v, given δ_w, then the running cost function for MV-MPC $\bar{J}^{MV}(x_k)$ is a Lyapunov function, hence MS-stability follows. □

Remark 3.1.2 *Due to the 'excess' terms $Q_{ex}^{v,j}$ and $Q_{ex}^{v,j,i}$, induced by the increasing state weight matrices Q_j^{MV} and $Q_{j,i}^{MV}$ over the prediction stages, the MV-MPC formulation does not require the terminal cost Q_f for the stability argument.*

Third Order Moment MPC

In this case we use the cost function as in Equation (3.3c), with $Q_f = 0$. The moment terms in Equation (3.3c) are found as follows.

Lemma 3.1.3 *For the stochastic dynamical system Σ^{add} in Equation* (3.1b)*, the MVS-MPC cost function (Equation* (3.3c)*) is equal to*

$$J^{MVS}(\xi_k) = \quad \xi_k^\top H_{MVS}^{N_p} \xi_k + f^M + \lambda_v f^V + \lambda_s f^S, \tag{3.15}$$

where

$$\begin{aligned} H_{MVS}^{N_p} = H_{MV}^{N_p} + H_S^{N_p}, \quad \xi_k = \begin{bmatrix} x_{0|k} \\ u_{[0,N_p-1]|k} \end{bmatrix}, \\ \bar{Q}^S := 2\bar{Q}\bar{\Upsilon}\bar{Q}\bar{\Upsilon}\bar{Q}, \quad \bar{\Upsilon} := T_F^{N_p}\bar{\Sigma}_w^{N_p}T_F^{N_p^\top}, \\ H_S^{N_p} = 12\lambda_s \begin{bmatrix} \Theta_{N_p}^\top \bar{Q}^S \Theta_{N_p} & \Theta_{N_p}^\top \bar{Q}^S T_B^{N_p} \\ T_B^{N_p^\top} \bar{Q}^S \Theta_{N_p} & T_B^{N_p^\top} \bar{Q}^S T_B^{N_p} \end{bmatrix}. \end{aligned} \tag{3.16}$$

Here, f^S is a constant term, defined by the 2^{nd}, 4^{th}, and 6^{th} order moments of $w_{[0,N_p-1]|k}$, i.e.,

$$f^S := \mathbb{E}\left\{ \left(\|w_{(\cdot)}\|^2_{T_F^{N_p^\top}\bar{Q}T_F^{N_p}} - Tr(T_F^{N_p^\top}\bar{Q}T_F^{N_p}) \right)^3 \right\},$$

and lastly $H_{MV}^{N_p}$ is given in Equation (3.10)*. Furthermore,*

$$\begin{aligned} \bar{J}^{MVS}(\xi_k) := & \quad J^{MVS}(x_k) - f^M - \lambda_v f^V - \lambda_s f^S, \\ = & \quad \xi_k^\top H_{MVS}^{N_p} \xi_k, \\ = & \quad \sum_{j=0}^{N_p-1} \bar{x}_{j|k}^\top Q_j^{MVS} \bar{x}_{j|k} + u_{j|k}^\top R u_{j|k} \\ & +2 \sum_{j=1}^{N_p-2} \sum_{i=j+1}^{N_p-1} \bar{x}_{j|k}^\top Q_{j,i}^{MVS} \bar{x}_{i|k}, \end{aligned} \tag{3.17}$$

where

$$\begin{aligned} Q_j^{MVS} = & \quad Q_j^{MV} + 12\lambda_s Q_j^S, \;\; Q_{l,i}^{MVS} = Q_{l,i}^{MV} + 12\lambda_s Q_{l,i}^S, \\ \bar{Q}^S = & \quad 2\bar{Q}\bar{\Upsilon}\bar{Q}\bar{\Upsilon}\bar{Q} =: 2 \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & Q_1^S & Q_{1,2}^S & \dots & Q_{1,N_p-1}^S \\ 0 & \star & Q_2^S & \dots & Q_{2,N_p-1}^S \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \star & \dots & \dots & Q_{N_p-1}^S \end{bmatrix}. \end{aligned}$$

Proof. Similar to the derivation given in Lemma 3.1.2 for MV-MPC case, we explicitly evaluate the third order central moment as,

$$\begin{aligned} \mathbb{S}\{J(x_k)\} = & \quad \mathbb{E}\{(J(x_k) - \mathbb{E}\{J(x_k)\})^3\}, \\ = & \quad \mathbb{E}\{(2\Delta J_1(x_0) + 2\Delta J_2(u) + \Delta J_3(w) - \Delta J_4)^3\}, \\ = & \quad 12\mathbb{E}\{x_0^\top \Theta_{N_p}^\top \bar{Q} T_F^{N_p} \beta_w^{N_p} T_F^{N_p^\top} \bar{Q}\Theta_{N_p} x_0 + \dots \\ & \quad x_0^\top \Theta_{N_p}^\top \bar{Q} T_F^{N_p} \beta_w^{N_p} T_F^{N_p^\top} \bar{Q} T_B^{N_p} u_{(\cdot)} + \dots \\ & \quad u_{(\cdot)}^\top T_B^{N_p^\top} \bar{Q} T_F^{N_p} \beta_w^{N_p} T_F^{N_p^\top} \bar{Q}\Theta_{N_p} x_0 + \dots \\ & \quad u_{(\cdot)}^\top T_B^{N_p^\top} \bar{Q} T_F^{N_p} \beta_w^{N_p} T_F^{N_p^\top} \bar{Q} T_B^{N_p} u_{(\cdot)}\} + f^S, \\ = & \quad \xi_k H_S^{N_p} \xi_k + f^S, \end{aligned}$$

where

$$\beta_w^{N_p} := w_{(\cdot)} w_{(\cdot)}^\top T_F^{N_p^\top} \bar{Q} T_F^{N_p} w_{(\cdot)} w_{(\cdot)}^\top - w_{(\cdot)} Tr(T_F^{N_p^\top} \bar{Q} T_F^{N_p} \bar{\Sigma}_w^{N_p}) w_{(\cdot)}^\top.$$

□

This cost is a function of only the nominal system, the nominal predictions of states $\bar{x}_{j|k}$ and the control actions deriving the system $u_{[0,N_p-1]|k}$. Then the MVS-MPC problem for system Σ^{add} is given as,

$$\begin{aligned} u^*_{[0,N_p-1]|k} = \quad & \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^{MVS}(\xi_k), \\ \text{subject to} \quad & \bar{x}_{i+1|k} = A\bar{x}_{i|k} + Bu_{i|k}, \\ & \bar{x}_{0|k} = x_k, i \in \mathbb{Z}_{[0,N_p-1]}. \end{aligned}$$

Furthermore, the stability of the Σ^{add} is established with similar deductions given for the MV-MPC case while incorporating Q_j^S and Q_{ji}^S terms within the derivation. Similar to the case in MV-MPC, we first provide several variables that are used in the derivation of the stability result

$$\begin{aligned} \mathbb{E}\{\hat{V}(x_{k+1})\} - V(x_k) = \quad & \xi_k M^{MVS}(\lambda_v, \lambda_s)\xi_k + g_w^{MVS}, \\ \bar{M}^{MVS}(\lambda_v, \lambda_s, \delta_w) = \quad & M^{MVS} + \begin{bmatrix} \delta_w \bar{g}_w^{MVS} & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned} \tag{3.18}$$

while the candidate Lyapunov function $V(x_k)$ is taken as the shifted MVS-MPC cost $\bar{J}^{MVS}(\xi_k)$.

Theorem 3.1.3 *For the stochastic dynamical system Σ^{add} in Equation* (3.1b)*, if the MVS-MPC problem is feasible for the initial condition x_0 and if $(\lambda_v, \lambda_s, \delta_w)$ triplet satisfy the condition $\bar{M}^{MVS}(\lambda_v, \lambda_s, \delta_w) \leq 0$, then the controlled system is MS-stable.*

Proof. The proof follows a similar reasoning with MV-MPC case (Theorem 3.1.2). First express at time k, $\mathbb{E}_k\{\hat{V}(x_{k+1})\} - V(x_k)$ and group the terms as

$$\begin{aligned} \mathbb{E}_k\{\hat{V}(x_{k+1})\} - V(x_k) = & -(||\bar{x}_{0|k}||_Q^2 + ||u_{0|k}||_R^2) + g_w^{MVS} \\ & + \bar{x}_{N_p|k}^\top (Q_{N_p-1}^{MVS} + K_f^\top R K_f) \bar{x}_{N_p|k} \\ & - \sum_{i=1}^{N_p-1} \bar{x}_{i|k}^\top Q_{ex}^{vs,i} \bar{x}_{i|k} - \sum_{j=2}^{N_p-2} \sum_{i=j+1}^{N_p-1} 2\bar{x}_{j|k}^\top Q_{ex}^{vs,j,i} \bar{x}_{i|k} \\ & + \sum_{j=1}^{N_p-2} \left(2\bar{x}_{j+1|k}^\top Q_{j,N_p-1}^{MVS} \bar{x}_{N_p|k} - 2\bar{x}_{1|k}^\top Q_{1,j+1}^{MVS} \bar{x}_{j+1|k} \right). \end{aligned}$$

where the terms $Q_{ex}^{vs,i}$ and $Q_{ex}^{vs,j,i}$ are found as

$$Q_{ex}^{vs,i} := Q_i^{MVS} - Q_{i-1}^{MVS}, \quad Q_{ex}^{vs,j,i} := Q_{j,i}^{MVS} - Q_{j-1,i-1}^{MVS},$$

and g_w^{MVS} contains the (trace) terms induced by $\mathbb{E}\{V_{k+1}\}$. Then use the Σ^{nom} dynamics to express $\mathbb{E}\{\hat{V}(x_{k+1})\} - V(x_k) = \xi_k^\top M^{MVS}(\lambda_v, \lambda_s)\xi_k + g_w^{MVS}$. The matrix $M^{MVS}(\lambda_v, \lambda_s)$

is defined similar to the $M^{MV}(\lambda_v)$ while the $\bar{Q}_\Delta^v$ term is replaced with $\bar{Q}_\Delta^{vs}$. Lastly, we use the bound δ_w in Assumption 3.1.1 to express the Lyapunov function condition as an LMI, i.e., $\bar{M}^{MVS}(\lambda_v, \lambda_s, \delta_w) \leq 0$, as in Equation (3.18). If $\bar{M}^{MVS}(\lambda_v, \lambda_s, \delta_w) \leq 0$ is satisfied for pair (λ_v, λ_s), given δ_w, then the running cost function for MVS-MPC $\bar{J}^{MVS}(\xi_k)$ is a Lyapunov function. □

3.1.2 Moment-based MPC for Tracking Problem

In this section, we construct the Mean, MV and MVS-MPC formulations for the unconstrained LTI system Σ^{add} for the output reference tracking problem. Here we consider two possible descriptions for linear systems. We define a nominal system, denoted with Σ^{nom} given in state-space representation as in Equation (3.19a) and a system with additive disturbances effecting the state and output equations, denoted with Σ^{add} as given in Equation (3.19b)

$$\Sigma^{nom} : \begin{cases} \bar{x}_{k+1} = & A\bar{x}_k + Bu_k, \\ \bar{y}_k = & C\bar{x}_k, \end{cases} \tag{3.19a}$$

$$\Sigma^{add} : \begin{cases} x_{k+1} = & Ax_k + Bu_k + Fw_k, \\ y_k = & Cx_k + v_k, \end{cases} \tag{3.19b}$$

where $x_k \in \mathbb{R}^n$, and $u_k \in \mathbb{R}^{n_u}$, and $y_k \in \mathbb{R}^{n_y}$ are the state, the control input (or the decision variable in the MPC problems) and the measured output at discrete time instant $k \in \mathbb{Z}_{\geq 0}$. The system Σ^{add} is subject to two type of uncertainties, denoted by w_k and v_k. These uncertainties are additive disturbances effecting the state and output equations, which are stochastic vector sequences with known pdfs, $f_{\tilde{w}}(w_k)$, for the disturbance w_k and $f_{\tilde{v}}(v_k)$, for the disturbance v_k.

Furthermore, the nature (or the characteristics) of uncertainties w_k and v_k are assumed to be independent and identically distributed (i.i.d.) random variables. We constrain the discussion to random variables that have finite moments to limit the technical discussion, as elaborated in next assumption.

Assumption 3.1.2 *The disturbance $w_k, k \in \mathbb{Z}_{\geq 0}$ in system Σ^{add} is distributed, for any k, with zero mean and time invariant covariance matrix $\Sigma_w \in \mathbb{R}^{n_w \times n_w}$, i.e.,*

$$\mathbb{E}\{w_k\} = 0_{n_w \times 1}, \qquad \Sigma_w := \mathbb{E}\{w_k w_k^\top\},$$

where $0_{n_w \times 1}$ is the zero vector of size n_w. Similarly, the disturbance v_k in system Σ^{add} is distributed with zero mean and time invariant covariance matrix $\Sigma_v \in \mathbb{R}^{n_v \times n_v}$. Furthermore, if the perturbations w_k or v_k are not Gaussian random variables, then

- *The covariance matrices are diagonal and time invariant;*

$$\begin{aligned} \Sigma_w = & \ \text{diag}(\sigma_{w_1}^2, \ldots, \sigma_{w_{n_w}}^2), \\ \Sigma_v = & \ \text{diag}(\sigma_{v_1}^2, \ldots, \sigma_{v_{n_v}}^2), \end{aligned}$$

where $\sigma_{w_i}^2$ is the variance of i^{th} component of random variable w_k and $\text{diag}(p), p \in R^{n_p}$ is the matrix of which the diagonal terms correspond to the elements of vector p.

- $\mathbb{E}\{[w_k]_i^4\}$ *and* $\mathbb{E}\{[v_k]_i^4\}$ *are finite, for each component* i *of random variables* w_k *or* v_k.

If the random variables w_k and v_k are distributed with Gaussian distribution function, then we denote it as $w_k \in \mathcal{N}(0, \Sigma_w)$ and $v_k \in \mathcal{N}(0, \Sigma_v)$.

Here we present the derivations for two different MPC formulations; case (i) stands for the cost function constructed directly using the input actions u_j, and in the case (ii) the cost function is constructed with the changes in the input action, i.e., $\Delta u_j := u_j - u_{j-1}$, to provide inherent integral action to the controller. For both of the cases we assume that a reference trajectory $r_k, k \in \mathbb{Z}_{\geq 0}$ is given which is desired to be tracked. The cost functions for case (i) and case (ii) are defined as[4]

$$J_{O,(i)}(\tilde{\xi}^O_{k,(i)}) = \sum_{j=0}^{N_p-1} (y_{j|k} - r_{j|k})^\top Q (y_{j|k} - r_{j|k}) + u_{j|k} R u_{j|k}, \tag{3.20a}$$

$$J_{O,(ii)}(\tilde{\xi}^O_{k,(ii)}) = \sum_{j=0}^{N_p-1} (y_{j|k} - r_{j|k})^\top Q (y_{j|k} - r_{j|k}) + \Delta u_{j|k} R \Delta u_{j|k}, \tag{3.20b}$$

where

$$\begin{aligned} \tilde{\xi}^{O^\top}_{k,(i)} &= \begin{bmatrix} x_{0|k}^\top & u_{[0,N_p-1]|k}^\top & w_{[0,N_p-1]|k}^\top & v_{[0,N_p-1]|k}^\top & r_{[0,N_p-1]|k}^\top \end{bmatrix}, \\ \tilde{\xi}^{O^\top}_{k,(i)} &= \begin{bmatrix} x_{0|k}^\top & u_{-1|k} & \Delta u_{[0,N_p-1]|k}^\top & w_{[0,N_p-1]|k}^\top & v_{[0,N_p-1]|k}^\top & r_{[0,N_p-1]|k}^\top \end{bmatrix}. \end{aligned} \tag{3.21}$$

The predicted output $y_{[0,N_p-1]|k}$ can be written as an algebraic function of $u_{[0,N_p-1]|k}$, $w_{[0,N_p-1]|k}$ and $v_{[0,N_p-1]|k}$, for case (i), as

$$y_{[0,N_p-1]|k} = \Theta_C x_0 + T_{B,C} u_{[0,N_p-1]|k} + T_{F,C} w_{[0,N_p-1]|k} + \bar{I}_{n_y}^{N_p} v_{[0,N_p-1]|k}, \tag{3.22a}$$

and as an algebraic function of input increments, apart from the disturbance variables, for the case (ii), as

$$\begin{aligned} y_{[0,N_p-1]|k} = \; & \Theta_C x_0 + T_{B,C}(1_{N_p} u_{-1|k} + T_\Delta \Delta u_{[0,N_p-1]|k}) + \\ & \dots \; T_{F,C} w_{[0,N_p-1]|k} + \bar{I}_{n_y}^{N_p} v_{[0,N_p-1]|k}, \end{aligned} \tag{3.22b}$$

in which the prediction matrices 1_{N_p} and T_Δ are defined as

$$1_{N_p} := \begin{bmatrix} I_{n_u} \\ I_{n_u} \\ \vdots \\ I_{n_u} \end{bmatrix}, \quad T_\Delta := \begin{bmatrix} I_{n_u} & 0 & \dots & 0 \\ I_{n_u} & I_{n_u} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ I_{n_u} & I_{n_u} & \dots & I_{n_u} \end{bmatrix}. \tag{3.22c}$$

With these definitions, the cost functions $J_{O,(i)}$ for case (i), and $J_{O,(ii)}$, for case (ii), are expressed as

$$\begin{aligned} J_{O,(i)}(\tilde{\xi}^O_{k,(i)}) &= \tilde{\xi}^{O^\top}_{k,(i)} H_{O,(i)} \tilde{\xi}^{O,(i)}_k, \\ J_{O,(ii)}(\tilde{\xi}^O_{k,(ii)}) &= \tilde{\xi}^{O^\top}_{k,(ii)} H_{O,(ii)} \tilde{\xi}^{O,(ii)}_k, \end{aligned} \tag{3.23}$$

[4]Due to its generality, we simplify the notation as $\xi^O_{k,(i)} = \xi_k$ wherever the control actions are decision variables and it is obvious that the outputs are used in the cost function.

where $H_{O,(i)}$ and $H_{O,(ii)}$ are constructed from $\Theta_C,\ T_{B,C},\ T_{F,C}$ with corresponding quadratic terms given in Equation (3.22c), i.e.,

$$H_{O,(i)} := \begin{bmatrix} \Theta_C^\top \bar{Q} \Theta_C & (\star) & (\star) & (\star) & (\star) \\ T_{B,C}^\top \bar{Q} \Theta_C & T_{B,C}^\top \bar{Q} T_{B,C} + \bar{R} & (\star) & (\star) & (\star) \\ T_{F,C}^\top \bar{Q} \Theta_C & T_{F,C}^\top \bar{Q} T_{B,C} & T_{F,C}^\top \bar{Q} T_{F,C} & (\star) & (\star) \\ \bar{I}_{n_y}^{N_p\top} \bar{Q} \Theta_C & \bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q} T_{F,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C} & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{F,C} & -\bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} & \bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} \end{bmatrix},$$

$$H_{O,(ii)} := \begin{bmatrix} \Theta_C^\top \bar{Q} \Theta_C & (\star) & (\star) & (\star) & (\star) & (\star) \\ T_{B,C,u_{-1}}^\top \bar{Q} \Theta_C & T_{B,C,u_{-1}}^\top \bar{Q} T_{B,C,u_{-1}} & (\star) & (\star) & (\star) & (\star) \\ T_{B,C,\Delta u}^\top \bar{Q} \Theta_C & T_{B,C,\Delta u}^\top \bar{Q} T_{B,C,u_{-1}} & T_{B,C,\Delta u}^\top \bar{Q} T_{B,C,\Delta u} + \bar{R} & (\star) & (\star) & (\star) \\ T_{F,C}^\top \bar{Q} \Theta_C & T_{F,C}^\top \bar{Q} T_{B,C,u_{-1}} & T_{F,C}^\top \bar{Q} T_{B,C,\Delta u} & T_{F,C}^\top \bar{Q} T_{F,C} & (\star) & (\star) \\ \bar{I}_{n_y}^{N_p\top} \bar{Q} \Theta_C & \bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C,u_{-1}} & \bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C,\Delta u} & \bar{I}_{n_y}^{N_p\top} \bar{Q} T_{F,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C,u_{-1}} & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C,\Delta u} & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{F,C} & -\bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} & \bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} \end{bmatrix},$$

while

$$T_{B,C,u_{-1}}^{N_p} = \ T_{B,C}^{N_p} 1_{N_p}, \qquad T_{B,C,\Delta u}^{N_p} = T_{B,C}^{N_p} T_{\Delta u}. \tag{3.24}$$

Next, the Mean, MV and MVS MPC formulations for the output tracking MPC configuration are constructed. Similar to the regulation case, first we define three cost functions from nominal system dynamics Σ^{nom} and then show the relations between these and the robust counterparts of $J_{O,(i)}$ and $J_{O,(ii)}$ constructed from the centralized moments, i.e.,

$$\bar{J}_{O,(i)}^{M}(\xi_{k,(i)}^{O}) = \ \xi_{k,(i)}^{O\top} H_{O,(i)}^{M} \xi_{k,(i)}^{O}, \tag{3.25a}$$

$$\bar{J}_{O,(i)}^{MV}(\xi_{k,(i)}^{O}) = \ \xi_{k,(i)}^{O\top} H_{O,(i)}^{MV} \xi_{k,(i)}^{O}, \tag{3.25b}$$

$$\bar{J}_{O,(i)}^{MVS}(\xi_{k,(i)}^{O}) = \ \xi_{k,(i)}^{O\top} H_{O,(i)}^{MVS} \xi_{k,(i)}^{O}, \tag{3.25c}$$

$$\bar{J}_{O,(ii)}^{M}(\xi_{k,(ii)}^{O}) = \ \xi_{k,(ii)}^{O\top} H_{O,(ii)}^{M} \xi_{k,(ii)}^{O}, \tag{3.25d}$$

$$\bar{J}_{O,(ii)}^{MV}(\xi_{k,(ii)}^{O}) = \ \xi_{k,(ii)}^{O\top} H_{O,(ii)}^{MV} \xi_{k,(ii)}^{O}, \tag{3.25e}$$

$$\bar{J}_{O,(ii)}^{MVS}(\xi_{k,(ii)}^{O}) = \ \xi_{k,(ii)}^{O\top} H_{O,(ii)}^{MVS} \xi_{k,(ii)}^{O}, \tag{3.25f}$$

where

$$\begin{aligned} \xi_{k,(i)}^{O\top} &= \begin{bmatrix} x_{0|k}^\top & u_{[0,N_p-1]|k}^\top & r_{[0,N_p-1]|k}^\top \end{bmatrix}, \\ \xi_{k,(ii)}^{O\top} &= \begin{bmatrix} x_{0|k}^\top & u_{-1|k}^\top & \Delta u_{[0,N_p-1]|k}^\top & r_{[0,N_p-1]|k}^\top \end{bmatrix}, \end{aligned} \tag{3.25g}$$

$$H_{O,(i)}^{M} = \begin{bmatrix} \Theta_C^\top \bar{Q} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q} \Theta_C & T_{B,C}^\top \bar{Q} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.25h}$$

$$H_{O,(i)}^{MV} = \begin{bmatrix} \Theta_C^\top \bar{Q}_O^{MV} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_O^{MV} \Theta_{N_p,C} & T_{B,C}^\top \bar{Q}_O^{MV} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MV} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MV} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MV} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.25i}$$

$$H_{O,(i)}^{MVS} = \begin{bmatrix} \Theta_C^\top \bar{Q}_O^{MVS} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_O^{MVS} \Theta_C & T_{B,C}^\top \bar{Q}_O^{MVS} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MVS} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MVS} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MVS} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \quad (3.25j)$$

$$H_{O,(ii)}^{M} = \begin{bmatrix} \Theta_C^\top \bar{Q} \Theta_C & (\star) & (\star) & (\star) \\ T_{B,C,u_{-1}}^\top \bar{Q} \Theta_C & T_{B,C,u_{-1}}^\top \bar{Q} T_{B,C,u_{-1}} & (\star) & (\star) \\ T_{B,C,\Delta u}^\top \bar{Q} \Theta_C & T_{B,C,\Delta u}^\top \bar{Q} T_{B,C,u_{-1}} & T_{B,C,\Delta u}^\top \bar{Q} T_{B,C,\Delta u} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C,u_{-1}} & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C,\Delta u} & \bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \quad (3.25k)$$

$$H_{O,(ii)}^{MV} = \begin{bmatrix} \Theta_C^\top \bar{Q}_O^{MV} \Theta_C & (\star) & (\star) & (\star) \\ T_{B,C,u_{-1}}^\top \bar{Q}_O^{MV} \Theta_C & T_{B,C,u_{-1}}^\top \bar{Q}_O^{MV} T_{B,C,u_{-1}} & (\star) & (\star) \\ T_{B,C,\Delta u}^\top \bar{Q}_O^{MV} \Theta_C & T_{B,C,\Delta u}^\top \bar{Q}_O^{MV} T_{B,C,u_{-1}} & T_{B,C,\Delta u}^\top \bar{Q}_O^{MV} T_{B,C,\Delta u} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MV} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MV} T_{B,C,u_{-1}} & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MV} T_{B,C,\Delta u} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MV} \bar{I}_{n_y}^{N_p\top} \end{bmatrix}, \quad (3.25l)$$

$$H_{O,(ii)}^{MVS} = \begin{bmatrix} \Theta_C^\top \bar{Q}_O^{MVS} \Theta_C & (\star) & (\star) & (\star) \\ T_{B,C,u_{-1}}^\top \bar{Q}_O^{MVS} \Theta_C & T_{B,C,u_{-1}}^\top \bar{Q}_O^{MVS} T_{B,C,u_{-1}} & (\star) & (\star) \\ T_{B,C,\Delta u}^\top \bar{Q}_O^{MVS} \Theta_C & T_{B,C,\Delta u}^\top \bar{Q}_O^{MVS} T_{B,C,u_{-1}} & T_{B,C,\Delta u}^\top \bar{Q}_O^{MVS} T_{B,C,\Delta u} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MVS} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MVS} T_{B,C,u_{-1}} & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MVS} T_{B,C,\Delta u} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_O^{MVS} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \quad (3.25m)$$

$$\begin{aligned} \bar{Q}_O^{MV} &= \bar{Q} + 4\lambda_v \bar{Q} \bar{\Upsilon}_O \bar{Q}, \\ \bar{Q}_O^{MVS} &= \bar{Q}_O^{MV} + 24\lambda_s \bar{Q} \bar{\Upsilon}_O \bar{Q} \bar{\Upsilon}_O \bar{Q}, \\ \bar{\Upsilon}_O &= T_{F,C} \bar{\Sigma}_w^{N_p} T_{F,C}^\top + \bar{\Sigma}_v^{N_p}, \\ \bar{\Sigma}_w^{N_p} &= I_{N_p} \otimes \Sigma_w, \\ \bar{\Sigma}_v^{N_p} &= I_{N_p} \otimes \Sigma_v, \end{aligned} \quad (3.25n)$$

Theorem 3.1.4 *Consider an LTI stochastic dynamical system Σ^{add} as in Equation* (3.19b) *subject to Assumption 3.1.2 and a reference trajectory $r_k, k \in \mathbb{Z}_{\geq 0}$;*

- *The cost functions for Mean MPC problem with the stochastic cost functions $J_{O,(i)}$ and $J_{O,(ii)}$ defined in Equation* (3.20) *satisfy*

$$\begin{aligned} J_{O,(i)}^M(\xi_{k,(i)}^O) &= \mathbb{E}\{J_{O,(i)}(\tilde{\xi}_{O,(i)})\}, \\ &= \xi_{k,(i)}^{O\top} H_{O,(i)}^M \xi_{k,(i)}^O + f_{O,(i)}^M, \\ J_{O,(ii)}^M(\xi_{k,(ii)}^O) &= \mathbb{E}\{J_{O,(ii)}(\tilde{\xi}_{O,(ii)})\}, \\ &= \xi_{k,(ii)}^{O\top} H_{O,(ii)}^M \xi_{k,(ii)}^O + f_{O,(ii)}^M, \end{aligned} \quad (3.26a)$$

Furthermore both $f_{O,(i)}^M$ and $f_{O,(ii)}^M$ terms are constant and

$$\begin{aligned} \bar{J}_{O,(i)}^M(\xi_{k,(i)}^O) &= J_{O,(i)}^M(\xi_{k,(i)}^O) - f_{O,(i)}^M, \\ \bar{J}_{O,(ii)}^M(\xi_{k,(ii)}^O) &= J_{O,(ii)}^M(\xi_{k,(ii)}^O) - f_{O,(ii)}^M, \end{aligned} \quad (3.26b)$$

and in particular

$$\begin{aligned}
\min_{u_{[0,N_p-1]|k}} J^M_{O,(i)}(\xi^O_{k,(i)}) &= \min_{u_{[0,N_p-1]|k}} \bar{J}^M_{O,(i)}(\xi^O_{k,(i)}) + f^M_{O,(i)},\\
\arg\min_{u_{[0,N_p-1]|k}} J^M_{O,(i)}(\xi^O_{k,(i)}) &= \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^M_{O,(i)}(\xi^O_{k,(i)}),\\
\min_{u_{[0,N_p-1]|k}} J^M_{O,(ii)}(\xi^O_{k,(ii)}) &= \min_{u_{[0,N_p-1]|k}} \bar{J}^M_{O,(ii)}(\xi^O_{k,(ii)}) + f^M_{O,(ii)},\\
\arg\min_{u_{[0,N_p-1]|k}} J^M_{O,(ii)}(\xi^O_{k,(ii)}) &= \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^M_{O,(ii)}(\xi^O_{k,(ii)}),
\end{aligned} \tag{3.26c}$$

- *The cost functions for MV MPC problem with the stochastic cost functions $J_{O,(i)}$ and $J_{O,(ii)}$ in Equation* (3.20) *are given by*

$$\begin{aligned}
J^{MV}_{O,(i)}(\xi^O_{k,(i)}) &= \mathbb{E}\{J_{O,(i)}(\tilde{\xi}_{O,(i)})\} + \lambda_v \mathbb{D}\{J_{O,(i)}(\tilde{\xi}_{O,(i)})\},\\
&= \xi^{O^\top}_{k,(i)} H^{MV}_{O,(i)} \xi^O_{k,(i)} + f^M_{O,(i)} + \lambda_v f^V_{O,(i)},\\
J^{MV}_{O,(ii)}(\xi^O_{k,(ii)}) &= \mathbb{E}\{J_{O,(ii)}(\tilde{\xi}_{O,(ii)})\} + \lambda_v \mathbb{D}\{J_{O,(ii)}(\tilde{\xi}_{O,(ii)})\},\\
&= \xi^{O^\top}_{k,(ii)} H^{MV}_{O,(ii)} \xi^O_{k,(ii)} + f^M_{O,(ii)} + \lambda_v f^V_{O,(ii)},
\end{aligned} \tag{3.27a}$$

Furthermore,

$$\begin{aligned}
\bar{J}^{MV}_{O,(i)}(\xi^O_{k,(i)}) &= J^{MV}_{O,(i)}(\xi^O_{k,(i)}) - f^M_{O,(i)} - \lambda_v f^V_{O,(i)},\\
\bar{J}^{MV}_{O,(ii)}(\xi^O_{k,(ii)}) &= J^{MV}_{O,(ii)}(\xi^O_{k,(ii)}) - f^M_{O,(ii)} - \lambda_v f^V_{O,(ii)},
\end{aligned} \tag{3.27b}$$

and in particular

$$\begin{aligned}
\min_{u_{[0,N_p-1]|k}} J^{MV}_{O,(i)}(\xi^O_{k,(i)}) &= \min_{u_{[0,N_p-1]|k}} \bar{J}^{MV}_{O,(i)}(\xi^O_{k,(i)}) + f^M_{O,(i)} + \lambda_v f^V_{O,(i)},\\
\arg\min_{u_{[0,N_p-1]|k}} J^{MV}_{O,(i)}(\xi^O_{k,(i)}) &= \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^{MV}_{O,(i)}(\xi^O_{k,(i)}),\\
\min_{u_{[0,N_p-1]|k}} J^{MV}_{O,(ii)}(\xi^O_{k,(ii)}) &= \min_{u_{[0,N_p-1]|k}} \bar{J}^{MV}_{O,(ii)}(\xi^O_{k,(ii)}) + f^M_{O,(ii)} + \lambda_v f^V_{O,(ii)},\\
\arg\min_{u_{[0,N_p-1]|k}} J^{MV}_{O,(ii)}(\xi^O_{k,(ii)}) &= \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^{MV}_{O,(ii)}(\xi^O_{k,(ii)}).
\end{aligned} \tag{3.27c}$$

- *The cost functions for MVS MPC problem with the stochastic cost functions $J_{O,(i)}$ and $J_{O,(ii)}$ in Equation* (3.20) *are given by*

$$\begin{aligned}
J^{MVS}_{O,(i)}(\xi^O_{k,(i)}) &= \mathbb{E}\{J_{O,(i)}(\tilde{\xi}_{O,(i)})\} + \lambda_v \mathbb{D}\{J_{O,(i)}(\tilde{\xi}_{O,(i)})\} + \lambda_s \mathbb{S}\{J_{O,(i)}(\tilde{\xi}_{O,(i)})\},\\
&= \xi^{O^\top}_{k,(i)} H^{MVS}_{O,(i)} \xi^O_{k,(i)} + f^M_{O,(i)} + \lambda_v f^V_{O,(i)} + \lambda_s f^S_{O,(i)},\\
J^{MVS}_{O,(ii)}(\xi^O_{k,(ii)}) &= \mathbb{E}\{J_{O,(ii)}(\tilde{\xi}_{O,(ii)})\} + \lambda_v \mathbb{D}\{J_{O,(ii)}(\tilde{\xi}_{O,(ii)})\} + \lambda_s \mathbb{S}\{J_{O,(ii)}(\tilde{\xi}_{O,(ii)})\},\\
&= \xi^{O^\top}_{k,(ii)} H^{MVS}_{O,(ii)} \xi^O_{k,(ii)} + f^M_{O,(ii)} + \lambda_v f^V_{O,(ii)} + \lambda_s f^S_{O,(ii)},
\end{aligned} \tag{3.28a}$$

Furthermore,

$$\begin{aligned}
\bar{J}^{MVS}_{O,(i)}(\xi^O_{k,(i)}) &= J^{MVS}_{O,(i)}(\xi^O_{k,(i)}) - f^M_{O,(i)} - \lambda_v f^V_{O,(i)} - \lambda_s f^S_{O,(i)},\\
\bar{J}^{MVS}_{O,(ii)}(\xi^O_{k,(ii)}) &= J^{MVS}_{O,(ii)}(\xi^O_{k,(ii)}) - f^M_{O,(ii)} - \lambda_v f^V_{O,(ii)} - \lambda_s f^S_{O,(ii)},
\end{aligned} \tag{3.28b}$$

and in particular

$$\begin{aligned}
\min_{u_{[0,N_p-1]|k}} J^{MVS}_{O,(i)}(\xi^O_{k,(i)}) = &\quad \min_{u_{[0,N_p-1]|k}} \bar{J}^{MVS}_{O,(i)}(\xi^O_{k,(i)}) + f^M_{O,(i)} + \lambda_v f^V_{O,(i)} + \lambda_s f^S_{O,(i)},\\
\arg\min_{u_{[0,N_p-1]|k}} J^{MVS}_{O,(i)}(\xi^O_{k,(i)}) = &\quad \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^{MVS}_{O,(i)}(\xi^O_{k,(i)}),\\
\min_{u_{[0,N_p-1]|k}} J^{MVS}_{O,(ii)}(\xi^O_{k,(ii)}) = &\quad \min_{u_{[0,N_p-1]|k}} \bar{J}^{MVS}_{O,(ii)}(\xi^O_{k,(ii)}) + f^M_{O,(ii)} + \lambda_v f^V_{O,(ii)} + \lambda_s f^S_{O,(ii)},\\
\arg\min_{u_{[0,N_p-1]|k}} J^{MVS}_{O,(ii)}(\xi^O_{k,(ii)}) = &\quad \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^{MVS}_{O,(ii)}(\xi^O_{k,(ii)}).
\end{aligned} \tag{3.28c}$$

Proof. The proof follows exactly same steps with the proofs of Lemmas 3.1.1-3.1.2-3.1.3, hence we refer the reader to the proofs in [317]. □

We stress the fact that the minimization of $J^M_{O,(i)}(\xi^O_{k,(i)})$ or $J^M_{O,(ii)}(\xi^O_{k,(ii)})$ over $u_{j|k}$ for mean MPC formulations[5] is equivalent to the minimization of the corresponding cost functions in Equation (3.25) in the sense that

$$\mathcal{P}^M_{O,(i)}: \begin{cases} u^{*,(i)}_{[0,N_p-1]|k} = & \arg\min_{u_{[0,N_p-1]|k}} \bar{J}^M_{O,(i)}(\xi^O_{k,(i)}),\\ \text{subject to} & \bar{x}_{i+1|k} = A\bar{x}_{i|k} + Bu_{i|k},\\ & \bar{y}_{i|k} = Cx_{i|k}, \bar{x}_{0|k} = x_k, i \in \mathbb{Z}_{[0,N_p-1]}. \end{cases} \tag{3.29a}$$

$$\mathcal{P}^M_{O,(ii)}: \begin{cases} u^{*,(ii)}_{[0,N_p-1]|k} = & \arg\min_{\Delta u_{[0,N_p-1]|k}} \bar{J}^M_{O,(ii)}(\xi^O_{k,(ii)}),\\ \text{subject to} & \bar{x}_{i+1|k} = A\bar{x}_{i|k} + Bu_{i|k},\\ & \bar{y}_{i|k} = Cx_{i|k}, \bar{x}_{0|k} = x_k, i \in \mathbb{Z}_{[0,N_p-1]},\\ & \Delta u_j = u_j - u_{j-1}. \end{cases} \tag{3.29b}$$

define the optimal solutions for the output tracking Mean MPC problems.

Remark 3.1.3 *In many cases, processes (with nonlinear behaviour) are operated in various operating conditions or over an operating trajectory. Once a process is controlled around multiple operating points (or equivalently around an trajectory) we observe, in general, a time varying linear dynamical behaviour in the deviation variables as an approximation of the full dynamics. If we assume access to nominally known time-varying linear dynamics, say in LTV (A_k, B_k, C_k, D_k) or in LPV ($A(p), B(p), C(p), D(p)$ setting, then the moment-based MPC formulations can be extended by incorporating time varying effects into the definitions of Θ_C and $T_{B,C}$.*

Moment-based MPC with Affine Disturbance Feedback Control Laws

A crucial observation in [138] is the incorporation of future uncertain perturbations $w_{j|k}$ into the control law formulation for $u_{j+i|k}, i \geq 1$. If one parametrizes the control law as

$$u_{j|k} = \quad h_{j|k} + \sum_{i=0}^{j-1} K_{i|k,j} w_{i|k}, \tag{3.30}$$

[5]The MV or MVS MPC cases, *mutadis mutandis*, follow with the same reasoning.

where

$$K_k = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ K_{0|k,1} & 0 & \dots & 0 & 0 \\ K_{0|k,2} & K_{1|k,2} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ K_{0|k,N_p-1} & K_{1|k,N_p-1} & \dots & K_{N_p-2|k,N_p-1} & 0 \end{bmatrix}, \tag{3.31}$$

then the future realizations of uncertainty are *causally* incorporated into the future control actions to correct the state (equivalently output) predictions. Then the control law formulation given in Equation (3.30) leads to a convex optimization problem if one considers the cost function $J^M_{O,(i)}$ in Equation (3.45a), see [138]. Hence, one might improve the closed-loop performance by incorporating the precompensator K_k while taking advantage of the convexity properties of the resulting MPC problem. In the following we extend the moment-based MPC formulations for affine disturbance feedback control parametrization case.

First, we evaluate the mean of the cost function given in Equation (3.20a) with the control actions are parameterized as in Equation (3.30). The explicit form of the resulting cost function, denoted with $J^M_{O,adf}(\xi^O_k)$, is summarized in Lemma 3.1.4.

Lemma 3.1.4 *Consider an LTI stochastic dynamical system Σ^{add} as in Equation* (3.19b) *subject to Assumption 3.1.2, a reference trajectory $r_k,\ k \in \mathbb{Z}_{\geq 0}$ and the cost function $J_{O,(i)}(\tilde{\xi}^O_k)$ in Equation* (3.20a)*. If the control actions in Mean MPC problem are paramterized as in Equation* (3.30)*, then the cost function of Mean-MPC problem is given as*

$$J^M_{O,adf}(\xi^O_k, K_k) = {\xi^O_k}^\top H^M_{adf} \xi^O_k + Tr\left(M^\top_{k,adf} \bar{Q} M_{k,adf} \bar{\Sigma}^{N_p}_w + \bar{Q}\bar{\Sigma}^{N_p}_v\right), \tag{3.32}$$

where

$$\begin{aligned} H^M_{adf} &:= \begin{bmatrix} \Theta_C^\top \bar{Q} \Theta_C & (\star) & (\star) \\ T^\top_{B,C} \bar{Q} \Theta_C & T^\top_{B,C} \bar{Q} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}^{N_p^\top}_{n_y} \bar{Q} \Theta_C & -\bar{I}^{N_p^\top}_{n_y} \bar{Q} T_{B,C} & \bar{I}^{N_p^\top}_{n_y} \bar{Q} \bar{I}^{N_p}_{n_y} \end{bmatrix} \\ M_{k,adf} &:= T_{B,C} K_k + T_{F,C}, \\ \xi^{O^\top}_{k,adf} &= \begin{bmatrix} x^\top_{0|k} & h^\top_{[0,N_p-1]|k} & r^\top_{[0,N_p-1]|k} \end{bmatrix}. \end{aligned} \tag{3.33}$$

Furthermore, by using the linearity properties of the trace operator,

$$\begin{aligned} \bar{J}^M_{O,adf}(\xi^O_{k,(i)}, K_k) = &\sum_{j=0}^{N_p-1} \bar{x}^\top_{j|k} Q \bar{x}_{j|k} + u^\top_{j|k} R u_{j|k} + \\ &\sum_{j=1}^{N_p-1} \sum_{i=j}^{N_p-1} Tr\left(M^\top_{i|k,j-1} Q M_{i|k,j-1} \Sigma_{w_{j-1|k}}\right), \\ \bar{J}^M_{O,adf}(\xi^O_{k,(i)}, K_k) = &\ J^M_{O,adf}(\xi^O_{k,(i)}, K_k) - Tr(\bar{Q}\bar{\Sigma}^{N_p}_v), \end{aligned} \tag{3.34}$$

where $\bar{x}_{j|k}$ is the prediction at time step j obtained from the nominal system Σ^{nom} in Equation (3.19a) *with the input actions $h_{[0,j-1]|k}$.*

Proof. First we express the evolution of Σ^{add} dynamics with the affine disturbance feedback parametrization of the input signal $u_{j|k} = h_{j|k} + \sum_{i=0}^{j-1} K_{i|k,j} w_{i|k}$, i.e.,

$$\Sigma_{adf}^{add} : \begin{cases} x_{j+1|k} = & Ax_{j|k} + Bh_{j|k} + Fw_{j|k} + B\sum_{i=0}^{j-1} K_{i|k,j} w_{i|k}, \\ y_{j|k} = & Cx_k + v_k. \end{cases}$$

Then the output predictions can be expressed as

$$y_{[0,N_p-1]|k} = \Theta_C x_{0|k} + T_{B,C} h_{[0,N_p-1]|k} + (T_{B,C}K_k + T_{F,C}) w_{[0,N_p-1]|k} + \bar{I}_{n_y}^{N_p\top} v_{[0,N_p-1]|k}.$$

Using the last expression in the cost function $J_{O,(i)}(\tilde{\xi}_{k,(i)}^{O}$ leads to a function similar to the cost function $J_{O,(i)}^{M}(\xi_{k,(i)}^{O})$ in Theorem 3.1.4, with the exception of term f^M. In the case of affine disturbance feedback parameterization, the f^M term contains the decision variables K_k in its definition, i.e.,

$$f^M = Tr\left((T_{B,C}^{N_p} K_k + T_{F,C}^{N_p})^\top \bar{Q} (T_{B,C}^{N_p} K_k + T_{F,C}^{N_p}) \bar{\Sigma}_w^{N_p} + \bar{Q} \bar{\Sigma}_v^{N_p} \right).$$

Then the proof follows by straightforward algebraic operations for calculating the cost function $\bar{J}_{O,adf}^{M}(\xi_{k,(i)}^{O}, K_k)$. □

So the minimization of $J_{O,adf}^{M}(\xi_{k,(i)}^{O}, K_k)$ in Equation (3.32) over $(h_{[0,N_p-1]|k}, K_k)$ is equivalent to the minimization of the cost function $\bar{J}_{O,adf}^{M}(\xi_{k,(i)}^{O}, K_k)$ in Equation (3.34) in the sense that

$$\mathcal{P}_{O,adf}^{M} : \begin{cases} (h_{[0,N_p-1]|k}^*, K_k^*) = & \arg \min_{h_{[0,N_p-1]|k}, K_k} \bar{J}_{O,adf}^{M}(k), \\ \text{subject to} & \bar{x}_{i+1|k} = A\bar{x}_{i|k} + Bh_{i|k}, \\ & \bar{y}_{i|k} = Cx_{i|k}, \bar{x}_{0|k} = x_k, i \in \mathbb{Z}_{[0,N_p-1]}, \end{cases} \tag{3.35}$$

defines an optimal solution $u_{[0,N_p-1]|k}^* = h_{[0,N_p-1]|k}^* + K_k^* w_{[0,N_p-1]|k}$ for the mean MPC problem with affine disturbance feedback structure.

Remark 3.1.4 *The moment-based MPC formulations for higher order moment cases, such as MV or MVS MPC, yield cost functions that are consisting of fourth (or more) order terms of the decision variables. This increases the complexity of the optimization problem tremendously and thus we do not report these cases here.*

3.1.3 Possible Extensions and Further Discussion on the Moment-based MPC

4th order Moment Case

In this section we consider the fourth order moment case, constructed as the linear combination of first four (centralized) moments. For this purpose, the cost function $J_{O,(i)}(\tilde{\xi}_{k,(i)}^{O})$ given in Equation (3.23) is manipulated[6].

$$J_{O,(i)}^{MVSK}(\xi_{k,(i)}^{O}) = J_{O,(i)}^{MVS}(\xi_{k,(i)}^{O}) + \lambda_K \mathbb{K}\{J_{O,(i)}(\tilde{\xi}_{k,(i)}^{O})\},$$

[6]For simplicity we assume that there is no output noise in the dynamics Σ^{add}, i.e., $v_k = 0,\ k \in \mathbb{Z}_{\geq 0}$.

where $\mathbb{K}\{J_{O,(i)}(\tilde{\xi}^O_{k,(i)})\} = \mathbb{E}\{(J_{O,(i)}(\tilde{\xi}^O_{k,(i)}) - \mathbb{E}\{J_{O,(i)}(\tilde{\xi}^O_{k,(i)})\})^4\}$. In this case the computational complexity of the resulting optimization problem has undesired properties. To demonstrate this fact, we explicitly express the fourth order moment MPC. Since we already have the explicit expressions for the first three terms derived in Theorem 3.1.4, only $\mathbb{K}\{J_{O,(i)}(\tilde{\xi}^O_{k,(i)})\}$ term is treated here. We express the term $\mathbb{K}\{J_{O,(i)}(\tilde{\xi}^O_{k,(i)})\}$ as

$$\begin{aligned}\mathbb{K}\{J_{O,(i)}(\tilde{\xi}^O_{k,(i)})\} =& \mathbb{E}\{(J_{O,(i)}(\tilde{\xi}^O_{k,(i)}) - \mathbb{E}\{J_{O,(i)}(\tilde{\xi}^O_{k,(i)})\})^4\},\\ =& \mathbb{E}\{(2\Delta J_1(x_0) + 2\Delta J_2(u) + \Delta J_3(w) - \Delta J_4)^4\},\\ =& {\xi^O_{k,(i)}}^\top H_K(x_k, u_{[0,N_p-1]|k})\xi^O_{k,(i)}\end{aligned} \tag{3.36a}$$

where

$$H_K(x,u) = \begin{bmatrix} \Theta_C^\top \bar{Q}_K(x,u)\Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_K(x,u)\Theta_C & T_{B,C}^\top \bar{Q}_K(x,u) T_{B,C} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_K(x,u)\Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_K(x,u) T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_K(x,u) \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.36b}$$

and

$$\begin{aligned} \bar{Q}_K(x,u) =& \ \mathbb{E}\left\{ \bar{Q} T_F w_{[0,N_p-1]|k} h_K(x,u,w)\, w_{[0,N_p-1]|k}^\top T_F^\top \bar{Q} \right\},\\ \Delta J_1(x_0) =& \ x_0^\top \Theta_C^\top \bar{Q} T_{F,C} w_{[0,N_p-1]|k},\\ \Delta J_2(u) =& \ u_{[0,N_p-1]|k}^\top T_{B,C}^\top \bar{Q} T_{F,C} w_{[0,N_p-1]|k},\\ \Delta J_3(w) =& \ w_{[0,N_p-1]|k}^\top T_{F,C}^\top \bar{Q} T_{F,C} w_{[0,N_p-1]|k},\\ \Delta J_4 =& \ Tr\left(T_{F,C}^\top \bar{Q} T_{F,C} \bar{\Sigma}_w\right). \end{aligned} \tag{3.36c}$$

In Equation (3.36c), the term $h_K(x,u,w)$ is given by

$$\begin{aligned} h_K(x,u,w) :=& \ 16\left(x_{0|k}^\top \Theta_C^\top \bar{Q} T_{F,C} w_{[0,N_p-1]|k} + u_{[0,N_p-1]|k}^\top T_{B,C}^\top \bar{Q} T_{F,C} w_{[0,N_p-1]|k}\right)^2 + \dots\\ & 24\left(w_{[0,N_p-1]|k}^\top T_{F,C}^\top \bar{Q} T_{F,C} w_{[0,N_p-1]|k} - Tr(T_{F,C}^\top \bar{Q} T_{F,C} \bar{\Sigma}_w^{N_p})\right)^2 + \dots\\ & 8\left(w_{[0,N_p-1]|k}^\top T_{F,C}^\top \bar{Q} T_{F,C} w_{[0,N_p-1]|k} Tr(T_{F,C}^\top \bar{Q} T_{F,C} \bar{\Sigma}_w^{N_p})\right). \end{aligned} \tag{3.36d}$$

Since the term $h_K(x,u,w)$ contains the second order forms depending on the state x_k and the input actions $u_{[0,N_p-1]|k}$, the cost function of 4^{th}-order moment MPC is resulting in a nonlinear matrix $H_K(x,u)$. Then the expression $\mathbb{K}\left\{J_{O,(i)}\left(\tilde{\xi}^O_{k,(i)}\right)\right\}$ contains fourth order terms of the optimization variables. This leads to an undesired increase in the complexity of the resulting optimization problem. Further research is required to express high order centralized moments in convex and compact form.

Observations on the Closed-loop Performance of Moment-based MPC

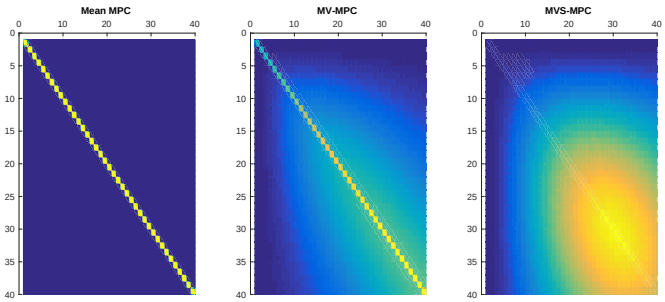
Here, we do not present stability results for the reference tracking problem via moment-based MPC. The stability problem for state regulation purposes is tackled in previous section, or in [317], which might be extended for the case of output reference tracking case by using the the mean-square stability (MS-stability) concept ([191]) and guaranteeing observability property via dynamics and the output weighting matrix.

One important observation on the moment-based MPC formulation is that for linear dynamics with additive disturbances, i.e., Σ^{add}, the mean, MV or MVS MPC problems can be expressed as deterministic optimization problems that are constrained to the nominal system dynamics, i.e., Σ^{nom}. This means that the resulting optimization problem is highly structured and can be solved with the same computational requirements of the nominal MPC problem. This structure has momentous advantages in comparison to the other robust MPC formulations by discarding the explicit maximization step over uncertainties or by elevating the problem from optimization over quantiles of the cost functions. The scenario based MPC formulations have similar computational complexity properties with moment-based MPC, yet these MPC methods are also prone to performance deterioration due to the possibility of extreme, yet highly unlikely, elements within the uncertainty scenarios set. In this line of reasoning, moment-based MPC is expected to reflect the aggregated, and on average true, effect of uncertainties within the dynamical evolution.

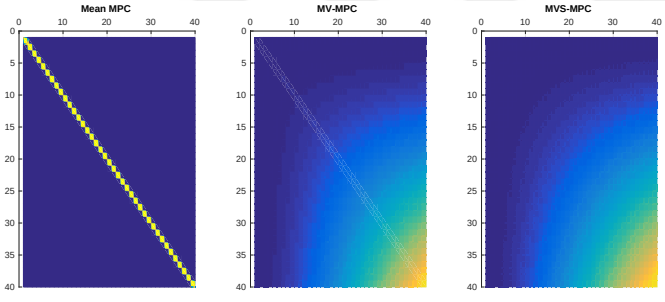
One further observation is that within the resulting moment-based MPC problems, the input weighting matrix R does not change over prediction stages j, $j \in \mathbb{Z}_{[0,N_p-1]}$ for mean, MV or MVS MPC formulations. On the other hand, the Q matrix, weighting the outputs (equivalently the states), accumulates the effects of predicted disturbances over the stages $j \geq 1$. At every prediction instant symmetric terms corresponding to the variance of disturbances upto that prediction stage is added to the Q_j^{MV}, or equivalently Q_j^{MVS}. Furthermore, the MV and MVS MPC formulations introduce cross-multiplication of outputs that are predicted to be corrupted with the same disturbance signals, since these outputs are cross correlated with each other. Thus the moment-based MPC weights $Q_{i,j}^{MV}$ reflect the cross-correlation in between the output predictions. Since at the first prediction step, the current output value is not correlated with predicted outputs at the later stages, the $Q_{0,j}^{MV}$ terms are equal to zero. We demonstrate the output weighting matrices in Figure 3.1, which visualizes the changes in mean (hence nominal) MPC, MV MPC and MVS MPC for three simple systems; we consider a two state system, one of them integrating the other state, and vary the pole locations for; a stable pair case ($p_i = 0.9$, Figure 3.1a), a marginally stable case ($p_i = 1$, Figure 3.1b) and an unstable pole pair case ($p_i = 1.1, i = 1, 2$, Figure 3.1c). In these examples, the system is excited by control input from first (integrated) state and measured from the second state, while disturbances are acting on both of the states. The noise covariances are selected as $\Sigma_w = I_2$, $\Sigma_v = 1$, the prediction horizon[7] is taken as $N_p = 100$ and risk aversion factors are taken as $\lambda_v = 1, \lambda_s = 1$.

The last observation that we mention here is connected to the frequency domain characteristics of the closed-loop systems. Since the controller resulting from an unconstrained MPC problem is equivalent to a static feedback controller, one can evaluate the changes in the closed-loop characteristics. Traditionally, the output weighting matrix Q is tuned, while keeping the R matrix constant, to result in desired closed-loop characteristics. Without any further analysis, it can be claimed that as the output weighting matrix grows in magnitude, the controller becomes more and more aggressive. Since in moment-based MPC, the output weighting matrix grows, the MV and MVS MPC formulations are expected to increase the bandwidth of the system and introduce lead action to the closed-loop process. We demon-

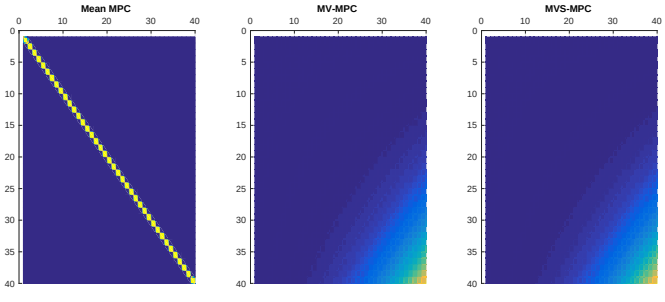
[7] In moment-based MPC problems, the prediction horizon can be extended till the nominal problem becomes intractable with the available computing devices, which is not the case in other RMPC methods.



(a) The values of $\bar{Q}$, $\bar{Q}^{MV}$ and $\bar{Q}^{MVS}$ matrices for system with stable poles ($p_i = 0.9$).



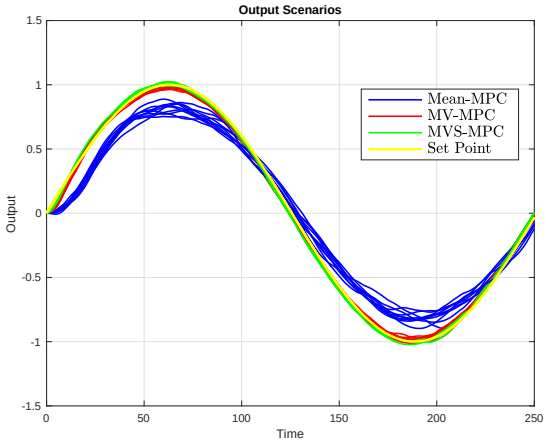
(b) The values of $\bar{Q}$, $\bar{Q}^{MV}$ and $\bar{Q}^{MVS}$ matrices for a marginally stable system.



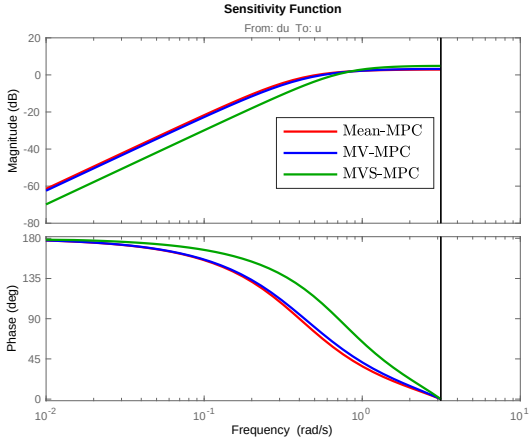
(c) The values of $\bar{Q}$, $\bar{Q}^{MV}$ and $\bar{Q}^{MVS}$ matrices for a system with unstable poles ($p_i = 1.1$).

Figure 3.1: Comparison of $\bar{Q}^{(\cdot)}$ matrix for different system structures.

strate this through the simulation example used above. We report the closed-loop sensitivity functions and the phase response of the closed-loop systems the marginally stable case in Figure 3.2. As can be observed from the figures, as the output (or state) weighting matrices for MV and MVS MPC cases, i.e., $Q_{j|k}^{MV}$ or $Q_{j|k}^{MVS}$, (stage-wise) increases in comparison



(a) The sine wave tracking with different moment-based MPC formulations.



(b) The sensitivity function of closed-loop system with different moment-based MPC formulations.

Figure 3.2: The sensitivity function of closed-loop system with different moment-based MPC formulations.

to the nominal output weighting matrix Q, the resulting controller gain for MV and MVS MPC cases increases. This leads to an increase in the closed-loop bandwidth. Describing the change in bandwidth (or phase response) according to the tuning parameter selection of λ_v or λ_s is left as a future research subject.

Linear Cost Functions

One drawback of moment-based MPC can be observed if one considers cost functions containing only the linear terms of outputs (equivalently the states). Let us consider a (linear) cost function $J_L : \mathbb{R}^{n_x} \times \mathbb{R}^{N_p n_u} \to \mathbb{R}$ as

$$\begin{aligned} J_L(x_k, u_{[0,N_p-1]|k}) = & \sum_{j=0}^{N_p-1} c_{y,j}^\top (y_{j|k} - r_{j|k}) + c_{u,j}^\top u_{j|k}, \\ = & \sum_{j=0}^{N_p-1} \langle c_{y,j}, y_{j|k} - r_{j|k} \rangle + \langle c_{u,j}, u_{j|k} \rangle, \end{aligned} \tag{3.37}$$

where $c_{y,i} \in \mathbb{R}^{n_y}, c_{u,i} \in \mathbb{R}^{n_u}, i = 0, \ldots, N_p - 1$.

It is straightforward to show that, by using the reasoning provided in Lemma 3.1.1, the Mean MPC the cost function $J_L^M(\xi_{k,(i)}^O) = \mathbb{E}\{J_L(\tilde{\xi}_{k,(i)}^O)\}$ is resulting with the nominal predictions of outputs, i.e.,

$$J_L^M(x_k, u_{[0,N_p-1]|k}) = \sum_{j=0}^{N_p-1} \langle c_{y,j}, \bar{y}_{j|k} - r_{j|k} \rangle + \langle c_{u,j}, u_{j|k} \rangle,$$

and hence resulting in the same optimal control action with the nominal MPC formulation. To overcome the detrimental effects of perturbations we incorporate high order statistics and observe that,

$$\mathbb{E}\{(J_L(\tilde{\xi}_{k,(i)}^O) - \mathbb{E}\{J_L(\tilde{\xi}_{k,(i)}^O)\})^m\} = \mathbb{E}\left\{ \left(\sum_{i=0}^{N_p-1} \left\langle c_{y,j}, C \sum_{i=0}^{j-1} A^{j-1-i} F w_{i|k} + v_{j|k} \right\rangle \right)^m \right\}. \tag{3.38}$$

These terms are functions of decision variables $u_{[0,N_p-1]|k}$ for any $m > 1$ and thus does not change the optimal control action.

In order to incorporate the effect of spread into the robust counterpart cost functions for higher order moments, one can consider polytopic norms, i.e., the infinity norm ($\|(\cdot)\|_\infty$) or the 1-norm ($\|(\cdot)\|_1$). However expectations of uncertain variables with these norms are far from trivial to express explicitly and generally lead to non-convex functions of initial condition $x_{0|k}$ and control actions $u_{j|k}$. Hence it is expected that moment-based MPC formulations result in undesirable increase in the computational complexity of the MPC problem. Simplifications and possible relaxations for the mentioned norms are currently an ongoing research activity.

Costs with Explicit Uncertain Terms

Another common selection for the MPC cost function is the cost functions containing disturbance terms explicitly. These type of costs are used to cancel the tail effects of uncertainties after the prediction horizon, see [235]. Here we demonstrate the effect of the explicit disturbance terms for the moment-based MPC formulations. For this purpose consider the cost function $J_W(\tilde{\xi}_k)$ which is defined through the outputs, inputs and additive disturbances of the system Σ^{add} as

$$J_W(\tilde{\xi}_k) = \sum_{j=0}^{N_p-1} (y_{j|k} - r_{j|k})^\top Q (y_{j|k} - r_{j|k}) + u_{j|k}^\top R_1 u_{j|k} - w_{j|k}^\top R_2 w_{j|k}. \tag{3.39}$$

For this cost function we summarize the effect of explicit disturbance terms in Lemma 3.1.5

Lemma 3.1.5 *Given a LTI dynamical system Σ^{add} as in Equation* (3.1b) *and a reference trajectory to follow $r_k, k \in \mathbb{Z}_{\geq 0}$, the cost functions for mean, MV and MVS MPC formulations, for the cost function J_W in Equation* (3.39)*, are given as*

$$\begin{aligned} J_W^M(\xi_k) &= \xi_k^\top H_W^M \xi_k + f_W^M, \\ J_W^{MV}(\xi_k) &= \xi_k^\top H_W^{MV} \xi_k + f_W^M + \lambda_v f_W^V, \\ J_W^{MVS}(\xi_k) &= \xi_k^\top H_W^{MVS} \xi_k + f_W^M + \lambda_v f_W^V + \lambda_s f_W^S, \end{aligned} \tag{3.40a}$$

where

$$\begin{aligned} \xi_k^\top &= \begin{bmatrix} x_0^\top & u_{[0,N_p-1]|k}^\top & r_{[0,N_p-1]|k}^\top \end{bmatrix}, \\ H_W^M &= \begin{bmatrix} \Theta_{N_p,C}^\top \bar{Q} \Theta_{N_p,C} & (\star) & (\star) \\ T_{B,C}^{N_p^\top} \bar{Q} \Theta_{N_p,C} & T_{B,C}^{N_p^\top} \bar{Q} T_{B,C}^{N_p} + \bar{R}_1 & (\star) \\ -\bar{I}_{N_p}^\top \bar{Q} \Theta_{N_p,C} & -\bar{I}_{N_p}^\top \bar{Q} T_{B,C}^{N_p} & \bar{I}_{N_p}^\top \bar{Q} \bar{I}_{N_p} \end{bmatrix}, \\ H_W^{MV} &= \begin{bmatrix} \Theta_{N_p,C}^\top \bar{Q}_W^{MV} \Theta_{N_p,C} & (\star) & (\star) \\ T_{B,C}^{N_p^\top} \bar{Q}_W^{MV} \Theta_{N_p,C} & (T_{B,C}^{N_p^\top} \bar{Q}_W^{MV} T_{B,C}^{N_p} + \bar{R}_1 & (\star) \\ -\bar{I}_{N_p}^\top \bar{Q}_W^{MV} \Theta_{N_p,C} & -\bar{I}_{N_p}^\top \bar{Q}_W^{MV} T_{B,C}^{N_p} & \bar{I}_{N_p}^\top \bar{Q}_W^{MV} \bar{I}_{N_p} \end{bmatrix}, \\ H_W^{MVS} &= \begin{bmatrix} \Theta_{N_p,C}^\top \bar{Q}_W^{MVS} \Theta_{N_p,C} & (\star) & (\star) \\ T_{B,C}^{N_p^\top} \bar{Q}_W^{MVS} \Theta_{N_p,C} & T_{B,C}^{N_p^\top} \bar{Q}_W^{MVS} T_{B,C}^{N_p} + \bar{R}_1 & (\star) \\ -\bar{I}_{N_p}^\top \bar{Q}_W^{MVS} \Theta_{N_p,C} & -\bar{I}_{N_p}^\top \bar{Q}_W^{MVS} T_{B,C}^{N_p} & \bar{I}_{N_p}^\top \bar{Q}_W^{MVS} \bar{I}_{N_p} \end{bmatrix}, \end{aligned} \tag{3.40b}$$

$$\begin{aligned} \bar{Q}_W^{MV} &= \bar{Q} + 4\lambda_v \bar{Q} \bar{\Upsilon}_W \bar{Q}, \\ \bar{Q}_W^{MVS} &= \bar{Q}_W^{MV} + 24\lambda_s \bar{Q} \beta_{W,R_2} \bar{Q}, \\ \bar{\Upsilon}_W &:= T_{F,C}^{N_p} \bar{\Sigma}_w^{N_p} T_{F,C}^{N_p^\top} + \bar{\Sigma}_v^{N_p}, \\ \beta_{W,R_2} &= T_{F,C}^{N_p} \bar{\Sigma}_w^{N_p} (T_{F,C}^{N_p^\top} \bar{Q} T_{F,C}^{N_p} - R_2) \bar{\Sigma}_w^{N_p} T_{F,C}^{N_p^\top}. \end{aligned} \tag{3.40c}$$

3.2 Generalization of Moment-based MPC Formulations for Additive Perturbations with Even Distribution Functions

In this section, we change the characteristics of additive uncertainties to generalize the results for the moment-based MPC towards random variables that are not distributed with Gaussian pdfs. We consider the LTI system with additive disturbances, i.e., Σ^{add}, but in this case we assume that the perturbations w_k and v_k are distributed with pdfs that have even characteristics (symmetric distribution function w.r.t. zero). Similar to the previous sections, we consider quadratic cost function $J_{O,(i)}(\tilde{\xi}_k)$ as in Equation (3.20a), but here we

denote the cost function with $J_{O,E}(\tilde{\xi}_k)$ to highlight the even pdf characteristics[8], i.e.,

$$J_{O,E}(\tilde{\xi}_k) = \sum_{j=0}^{N_p-1} (y_{j|k} - r_{j|k})^\top Q(y_{j|k} - r_{j|k}) + u_{j|k} R u_{j|k}, \tag{3.41}$$

where $\tilde{\xi}_k = \begin{bmatrix} x_{0|k}^\top & u_{[0,N_p-1]|k}^\top & w_{[0,N_p-1]|k}^\top & v_{[0,N_p-1]|k}^\top & r_{[0,N_p-1]|k}^\top \end{bmatrix}^\top$. Under this setting, we formulate the moment-based MPC problems, specifically mean, MV or MVS MPC problems, for non-Gaussian but even random variables. The deterministic cost functions projected via moment operators are defined as

$$J_{O,E}^{M}(\xi_k) := \mathbb{E}\{J_{O,E}(\tilde{\xi}_k)\}, \tag{3.42a}$$
$$J_{O,E}^{MV}(\xi_k) := \mathbb{E}\{J_{O,E}(\tilde{\xi}_k)\} + \lambda_v \mathbb{D}\{J_{O,E}(\tilde{\xi}_k)\}, \tag{3.42b}$$
$$J_{O,E}^{MVS}(\xi_k) := \mathbb{E}\{J_{O,E}(\tilde{\xi}_k)\} + \lambda_v \mathbb{D}\{J_{O,E}(\tilde{\xi}_k)\} + \lambda_s \mathbb{S}\{J_{O,E}(\tilde{\xi}_k)\}, \tag{3.42c}$$

where ξ_k is defined as in Equation (3.21). Before providing the explicit forms of the cost functions of Mean, MV and MVS MPC, we use the results provided in Lemma 3.2.1 for the calculation of higher order centralized moments of the quadratic forms in the cost function $J_{O,E}(\tilde{\xi}_k)$.

Lemma 3.2.1 *Consider an LTI stochastic dynamical system Σ^{add} as in Equation* (3.19b)*, a reference trajectory $r_k, k \in \mathbb{Z}_{\geq 0}$ and disturbances w_k and v_k in Σ^{add} being non-Gaussian random variables that are subject to Assumption 3.1.2. Given the quadratic cost function $J_{O,E}(\tilde{\xi}_k)$ as in Equation* (3.41)*, then, for any $k \in \mathbb{Z}_{\geq 0}$, $p_1 \in \mathbb{R}^{n_w}, P_1 \in \mathbb{R}^{n_w \times n_w}$ and $P = P^\top \in \mathbb{R}^{n_w \times n_w}$, the following statements are true;*

- *For $J_{O,E}^{M}(\xi_k)$;*

$$\mathbb{E}\{p_1^\top w_k\} = \quad 0,$$

- *For $J_{O,E}^{MV}(\xi_k)$;*

$$\begin{aligned} \mathbb{E}\{p_1^\top w_k w_k^\top P_1 w_k\} &= \quad 0, \\ \mathbb{E}\{p_1^\top w_k w_k^\top p_1\} &= \quad p_1^\top \bar{\Sigma}_w^{N_p} p_1. \end{aligned}$$

- *For $J_{O,E}^{MVS}(\xi_k)$;*

$$\begin{aligned} \mathbb{E}\{p_1^\top w_k w_k^\top P w_k w_k^\top p_1\} &= \quad p_1^\top \beta_{w,1} p_1, \\ \mathbb{E}\{p_1^\top w_k w_k^\top p_1 \mathrm{Tr}(P\Sigma_w)\} &= \quad p_1^\top \beta_{w,2} p_1, \end{aligned}$$

where

$$\beta_{w,1} = \begin{bmatrix} \mathbb{E}\{[w]_1^4\} P_{1,1} + f_\beta(1) & 2\sigma_1^2\sigma_2^2 P_{1,2} & \dots & 2\sigma_1^2\sigma_{n_w}^2 P_{1,n_w} \\ 2\sigma_2^2\sigma_1^2 P_{2,1} & \mathbb{E}\{[w]_2^4\} P_{2,2} + f_\beta(2) & \ddots & 2\sigma_2^2\sigma_{n_w}^2 P_{n_w,2} \\ \vdots & \ddots & \ddots & \vdots \\ 2\sigma_1^2\sigma_{n_w}^2 P_{n_w,1} & 2\sigma_2^2\sigma_{n_w}^2 P_{n_w,2} & \dots & \mathbb{E}\{[w]_{n_w}^4\} P_{n_w,n_w} + f_\beta(n_w) \end{bmatrix}, \tag{3.43a}$$

[8] We, also, assume that the state x_k is correctly and fully measurable at every time instant, for simplicity.

$$\beta_{w,2} = \begin{bmatrix} \sigma_1^2 \sum_{i=1}^{n_w} P_{i,i}\sigma_i^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 \sum_{i=1}^{n_w} P_{i,i}\sigma_i^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{n_w}^2 \sum_{i=1}^{n_w} P_{i,i}\sigma_i^2 \end{bmatrix}, \tag{3.43b}$$

where $f_\beta(i) = \sum_{j \in \mathcal{J}_i} P_{j,j}\sigma_j^2\sigma_i^2$, $\mathcal{J}_i = \mathbb{Z}_{[1,n_w]} \backslash \{i\}$.

Proof.
For the first and the second claims, we show that the odd order centralized moments of random variables with even pdf are zero. Assuming $g(w)$ is an odd function, and denoting the pdf of w with $f_{\tilde{w}}(w)$,

$$\mathbb{E}\{g(w)\} = \int_{-\infty}^{\infty} g(w) f_{\tilde{w}}(w) dw = -\int_{0}^{\infty} g(w) f_{\tilde{w}}(w) dw + \int_{0}^{\infty} g(w) f_{\tilde{w}}(w) dw = 0.$$

The third claim is a trivial reformulation of the definition of the covariance matrix. Lastly, for the terms in MVS-MPC formulations, it is straightforward to see the claim for $\beta_{w,2}$. For $\beta_{w,1}$, explicitly evaluating the expectation operation yields the claimed result by using the properties of expectations of odd order combinations. □

With the results stated in Lemma 3.2.1, one can make use of different types of uncertainty distributions to adjust the perceived affects of uncertainty. Next, we state that the deterministic counterparts of moment-based MPC cost functions for the Theorem 3.2.1.

$$\bar{J}_{O,E}^{M}(\xi_k) = \quad \xi_k^\top H_{O,E}^{M} \xi_k, \tag{3.44a}$$

$$\bar{J}_{O,E}^{MV}(\xi_k) = \quad \xi_k^\top H_{O,E}^{MV} \xi_k, \tag{3.44b}$$

$$\bar{J}_{O,E}^{MVS}(\xi_k) = \quad \xi_k^\top H_{O,E}^{MVS} \xi_k, \tag{3.44c}$$

where

$$H_{O,E}^{M} = \begin{bmatrix} \Theta_C^\top \bar{Q} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q} \Theta_C & T_{B,C}^\top \bar{Q} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.44d}$$

$$H_{O,E}^{MV} = \begin{bmatrix} \Theta_C^\top \bar{Q}_{O,E}^{MV} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_{O,E}^{MV} \Theta_{N_p,C} & T_{B,C}^\top \bar{Q}_{O,E}^{MV} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,E}^{MV} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,E}^{MV} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,E}^{MV} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.44e}$$

$$H_{O,E}^{MVS} = \begin{bmatrix} \Theta_C^\top \bar{Q}_{O,E}^{MVS} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_{O,E}^{MVS} \Theta_C & T_{B,C}^\top \bar{Q}_{O,E}^{MVS} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,E}^{MVS} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,E}^{MVS} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,E}^{MVS} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.44f}$$

$$\begin{aligned} \bar{Q}_{O,E}^{MV} &= \bar{Q} + 4\lambda_v \bar{Q}\bar{\Upsilon}_{O,E}\bar{Q}, \\ \bar{Q}_{O,E}^{MVS} &= \bar{Q}_{O,E}^{MV} + 24\lambda_s \bar{Q}\tilde{\Upsilon}_{O,E}\bar{Q}, \end{aligned} \tag{3.44g}$$

$$\begin{aligned} \bar{\Upsilon}_{O,E} &= T_{F,C}\bar{\Sigma}_w^{N_p} T_{F,C}^{\top} + \bar{\Sigma}_v^{N_p}, \\ \tilde{\Upsilon}_{O,E} &:= \bar{\Gamma}_{w,E}(T_{F,C}^{N_p^{\top}} \bar{Q} T_{F,C}^{N_p}) + \bar{\Gamma}_{v,E}(\bar{Q}), \end{aligned} \tag{3.44h}$$

$$\bar{\Gamma}_{w,E}(Q) = \bar{\Gamma}_{w,E}^{\top}(Q) := \begin{cases} (\mathbb{E}\{[\tilde{w}]_i^4\} - \sigma_{\tilde{w}_i}^4)Q(i,i), & (i,i)^{\text{th}} \text{ entry} \\ 2\sigma_{\tilde{w}_i}^2 \sigma_{\tilde{w}_j}^2 Q_{i,j} & (i,j)^{\text{th}} \text{ entry} \end{cases} \tag{3.44i}$$

$$\begin{aligned} \bar{\Sigma}_w^{N_p} &= I_{N_p} \otimes \Sigma_w, \\ \bar{\Sigma}_v^{N_p} &= I_{N_p} \otimes \Sigma_v, \end{aligned} \tag{3.44j}$$

Theorem 3.2.1 *Consider an LTI stochastic dynamical system Σ^{add} as in Equation* (3.19b) *subject to Assumption 3.1.2 and a reference trajectory $r_k, k \in \mathbb{Z}_{\geq 0}$. Then;*

- *The cost function for Mean MPC problem with the stochastic cost function $J_{O,E}(\tilde{\xi}_k)$ defined in Equation* (3.41) *satisfy*

$$\begin{aligned} J_{O,E}^{M}(\xi_k) &= \mathbb{E}\{J_{O,E}(\tilde{\xi}_k)\}, \\ &= \xi_k^{\top} H_{O,E}^{M} \xi_k + f_{O,E}^{M}, \end{aligned} \tag{3.45a}$$

Furthermore,

$$\bar{J}_{O,E}^{M}(\xi_k) = J_{O,E}^{M}(\xi_k) - f_{O,E}^{M}, \tag{3.45b}$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} J_{O,E}^{M}(\xi_k) &= \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,E}^{M}(\xi_k) + f_{O,E}^{M}, \\ \arg\min_{u_{[0,N_p-1]|k}} J_{O,E}^{M}(\xi_k) &= \arg\min_{u_{[0,N_p-1]|k}} \bar{J}_{O,E}^{M}(\xi_k), \end{aligned} \tag{3.45c}$$

- *The cost function for MV MPC problem with the stochastic cost function $J_{O,E}(\tilde{\xi}_k)$ Equation* (3.41) *is given by*

$$\begin{aligned} J_{O,E}^{MV}(\xi_k) &= \mathbb{E}\{J_{O,E}(\tilde{\xi}_k)\} + \lambda_v \mathbb{D}\{J_{O,E}(\tilde{\xi}_k)\}, \\ &= \xi_k^{\top} H_{O,E}^{MV} \xi_k + f_{O,E}^{M} + \lambda_v f_{O,E}^{V}, \end{aligned} \tag{3.46a}$$

Furthermore,

$$\bar{J}_{O,E}^{MV}(\xi_k) = J_{O,E}^{MV}(\xi_k) - f_{O,E}^{M} - \lambda_v f_{O,E}^{V}, \tag{3.46b}$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} J_{O,E}^{MV}(\xi_k) &= \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,E}^{MV}(\xi_k) + f_{O,E}^{M} + \lambda_v f_{O,E}^{V}, \\ \arg\min_{u_{[0,N_p-1]|k}} J_{O,E}^{MV}(\xi_k) &= \arg\min_{u_{[0,N_p-1]|k}} \bar{J}_{O,E}^{MV}(\xi_k). \end{aligned} \tag{3.46c}$$

- *The cost function for MVS MPC problem with the stochastic cost function* $J_{O,E}(\tilde{\xi}_k)$ *in Equation* (3.41) *is given by*

$$\begin{aligned} J_{O,E}^{MVS}(\xi_k) = &\ \mathbb{E}\{J_{O,E}(\tilde{\xi}_k)\} + \lambda_v \mathbb{D}\{J_{O,E}(\tilde{\xi}_k)\} + \lambda_s \mathbb{S}\{J_{O,E}(\tilde{\xi}_k)\}, \\ = &\ \xi_k^\top H_{O,E}^{MVS} \xi_k + f_{O,E}^M + \lambda_v f_{O,E}^V + \lambda_s f_{O,E}^S. \end{aligned} \tag{3.47a}$$

Furthermore,

$$\bar{J}_{O,E}^{MVS}(\xi_k) = \ J_{O,E}^{MVS}(\xi_k) - f_{O,E}^M - \lambda_v f_{O,E}^V - \lambda_s f_{O,E}^S, \tag{3.47b}$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1|k}} J_{O,E}^{MVS}(\xi_k) = &\ \min_{u_{[0,N_p-1|k}} \bar{J}_{O,E}^{MVS}(\xi_k) + f_{O,E}^M + \lambda_v f_{O,E}^V + \lambda_s f_{O,E}^S, \\ \arg\min_{u_{[0,N_p-1|k}} J_{O,E}^{MVS}(\xi_k) = &\ \arg\min_{u_{[0,N_p-1|k}} \bar{J}_{O,E}^{MVS}(\xi_k). \end{aligned} \tag{3.47c}$$

Proof. The proof follows from the results of Lemmas 3.2.1-3.1.1-3.1.2-3.1.3. □

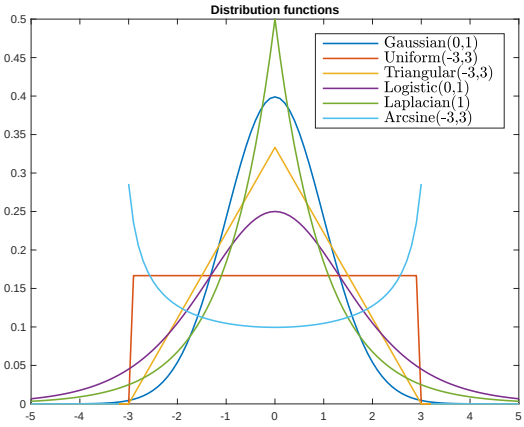
To highlight the dependency on the uncertainty model we present Table 3.1, in which the variance and the fourth order moment values of some standard even pdf random variables are reported. By selecting the distribution function, the designer is able to adjust the risk-awareness of the closed-loop system, see Figure 3.3.

Even Uncertainty Type	Characteristic Values	Variance Term	4^{th} Order Term
Uniform Distribution	$\tilde{w} \in \mathcal{U}(-a, a)$	$\frac{1}{12} 4a^2 I_{N_p}$	$\frac{1}{80} 16a^4 I_{N_p}$
Triangular Distribution	$\tilde{w} \in \mathcal{T}(-a, a, 0)$	$\frac{1}{18} 3a^2 I_{N_p}$	$\frac{1}{135} 9a^4 I_{N_p}$
Logistic Distribution	$\tilde{w} \in \mathcal{L}o(0, b)$	$\frac{1}{3}\pi^2 b^2 I_{N_p}$	$\frac{7}{15}\pi^4 b^4 I_{N_p}$
Laplacian Distribution	$\tilde{w} \in \mathcal{L}a(0, b)$	$2b^2 I_{N_p}$	$24b^4 I_{N_p}$
Arcsine Distribution	$\tilde{w} \in \mathcal{A}(-a, a)$	$\frac{1}{8} 4a^2 I_{N_p}$	$\frac{3}{8} a^4 I_{N_p}$

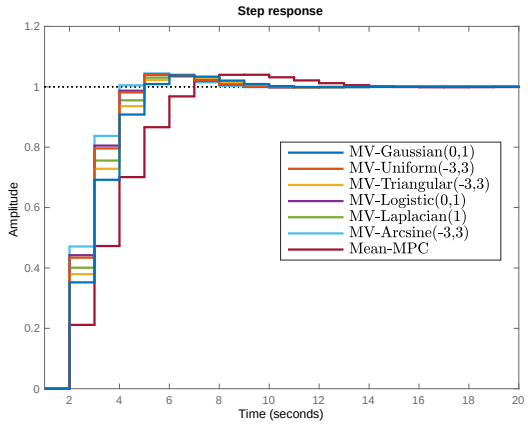
Table 3.1: An overview of even pdfs and corresponding MV and MVS weighting terms.

3.3 Output Tracking Moment-based MPC with Corrupted Initial Condition

So far, we assumed the current state is fully accessible by the MPC controller to be evaluated in the optimization routine. In what follows we present the effect of the uncertainties in the



(a) The i.i.d pdfs of uncertain variables w_k and v_k used in the simulation.



(b) The step response of the controlled systems with different moment-based MPC formulations.

Figure 3.3: The dependency of moment-based MPC formulations on the uncertainty model.

initialization point of the MPC problem. First we formulate the mean, MV and MVS MPC problems for the system Σ^{nom} with the assumption that the initialized state is known upto some uncertainty. Later we incorporate the additive perturbations and provide an algorithm for output feedback MPC running with an state estimator in closed-loop.

3.3.1 Cost Reformulations for Moment-based MPC With Initial Condition Mismatch

Here, we assume that the initial condition used in the MPC problem, i.e., $x_{0|k}$, is differing from the true state x_k, but it is known upto some extent that is described by an uncertainty with known characteristics, i.e.,

$$x_k = \bar{x}_{0|k} + \zeta_k, \tag{3.48}$$

where we desire that the state that will be used in the moment-based MPC problem is predicted from the initial condition $\bar{x}_{0|k}$ and with the dynamics Σ^{nom} in Equation (3.19a). Since the distribution characteristics of the random variable ζ_k is known, one can incorporate the uncertain effects into the MPC formulation, similar to the additive perturbations. Here, for simplicity, we assume that the random vector ζ_k is distributed with Gaussian distribution function, i.e., $\zeta_k \in \mathcal{N}(0, \Sigma_{\zeta_k})$. We consider the output reference tracking problem for Σ^{nom}, hence no w_k or v_k effecting the evolution, thus using the cost function $J_{O,I}(\tilde{\xi}_k^O)$,

$$J_{O,I}(\bar{x}_k, u_{[0,N_p-1]|k}, \zeta_k) = \sum_{j=0}^{N_p-1} (\bar{y}_{j|k} - r_{j|k})^\top Q(\bar{y}_{j|k} - r_{j|k}) + u_{j|k}^\top R u_{j|k}, \tag{3.49}$$

where $Q \succeq 0, R \succ 0$ and, in contrast to the previous MPC problems, in this case the outputs $\bar{y}_{j|k}, j \in \mathbb{Z}_{[0,N_p-1]|k}$ are random variables, due to ζ_k. Next we state the robust counterpart MPC problems for the mean, MV and MVS MPc problems, where these problems are defined through the cost functions,

$$\begin{aligned}
J_{O,I}^M(\xi_k^O) &:= \mathbb{E}\{J_{O,I}(\tilde{\xi}_k^O)\}, \\
J_{O,I}^{MV}(\xi_k^O) &:= \mathbb{E}\{J_{O,I}(\tilde{\xi}_k^O)\} + \lambda_v \mathbb{D}\{J_{O,I}(\tilde{\xi}_k^O)\}, \\
J_{O,I}^{MVS}(\xi_k^O) &:= \mathbb{E}\{J_{O,I}(\tilde{\xi}_k^O)\} + \lambda_v \mathbb{D}\{J_{O,I}(\tilde{\xi}_k^O)\} + \lambda_s \mathbb{S}\{J_{O,I}(\tilde{\xi}_k^O)\},
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\xi}_k^O &= \begin{bmatrix} \bar{x}_{0|k}^\top & u_{[0,N_p-1]|k}^\top & \zeta_k^\top & r_{[0,N_p-1]|k} \end{bmatrix}^\top, \\
\xi_k^O &= \begin{bmatrix} \bar{x}_{0|k}^\top & u_{[0,N_p-1]|k}^\top & r_{[0,N_p-1]|k} \end{bmatrix}^\top,
\end{aligned}$$

We, also, define the following nominal MPC cost functions as,

$$\bar{J}_{O,I}^M(\xi_k^O) = {\xi_k^O}^\top H_{O,I}^M \xi_k^O, \tag{3.50a}$$

$$\bar{J}_{O,I}^{MV}(\xi_k^O) = {\xi_k^O}^\top H_{O,I}^{MV} \xi_k^O, \tag{3.50b}$$

$$\bar{J}_{O,I}^{MVS}(\xi_k^O) = {\xi_k^O}^\top H_{O,I}^{MVS} \xi_k^O, \tag{3.50c}$$

where

$$H_{O,I}^M = \begin{bmatrix} \Theta_C^\top \bar{Q} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q} \Theta_C & T_{B,C}^\top \bar{Q} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.50d}$$

$$H_{O,I}^{MV} = \begin{bmatrix} \Theta_C^\top \bar{Q}_{O,I}^{MV} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_{O,I}^{MV} \Theta_{N_p,C} & T_{B,C}^\top \bar{Q}_{O,I}^{MV} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,I}^{MV} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,I}^{MV} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,I}^{MV} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.50e}$$

$$H_{O,I}^{MVS} = \begin{bmatrix} \Theta_C^\top Q^- I_O^{MVS} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_{O,I}^{MVS} \Theta_C & T_{B,C}^\top \bar{Q}_{O,I}^{MVS} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,I}^{MVS} \Theta_C & -\bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,I}^{MVS} T_{B,C} & \bar{I}_{n_y}^{N_p\top} \bar{Q}_{O,I}^{MVS} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \quad (3.50f)$$

$$\begin{aligned} \bar{Q}_{O,I}^{MV} &= \bar{Q} + 4\lambda_v \bar{Q} \bar{\Upsilon}_{O,I} \bar{Q}, \\ \bar{Q}_{O,I}^{MVS} &= \bar{Q}_{O,I}^{MV} + 24\lambda_s \bar{Q} \bar{\Upsilon}_{O,I} \bar{Q} \bar{\Upsilon}_{O,I} \bar{Q}, \\ \bar{\Upsilon}_{O,I} &= \Theta_C \Sigma_{\zeta_k} \Theta_C^\top. \end{aligned} \quad (3.50g)$$

Theorem 3.3.1 *Given an LTI dynamical system Σ^{nom} as in Equation* (3.19a) *and a reference trajectory $r_k, k \in \mathbb{Z}_{\geq 0}$. Consider the moment-based MPC problem with cost function $J_{O,I}(\tilde{\xi}_k^O$ as in Equation* (3.49) *and assume that there exists an initial condition mismatch described as in Equation* (3.48) *with the characteristics given as $\zeta_k \in \mathcal{N}(0, \Sigma_{\zeta_k})$. Then;*

- *The cost function for Mean MPC problem with the stochastic cost function $J_{O,I}(\tilde{\xi}_k^O)$, i.e., $J_{O,I}^M(\xi_k^O)$, satisfy*

$$\begin{aligned} J_{O,I}^M(\xi_k^O) &= \mathbb{E}\{J_{O,I}(\tilde{\xi}_k^O)\}, \\ &= \xi_k^{O\top} H_{O,I}^M \xi_k^O + f_{O,I}^M, \end{aligned} \quad (3.51a)$$

Furthermore the $f_{O,I}^M$ term is constant and

$$\bar{J}_{O,I}^M(\xi_k^O) = J_{O,I}^M(\xi_k^O) - f_{O,I}^M, \quad (3.51b)$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} J_{O,I}^M(\xi_k^O) &= \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,I}^M(\xi_k^O) + f_{O,I}^M, \\ \arg\min_{u_{[0,N_p-1]|k}} J_{O,I}^M(\xi_k^O) &= \arg\min_{u_{[0,N_p-1]|k}} \bar{J}_{O,I}^M(\xi_k^O). \end{aligned} \quad (3.51c)$$

- *The cost function for MV MPC problem with the stochastic cost function $J_{O,I}(\tilde{\xi}_k^O)$, i.e., $J_{O,I}^{MV}(\xi_k^O)$, satisfy*

$$\begin{aligned} J_{O,I}^{MV}(\xi_k^O) &= \mathbb{E}\{J_{O,I}(\tilde{\xi}_k^O)\} + \lambda_v \mathbb{D}\{J_{O,I}(\tilde{\xi}_k^O)\}, \\ &= \xi_k^{O\top} H_{O,I}^{MV} \xi_k^O + f_{O,I}^M + \lambda_v f_{O,I}^V, \end{aligned} \quad (3.52a)$$

Furthermore,

$$\bar{J}_{O,I}^{MV}(\xi_k^O) = J_{O,I}^{MV}(\xi_k^O) - f_{O,I}^M - \lambda_v f_{O,I}^V, \quad (3.52b)$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} J_{O,I}^{MV}(\xi_k^O) &= \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,I}^{MV}(\xi_k^O) + f_{O,I}^M + \lambda_v f_{O,I}^V, \\ \arg\min_{u_{[0,N_p-1]|k}} J_{O,I}^{MV}(\xi_k^O) &= \arg\min_{u_{[0,N_p-1]|k}} \bar{J}_{O,I}^{MV}(\xi_k^O). \end{aligned} \quad (3.52c)$$

- *The cost function for MVS MPC problem with the stochastic cost function $J_{O,I}$, i.e., $J_{O,I}^{MVS}(\xi_k^O)$, satisfy*

$$\begin{aligned} J_{O,I}^{MVS}(\xi_k^O) = &\quad \mathbb{E}\{J_{O,I}(\tilde{\xi}_O)\} + \lambda_v \mathbb{D}\{J_{O,I}(\tilde{\xi}_k^O)\} + \lambda_s \mathbb{S}\{J_{O,I}(\tilde{\xi}_k)\}, \\ = &\quad {\xi_k^O}^\top H_{O,I}^{MVS} \xi_k^O + f_{O,I}^M + \lambda_v f_{O,I}^V + \lambda_s f_{O,I}^S, \end{aligned} \tag{3.53a}$$

Furthermore,

$$\bar{J}_{O,I}^{MVS}(\xi_k^O) = \quad J_{O,I}^{MVS}(\xi_k^O) - f_{O,I}^M - \lambda_v f_{O,I}^V - \lambda_s f_{O,I}^S, \tag{3.53b}$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} J_{O,I}^{MVS}(\xi_k^O) = &\quad \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,I}^{MVS}(\xi_k^O) + f_{O,I}^M + \lambda_v f_{O,I}^V + \lambda_s f_{O,I}^S, \\ \arg\min_{u_{[0,N_p-1]|k}} J_{O,I}^{MVS}(\xi_k^O) = &\quad \arg\min_{u_{[0,N_p-1]|k}} \bar{J}_{O,I}^{MVS}(\xi_k^O). \end{aligned} \tag{3.53c}$$

Proof. The proof follows trivially from the proofs given in Lemmas 3.1.1-3.1.2-3.1.3, with redefining the ζ_k as the exogenous signal $w_{0|k}$ which is assuming nonzero value only at prediction stage $j = 0$ and setting all other disturbance to zero. □

3.3.2 Reference Tracking with Output Feedback Moment-based MPC

In this section we present a complete result which incorporates both of the additive perturbations w_k and v_k while the MPC controller is not receiving the full state information, i.e., $\zeta_k \neq 0$. For brevity, we assume that the controller is receiving the state estimates from a Kalman filter that is designed beforehand. This means that the estimated state of Kalman filter ($\hat{x}_{k|k}$) is used as the initialization state of the MPC problem $\bar{x}_{0|k}$.

Here we assume that Kalman filter is initialized with the initial estimate $\hat{x}_{0|0} \neq x_0$, and the initial state covariance matrix $P_{0|0} := \mathbb{E}\{(x_0 - \hat{x}_{0|0})(x_0 - \bar{x}_{0|0})^\top\}$. Furthermore, we assume that at every time instant, the Kalman filter outputs an estimate of the true state, i.e., $\hat{x}_{k|k}$ with a corresponding state covariance matrix $P_{k|k}$ via the Kalman update equations with the time varying gain matrix $L_k := P_{k|k-1}C^\top(C^\top P_{k|k-1}C + \Sigma_v)^{-1}$, i.e.;
for the prediction step:

$$\begin{aligned} \hat{x}_{k|k-1} = &\quad A\hat{x}_{k-1|k-1} + Bu_{k-1}, \\ P_{k|k-1} = &\quad AP_{k-1|k-1}A^\top + F\Sigma_w F^\top; \end{aligned} \tag{3.54a}$$

and for the correction step:

$$\begin{aligned} \hat{x}_{k|k} = &\quad \hat{x}_{k|k-1} + L_k(y_k - C\hat{x}_{k|k-1}), \\ P_{k|k} = &\quad P_{k|k-1} - AP_{k|k-1}C^\top(C^\top P_{k|k-1}C^\top + \Sigma_v)^{-1}CP_{k|k-1}A^\top. \end{aligned} \tag{3.54b}$$

Once these estimates are accessible to the MPC controller, we set $\bar{x}_{0|k} = \hat{x}_{k|k}$ and $\Sigma_{\zeta_k} = P_{k|k}$ and run the moment-based MPC algorithm. Next we construct the cost function

$J_{O,A}(\tilde{\xi}_k^O)$ from the stochastic output predictions generated via the system dynamics Σ^{add} similar to the case in Equation (3.2a), i.e.,

$$J_{O,A}(\tilde{\xi}_k^O) = \sum_{j=0}^{N_p-1} (y_{j|k} - r_{j|k})^\top Q(y_{j|k} - r_{j|k}) + u_{j|k}^\top R u_{j|k}, \tag{3.55}$$

where $Q \succeq 0, R \succ 0$ and

$$\tilde{\xi}_k^{O^\top} = \begin{bmatrix} \bar{x}_{0|k}^\top & u_{[0,N_p-1]|k}^\top & \zeta_k^\top & w_{[0,N_p-1]|k}^\top & v_{[0,N_p-1]|k}^\top & r_{[0,N_p-1]|k}^\top \end{bmatrix}.$$

Similar to the previous moment-based MPC cases, we show that an MPC problem can be formulated by using nominal dynamics Σ^{nom} which is equivalent[9] to the moment-based robust counterparts of $J_{O,A}\tilde{\xi}_k^O$. We first define the deterministic cost functions as follows;

$$\bar{J}_{O,A}^M(\xi_k^O) = \quad \xi_k^{O^\top} H_{O,A}^M \xi_k^O, \tag{3.56a}$$

$$\bar{J}_{O,A}^{MV}(\xi_k^O) = \quad \xi_k^{O^\top} H_{O,A}^{MV} \xi_k^O, \tag{3.56b}$$

$$\bar{J}_{O,A}^{MVS}(\xi_k^O) = \quad \xi_k^{O^\top} H_{O,A}^{MVS} \xi_k^O, \tag{3.56c}$$

where

$$\xi_k^O = \begin{bmatrix} \bar{x}_{0|k} & u_{[0,N_p-1]|k} & r_{[0,N_p-1]|k} \end{bmatrix} \tag{3.56d}$$

$$H_{O,A}^M = \begin{bmatrix} \Theta_C^\top \bar{Q} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q} \Theta_C & T_{B,C}^\top \bar{Q} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p^\top} \bar{Q} \Theta_C & -\bar{I}_{n_y}^{N_p^\top} \bar{Q} T_{B,C} & \bar{I}_{n_y}^{N_p^\top} \bar{Q} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.56e}$$

$$H_{O,A}^{MV} = \begin{bmatrix} \Theta_C^\top \bar{Q}_{O,A}^{MV} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_{O,A}^{MV} \Theta_{N_p,C} & T_{B,C}^\top \bar{Q}_{O,A}^{MV} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p^\top} \bar{Q}_{O,A}^{MV} \Theta_C & -\bar{I}_{n_y}^{N_p^\top} \bar{Q}_{O,A}^{MV} T_{B,C} & \bar{I}_{n_y}^{N_p^\top} \bar{Q}_{O,A}^{MV} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.56f}$$

$$H_{O,A}^{MVS} = \begin{bmatrix} \Theta_C^\top Q, \bar{}\, A_O^{MVS} \Theta_C & (\star) & (\star) \\ T_{B,C}^\top \bar{Q}_{O,A}^{MVS} \Theta_C & T_{B,C}^\top \bar{Q}_{O,A}^{MVS} T_{B,C} + \bar{R} & (\star) \\ -\bar{I}_{n_y}^{N_p^\top} \bar{Q}_{O,A}^{MVS} \Theta_C & -\bar{I}_{n_y}^{N_p^\top} \bar{Q}_{O,A}^{MVS} T_{B,C} & \bar{I}_{n_y}^{N_p^\top} \bar{Q}_{O,A}^{MVS} \bar{I}_{n_y}^{N_p} \end{bmatrix}, \tag{3.56g}$$

$$\begin{aligned} \bar{Q}_{O,A}^{MV} &= \bar{Q} + 4\lambda_v \bar{Q} \bar{\Upsilon}_{O,A} \bar{Q}, \\ \bar{Q}_{O,A}^{MVS} &= \bar{Q}_{O,A}^{MV} + 24\lambda_s \bar{Q} \bar{\Upsilon}_{O,A} \bar{Q} \bar{\Upsilon}_{O,A} \bar{Q}, \\ \bar{\Upsilon}_{O,A} &:= T_{F,C} \bar{\Sigma}_w^{N_p} T_{F,C}^\top + \bar{\Sigma}_v^{N_p} + \Theta_C \Sigma_{\zeta_k} \Theta_C^\top, \end{aligned} \tag{3.56h}$$

The overall effect of the uncertainties w_k, v_k and the initial condition mismatch ζ_k, all of which are independent from each other over time, is incorporated to the moment-based MPC control structures as in Theorem 3.3.2.

[9] In the sense that the optimal input actions are equal to each other.

Theorem 3.3.2 *Given an LTI dynamical system Σ^{add} as in Equation* (3.19b)*, a reference trajectory $r_k, k \in \mathbb{Z}_{\geq 0}$, and consider the output feedback moment-based (Mean, MV and MVS) MPC problems with cost function $J_{O,A}(\tilde{\xi}_k^O)$ as in Equation* (3.55)*. Assume that a Kalman filter is designed for the system Σ^{add} which results in the update relations as in Equations* (3.54a)-(3.54b)*. Then;*

- *The cost function for Mean MPC problem with the stochastic cost function $J_{O,A}(\tilde{\xi}_k^O)$, i.e., $J_{O,A}^M(\xi_k^O)$, satisfy*

$$\begin{aligned} J_{O,A}^M(\xi_k^O) = &\quad \mathbb{E}\{J_{O,A}(\tilde{\xi}_k^O)\}, \\ = &\quad {\xi_k^O}^\top H_{O,A}^M \xi_k^O + f_{O,A}^M, \end{aligned} \tag{3.57a}$$

Furthermore the $f_{O,A}^M$ term is constant and

$$\bar{J}_{O,A}^M(\xi_k^O) = \quad J_{O,A}^M(\xi_k^O) - f_{O,A}^M, \tag{3.57b}$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} J_{O,A}^M(\xi_k^O) = &\quad \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,A}^M(\xi_k^O) + f_{O,A}^M, \\ \arg\min_{u_{[0,N_p-1]|k}} J_{O,A}^M(\xi_k^O) = &\quad \arg\min_{u_{[0,N_p-1]|k}} \bar{J}_{O,A}^M(\xi_k^O). \end{aligned} \tag{3.57c}$$

- *The cost function for MV MPC problem with the stochastic cost function $J_{O,A}(\tilde{\xi}_k^O)$, i.e., $J_{O,A}^{MV}(\xi_k^O)$, satisfy*

$$\begin{aligned} J_{O,A}^{MV}(\xi_k^O) = &\quad \mathbb{E}\{J_{O,A}(\tilde{\xi}_k^O)\} + \lambda_v \mathbb{D}\{J_{O,A}(\tilde{\xi}_k^O)\}, \\ = &\quad {\xi_k^O}^\top H_{O,A}^{MV} \xi_k^O + f_{O,A}^M + \lambda_v f_{O,A}^V, \end{aligned} \tag{3.58a}$$

Furthermore,

$$\bar{J}_{O,A}^{MV}(\xi_k^O) = \quad J_{O,A}^{MV}(\xi_k^O) - f_{O,A}^M - \lambda_v f_{O,A}^V, \tag{3.58b}$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} J_{O,A}^{MV}(\xi_k^O) = &\quad \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,A}^{MV}(\xi_k^O) + f_{O,A}^M + \lambda_v f_{O,A}^V, \\ \arg\min_{u_{[0,N_p-1]|k}} J_{O,A}^{MV}(\xi_k^O) = &\quad \arg\min_{u_{[0,N_p-1]|k}} \bar{J}_{O,A}^{MV}(\xi_k^O). \end{aligned} \tag{3.58c}$$

- *The cost function for MVS MPC problem with the stochastic cost function $J_{O,A}(\tilde{\xi}_k^O)$, i.e., $J_{O,A}^{MVS}(\xi_k^O)$, satisfy*

$$\begin{aligned} J_{O,A}^{MVS}(\xi_k^O) = &\quad \mathbb{E}\{J_{O,A}(\tilde{\xi}_O)\} + \lambda_v \mathbb{D}\{J_{O,A}(\tilde{\xi}_k^O)\} + \lambda_s \mathbb{S}\{J_{O,A}(\tilde{\xi}_k^O)\}, \\ = &\quad {\xi_k^O}^\top H_{O,A}^{MVS} \xi_k^O + f_{O,A}^M + \lambda_v f_{O,A}^V + \lambda_s f_{O,A}^S, \end{aligned} \tag{3.59a}$$

Furthermore,

$$\bar{J}_{O,A}^{MVS}(\xi_k^O) = \quad J_{O,A}^{MVS}(\xi_k^O) - f_{O,A}^M - \lambda_v f_{O,A}^V - \lambda_s f_{O,A}^S, \tag{3.59b}$$

and in particular

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} J_{O,A}^{MVS}(\xi_k^O) = & \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,A}^{MVS}(\xi_k^O) + f_{O,A}^M + \lambda_v f_{O,A}^V + \lambda_s f_{O,A}^S, \\ \arg \min_{u_{[0,N_p-1]|k}} J_{O,A}^{MVS}(\xi_k^O) = & \arg \min_{u_{[0,N_p-1]|k}} \bar{J}_{O,A}^{MVS}(\xi_k^O). \end{aligned} \tag{3.59c}$$

Proof. First observe that a Kalman filter operating with outputs of an unconstrained linear system leads to an observer gain L_k, hence one can make use of separation principle for observer and controller design. Next, we denote the outputs of the Kalman filter $\hat{x}_{k|k}$ and $P_{k|k}$, in which, by definition, $x_k = \hat{x}_{k|k} + \zeta_k$, leading to initial condition mismatch as in Equation (3.48) and the initial condition error characteristics given as $\zeta_k \sim \mathcal{N}(0, \Sigma_{\zeta_k})$. Set $\hat{x}_k$ as the initial condition of the MPC routine, i.e., $\bar{x}_{0|k}$ with the system Σ^{nom} and use the results in Lemma 3.1.4 and 3.3.1 with the assumption that the uncertain effects are independent from each other. □

Now we state the full moment-based MPC problem for Mean, MV or MVS MPC cases. We denote the MPC problem with $\mathcal{P}^{(\cdot)}$, where $(\cdot)$ can be any of M, MV or MVS.

$$\mathcal{P}_{O,A}^{(\cdot)} : \begin{cases} \min\limits_{u_{[0,N_p-1]|k}} & \bar{J}_{O,A}^{(\cdot)}(\xi_k^O), \\ \text{subject to} & \bar{x}_{i+1|k} = A\bar{x}_{i|k} + Bu_{i|k}, \\ & \bar{y}_{i|k} = C\bar{x}_{i|k}, i \in \mathbb{Z}_{[0,N_p-1]}, \\ & \bar{x}_{0|k} = \hat{x}_{k|k}, \Sigma_{\zeta_k} = P_{k|k}; \end{cases} \tag{3.60}$$

where the optimal solution for all of the problems is denoted with $u^*_{[0,N_p-1]|k}$. Furthermore for these cost functions of moment-based MPC formulations, one can apply Algorithm 1 for the output feedback MPC control law calculation.

Remark 3.3.1 *Even though the value of ζ_k, the deviation of estimated state from the true state, depends on instantaneous realizations of w_j and v_j, $j \in \mathbb{Z}_{[0,k]}$, Assumption 3.1.2 leads to independence of ζ_k from $w_{j|k}$ over prediction stages $j \in \mathbb{Z}_{[0,N_p-1]}$.*

3.4 Simulation Examples for Moment-based MPC with Perturbations

3.4.1 Regulation Problem

We apply the moment MPC technique to an example from [295]. The two state system with the control action and disturbances effective on both of the states is given as,

$$\begin{aligned} x_{k+1} = & \quad Ax_k + Bu_k + w_k, \\ A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, & \quad B = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}, \end{aligned}$$

where $w_k \sim \mathcal{N}(0, 0.5)$, i.i.d. over the time index k. We present the trajectories initiated from two different initial conditions for M-,MV- and MVS-MPC, comparing the closed-loop performances, in the first case the disturbances are set to zero and in the second case

Algorithm 1 Output Feedback Moment-based MPC Algorithm

Input: A system Σ^{add} with known covariance matrices Σ_w, Σ_v together with an initial state guess $\hat{x}_0$ and initial guess of covariance matrix Σ_{ζ_0}. Set $k = 0$, $\bar{x}_{0|0} = \hat{x}_{0|0}$ and $P_{0|0} = \Sigma_{\zeta_0}$.

1: **procedure** RECEDING HORIZON OUTPUT FEEDBACK MOMENT MPC

2: Solve the corresponding moment-based MPC problem in $\mathcal{P}^{(\cdot)}$ in Equations (3.60)according to the cost functions given in Theorem 3.3.2;

3: Apply the first term in the found optimal control sequence to the plant Σ^{add}, proceed to time $\hat{k} = k + 1$ and measure the new output $y_{\hat{k}}$;

4: Evaluate the Kalman updates, with the gain matrix $L_{\hat{k}|k} := P_{\hat{k}|k}C^\top(C^\top P_{\hat{k}|k}C + \Sigma_v)^{-1}$;

5: i) Prediction step:

$$\begin{aligned} \hat{x}_{\hat{k}|k} &= A\hat{x}_{k|k} + Bu^*_{k|k}, \\ P_{\hat{k}|k} &= AP_{k|k}A^\top + F\Sigma_w F^\top; \end{aligned}$$

ii) Correction step:

$$\begin{aligned} \hat{x}_{\hat{k}|\hat{k}} &= \hat{x}_{\hat{k}|k} + L_{\hat{k}|k}(y_{\hat{k}} - C\hat{x}_{\hat{k}|k}), \\ P_{\hat{k}|\hat{k}} &= P_{\hat{k}|k} - AP_{\hat{k}|k}C^\top(C^\top P_{\hat{k}|k}C^\top + \Sigma_v)^{-1}CP_{\hat{k}|k}A^\top; \end{aligned}$$

6: Set $\bar{x}_{0|\hat{k}} = \hat{x}_{\hat{k}|\hat{k}}$ and $\Sigma_{\zeta_{\hat{k}}} = P_{\hat{k}|\hat{k}}$.

7: Set $k = \hat{k}$ and return to Step 1.

non-zero disturbances are assumed in simulation. The prediction horizon is taken as 10 steps, the nominal weights Q and R are taken as

$$Q = \begin{bmatrix} \frac{1}{70} & 0 \\ 0 & \frac{1}{12} \end{bmatrix}, \quad R = 0.5$$

and for all of the simulations we omit Q_f and K_f terms. Lastly the λ_v and λ_s values are taken as 10 and 1000, respectively. Under this construction, the simulation results are visualized in Figure 3.4. The results are in coherence with the expectations, the trajectories corresponding to different moment MPC formulations are differing due to the aggressiveness of the high moment MPC controllers.

3.4.2 Tracking Problem

We consider a quadruple-tank process ([171]) to demonstrate the effective closed-loop operation of moment-based MPC constructions. The quadruple-tank system consists of four tanks where liquid is poured into these tanks via two actuators. The first actuator pours liquid into Tank 1 and Tank 4, while the second actuator is filling the Tank 2 and Tank 3. Furthermore the tanks are interacting with each other, i.e., the upper tanks (Tank 3 and Tank 4) are pouring liquid into the lower tanks (Tank 1 and Tank 2, respectively). We consider the states of the process as the liquid levels in the tanks and the liquid levels at Tank 1 and Tank

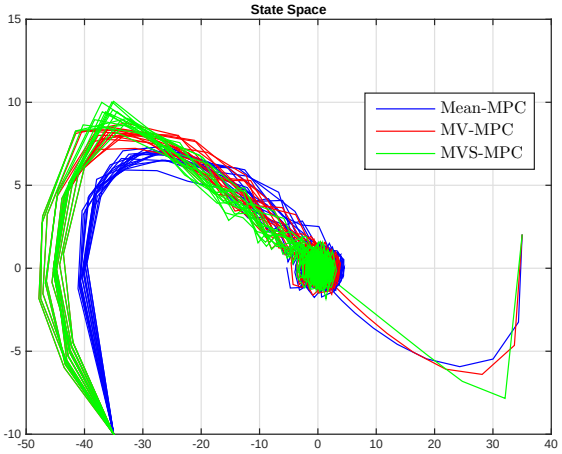


Figure 3.4: State trajectories for different formulations of moment-based MPC.

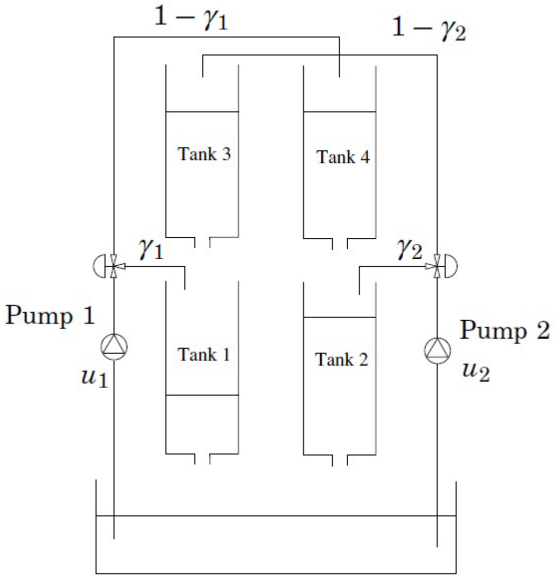


Figure 3.5: The four tank system interconnection scheme.

2 are tracked, see Figure 3.5. Hence the process can be described by four states, two inputs and two outputs. Furthermore we assume that the process is operating around a linearization

point ([350]), which yields the linearized dynamics with the uncertain perturbations;

$$\delta\dot{x}(t) = \begin{bmatrix} -\frac{1}{T_1} & 0 & \frac{A_4}{A_1T_3} & 0 \\ 0 & -\frac{1}{T_2} & 0 & \frac{A_4}{A_2T_4} \\ 0 & 0 & -\frac{1}{T_3} & 0 \\ 0 & 0 & 0 & \frac{1}{T_4} \end{bmatrix} \delta x(t) + \begin{bmatrix} \frac{\gamma_1 k_1}{A_1} & 0 \\ 0 & \frac{\gamma_2 k_2}{A_2} \\ 0 & \frac{(1-\gamma_2)k_2}{A_3} \\ \frac{(1-\gamma_1)k_1}{A_4} & 0 \end{bmatrix} \delta u(t) + w(t),$$

$$\delta y(t) = \begin{bmatrix} k_c & 0 & 0 & 0 \\ 0 & k_c & 0 & 0 \end{bmatrix} \delta x(t) + v(t).$$

The parameter and operating point values are provided in Table 3.2 and Table 3.3. Then various MPC controllers are implemented to compare the resulting trajectories. We compare mean (hence nominal), MV, MVS MPC formulations, with Gaussian pdf characteristics. Furthermore, in order to demonstrate the differences between MPC formulations considering different random variables, we also incorporate the MV MPC with uniform distribution, MV MPC with triangular distribution and MV MPC in output feedback formulation (with Kalman filter). The parameters used in the MPC construction are summarized in Table 3.4.

Parameters	Values	Units
A_1, A_2, A_3, A_4	28, 28, 32, 32	$[cm^2]$
$a_1, a_2, a_3 a_4$	0.071, 0.071, 0.057, 0.057	$[cm^2]$
k_1, k_2	3.33, 3.35	$[\frac{cm^2}{Vs}]$
k_c	0.5	$[\frac{V}{cm}]$
g	981	$[\frac{cm}{s^2}]$
γ_1, γ_2	0.25, 0.35	-
$T_i = \frac{A_i}{a_i}\sqrt{\frac{2x_i^*}{g}}$	-	-

Table 3.2: The parameter values of the quadruple-tank system.

Parameters	Values	Units
x_1^*, x_2^*	8.2444, 19.0163	$[cm]$
x_3^*, x_4^*	4.3146, 8.8065	$[cm]$
u_1^*, u_2^*	3, 3	$[V]$

Table 3.3: The selected operating point for the quadruple tank system.

The simulation results are visualized in Figure 3.6 and Figure 3.7. In Figure 3.6 we visualize the zero disturbance case, where the additive perturbations are set to zero. As expected, the MVS MPC (Gaussian) has the fastest response compared to MV MPC (Gaussian) and mean MPC. This is due to the extra terms effecting the output weighting matrix $\bar{Q}^{MVS}$. Furthermore, in comparison to the uniform and triangular cases, the Gaussian pdf

Parameters	Meaning	Value
N_p	Prediction horizon	50
Q	Output error weighting	$50I_2$
R	Input weighting	$0.1\ I_2$
Σ_w	State covariance matrix	$0.01\ I_4$
Σ_v	Output covariance matrix	$0.01\ I_2$
λ_v	Variance weight	0.02
λ_s	Skewness weight	0.03
$\lambda_{v,U}$	Variance weight (uniform)	0.01
$\lambda_{v,T}$	Variance weight (triangular)	0.01

Table 3.4: Parameters used in various moment-based MPC formulations.

MV-MPC provides the fastest convergence towards the reference values. In Figure 3.7a we visualize ten different simulations where the additive perturbations (with Gaussian pdfs) are present. Again, the MVS MPC has the best results compared to MV MPC and mean MPC, the variance of the outputs are also suppressed more than the other two cases. However this observation is not correct for any choice of variance and skewness weights, i.e., λ_v and λ_s. As one increases these weights, the closed-loop trajectories first slow down in convergence towards the reference and then the system becomes unstable. Further analysis on the selection of the moment weights are left out of this study and will be reported in future works. In Figure 3.7b we visualize five different simulation scenarios where the MV-MPC constructions are compared for tracking a sine wave. If one compares the results for the state feedback MV MPC with the output feedback MV MPC, the effect of not measuring the states is observable in the trajectories, the overall performance deteriorates, while the effect of initial condition covariance matrix suppresses this deterioration. Furthermore the MV MPC constructions that consider triangular or uniform distribution functions on the nature of uncertainties are leading to slight deterioration in the closed-loop performance. This drawback is due to higher variance (and third central moment) values of the non-Gaussian pdf characteristics, which penalizes the mean MPC behaviour much more than the MV MPC constructed for Gaussian disturbances.

3.5 Conclusions on Moment-based MPC Problems for Linear Systems with Additive Perturbations

In this chapter, we have presented a novel MPC algorithm which is based on the statistics (centralized moments) of the cost functionals for linear stochastic systems with additive uncertainty. We introduce and evaluate the mean, mean-variance, mean-variance-skewness MPC problems for state regulation and output tracking configuration. For the state regulation problem and all three MPC formulations, we provide the algebraic conditions on the stability of the closed-loop system. We, then, extend the discussion towards non-Gaussian disturbances and report the mean, MV and MVS MPC problems for even pdfs. Lastly, we discuss the effects of output feedback MPC configuration, thus assuming that the current

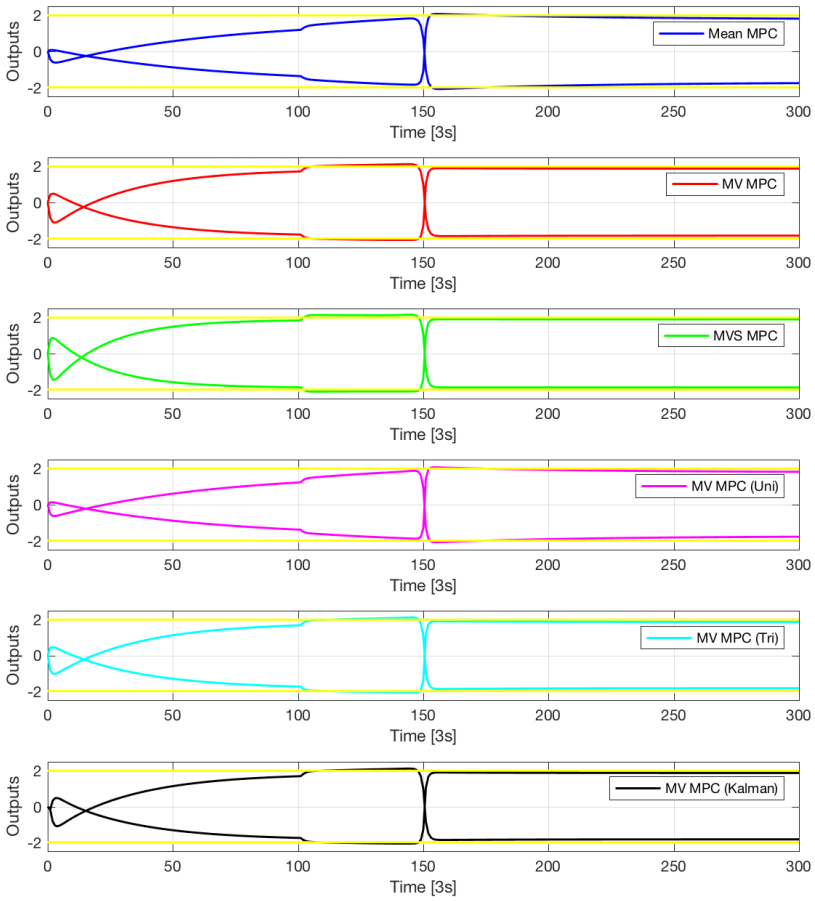
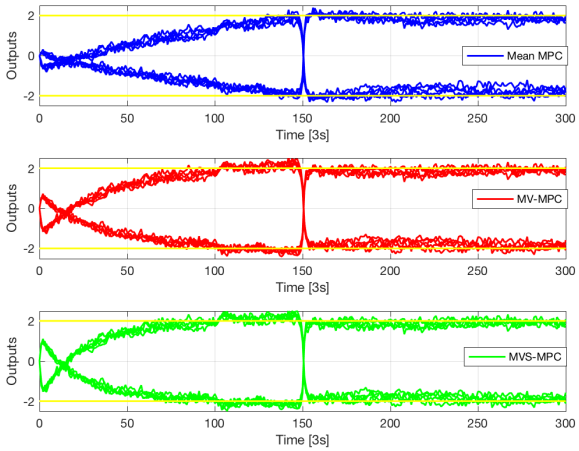


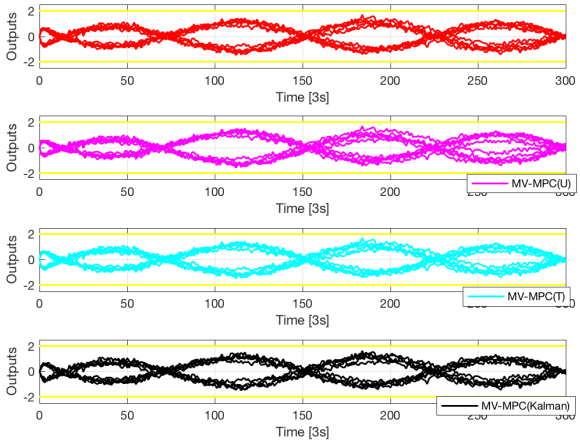
Figure 3.6: The output reference tracking with different moment-based MPC formulations for no disturbance case.

state is not fully accessible to the controller, for which we have shown that this stochastic MPC formulation reduces to the additive moment-based MPC formulations via incorporating the initial condition uncertainty (mismatch) as a new external additive perturbation occurring only at the current time instant.

Brief conclusions on the moment-based MPC for linear systems with additive perturbations can be stated as;



(a) The output reference tracking with different moment-based MPC formulations with disturbances case.



(b) The sine wave tracking with different MV-MPC formulations.

Figure 3.7: Simulations with disturbances.

- The mean-MPC problem for linear systems with additive uncertainties is equivalent with nominal MPC, meaning that MPC problems that are formulated without incorporating any uncertain effects are on average robust to the perturbations.
- The MV and MVS MPC can be expressed in terms of nominal dynamics, leading to stage-wise varying state weighting matrices over the prediction horizon. This de-

creases the computational complexity drastically, in comparison to the other robust MPC techniques.

- For additive perturbations case, one only needs to design a nominal MPC problem with varying weights to improve the disturbance rejection properties of the closed-loop system.
- In MV (or equivalently MVS) MPC formulation, the state (equivalently output) weights are affected by the noise characteristics, the moments of the random variables, and their evolution in the dynamical process within the prediction horizon.
- Modelling the distribution function of the random variables affecting the process has important consequences in moment-based MPC formulations. As the spread of the realizations gets larger, leading to larger variances, the controller becomes more and more aggressive to suppress the disturbances.

Chapter 4

Moment-based Model Predictive Control Problem with Plant-Prediction Model Mismatch

Think geometrically, prove algebraically.

John Torrence Tate - Rational Points on Elliptic Curves

In this chapter, we present a model predictive control (MPC) strategy based on the moments of the state variables and the cost functional constructed from prediction models described with uncertain parameters. The statistical properties of the state predictions are calculated through the open loop iteration of uncertain dynamics and used in the formulation of the MPC cost function. Here we treat the cases where the uncertain model parameters are unknown and time-invariant or independently varying over time iterations. We show that the moment-based MPC formulation yields predictive control problems that are described by nominal dynamics, hence computationally tractable compared to the existing robust MPC formulations, while providing statistical robustness properties. Due to combinatorial nature of the the robust counterpart cost function, we provide two approximation techniques, which are also shown to be acting similar to the discussed theoretical results. We apply the proposed MPC technique to a simple simulation example to demonstrate its effectiveness.

In Section 4.1 we present the uncertain models that are considered in moment-based MPC formulation and in Section 4.2 we present the moment MPC results for LTI systems with multiplicative uncertain factors. We provide the requirements for robust stability property in this section. Then in Section 4.3 we present the results for the moment MPC problems for dynamical systems with time-invariant multiplicative uncertain effects.

[0]Substantial content of this chapter is also published or presented in:

- M.B. Saltik, L. Özkan, S. Weiland. Moment-based Model Predictive Control for Linear Systems Part 2: Multiplicative Perturbations Case.
- R. Zhang, M.B. Saltik, L. Özkan, Study of Moment-based MPC formulations and their Connection to Classical Control.

4.1 Introduction on Dynamic and Static Model Mismatch

In this chapter, we consider linear systems under the effect of multiplicative uncertainty. We define a nominal system, denoted with Σ^{nom} given in state-space representation as in Equation (4.1a), and two systems effected by multiplicative uncertainty, the time-invariant (TI) uncertainty case and the time-varying (TV) uncertainty case, denoted with Σ^{mul}_{TI} and Σ^{mul}_{TV} as in Equation (4.1b) and (4.1c), respectively;

$$\Sigma^{nom} : \begin{cases} \bar{x}_{k+1} = & \bar{A}\bar{x}_k + \bar{B}u_k, \\ \bar{y}_k = & \bar{C}\bar{x}_k, \end{cases} \tag{4.1a}$$

$$\Sigma^{mul}_{TI} : \begin{cases} x_{k+1} = & A(\delta)x_k + B(\gamma)u_k, \\ y_k = & C(\delta)x_k, \end{cases} \tag{4.1b}$$

$$\Sigma^{mul}_{TV} : \begin{cases} x_{k+1} = & A(\delta_k)x_k + B(\gamma_k)u_k, \\ y_k = & C(\delta_k)x_k, \end{cases} \tag{4.1c}$$

where $x_k \in \mathbb{R}^n$, and $u_k \in \mathbb{R}^{n_u}$, and $y_k \in \mathbb{R}^{n_y}$ are the state and the control input (or the decision variable in the MPC problems) and the measured output at discrete time instant $k \in \mathbb{Z}_{\geq 0}$. These plants are subject to two types of uncertainties denoted by, for the case of TV uncertainties, δ_k and γ_k for $k \in \mathbb{Z}_{\geq 0}$. These uncertainties are describing the perturbations in the system matrices A and B which are stochastic vector sequences with known probability distribution functions (pdf), denoted with $f_{\tilde{\delta}}(\delta_k)$, for the random variable δ_k, and $f_{\tilde{\gamma}}(\gamma_k)$, for the random variable γ_k. The uncertain effects $\delta \in \mathbb{R}^{n_\delta}$ and $\gamma \in \mathbb{R}^{n_\gamma}$, for the TI case, are assumed to be effecting the system matrices affinely, i.e.,

$$A(\delta) = \bar{A} + \sum_{i=1}^{n_\delta} A_i[\delta]_i, \quad B(\gamma) = \bar{B} + \sum_{i=1}^{n_\gamma} B_i[\gamma]_i, \tag{4.2}$$

where, for brevity, we express the case for the time-invariant system dynamics Σ^M_{TI}, which also holds for the case of dynamics Σ^M_{TV}. Here the weighting matrices $A_{[\delta_k]_i}$ and $B_{[\gamma_k]_i}$ are introduced to scale the uncertainties such that one can obtain standard distributions for the random variables δ_k and γ_k. For the system Σ^{nom} the future predictions of state, at any time $k \in \mathbb{Z}_{\geq 0}$, can be expressed as

$$\bar{x}_{j|k} = A^j \bar{x}_{0|k} + \sum_{i=0}^{j-1} A^{j-1-i} B u_{i|k}, \tag{4.3a}$$

while the predictions from the uncertain dynamical systems can be expressed as,

$$\begin{aligned} x_{0|k} = & \ x_k, \\ x_{j|k} = & \ \prod_{i=0}^{j-1} A(\delta_{i|k}) x_{0|k} + \sum_{i=0}^{j-1} \prod_{l=i+1}^{j-1} A(\delta_l) B(\gamma_i) u_{i|k}, \qquad j \geq 1, \end{aligned} \tag{4.3b}$$

where $\prod_{i=l_1}^{l_2} = I$ if $l_1 > l_2$, $j \in \mathbb{Z}_{[0,N_p-1]}$ and $x_{j|k}$ denotes the prediction x_{k+j} parameterized with current state x_k, the exogenous signals $u_{i|k}, \delta_{i|k}$ and $\gamma_{i|k}, i \in \mathbb{Z}_{[0,j-1]}$ with respect to

system dynamics.
For the nominal system Σ^{nom} and for both of the multiplicative cases $\Sigma^{mul}_{(\cdot)}$, the matrices $(\bar{A}, \bar{B})$ are assumed to be a stabilizable pair.

Furthermore, the nature of uncertainties are assumed to be as follows;
(i) For the TI dynamics Σ^{mul}_{TI} case, the uncertain vectors δ and γ are assumed to be time invariant (such as identification mismatches) where $\delta \in \mathbb{R}^{n_\delta}$ and $\gamma \in \mathbb{R}^{n_\gamma}$ are Gaussian random vectors with zero mean and diagonal covariance matrices $\Sigma_\delta \in \mathbb{R}^{n_\delta \times n_\delta}$ or $\Sigma_\gamma \in \mathbb{R}^{n_\gamma \times n_\gamma}$, with the individual variance elements of the uncertain vectors are denoted as $\sigma^2_{[\delta]_i}$ for the variance of the element $[\delta]_i$, and $\sigma^2_{[\gamma]_i}$ for the variance of the element $[\gamma]_i$;
(ii) For the TV dynamics Σ^{mul}_{TV} case, the uncertain vectors $\delta_k \in \mathbb{R}^{n_\delta}$ and $\gamma_k \in \mathbb{R}^{n_\gamma}$ are i.i.d. Gaussian vectors with zero mean and diagonal covariance matrices $\Sigma_\delta \in \mathbb{R}^{n_\delta \times n_\delta}$ or $\Sigma_\gamma \in \mathbb{R}^{n_\gamma \times n_\gamma}$, with the same notation for the individual variance elements of the uncertain vectors as in the TI case.

Remark 4.1.1 *In here we assume that the uncertainties δ and γ which are attached to the system matrices $A(\delta)$ and $B(\gamma)$ are independent from each other, instead of $A(\delta)$ and $B(\delta)$. Furthermore, we also assumed that there is no covariance in between the elements of these vectors, i.e., Σ_δ and Σ_γ are diagonal covariance matrices. These assumptions are indeed not realistic in many cases. However the derivations that we present in here can easily be extended to incorporate these cases with extra, and tedious, algebra.*

4.2 Moment-based MPC for Time-Varying Multiplicative Uncertainty in Dynamics

In this section, we consider perturbations in A and B matrices that are varying independently over time, the dynamics represented with Σ^{mul}_{TV} in Equation (4.1c). Here the uncertainty in the model is defined as

$$A(\delta_k) = \bar{A} + \sum_{i=1}^{n_\delta} A_i [\delta_k]_i, \qquad B(\gamma_k) = \bar{B} + \sum_{i=1}^{n_\gamma} B_i [\gamma_k]_i.$$

In MPC problems we consider a cost function which is optimized over to decide on the control actions to be applied to the process. Assume that the time instant $k \in \mathbb{Z}_{\geq 0}$ is fixed. Then we set the cost function $J_M : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u N_p} \times \mathbb{R}^{n_\delta N_p} \times \mathbb{R}^{n_\gamma N_p} \to \mathbb{R}$ for a generic MPC problem as the sum of the quadratic terms running from current time instant k till $k + N_p - 1$ with a prediction horizon $N_p > 0$, i.e.,

$$J(\tilde{\xi}_k) = \sum_{j=0}^{N_p - 1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k}, \tag{4.4}$$

in which the uncertain states are predicted via dynamics Σ^{mul}_{TV} and

$$\tilde{\xi}_k = \begin{bmatrix} x_k^\top & u_{[0,N_p-1]|k}^\top & \delta_{[0,N_p-1]|k}^\top & \gamma_{[0,N_p-1]|k}^\top \end{bmatrix}^\top.$$

Using the parametrized state predictions given in Equation (5.2c), it can be easily seen that the cost function in Equation (4.4) is uncertain, thus a projection step is required to transform the uncertain cost into a deterministic cost function which is then used in the optimization procedure. The projection to be used is a design choice, hence highly effective on the resulting optimal control actions. In this section we analyze the linear combinations of (centralized) moments of the cost function $J(\tilde{\xi}_k)$ in Equation (4.4) and provide the closed form MPC problems with the robust counterpart cost functions. Next we start our discussion with Mean-MPC case, where the expectation operator is used to transform the uncertain cost function.

4.2.1 Mean MPC with Time-Varying Multiplicative Uncertainty in Dynamics

We evaluate the mean of the cost function $J_M(\tilde{\xi}_k)$, in Equation (4.4), i.e.,

$$\begin{aligned} J^M_{M,TV}(\xi_k) = &\quad \mathbb{E}\left\{ \xi_k^\top \begin{bmatrix} \tilde{\Theta}^\top \bar{Q} \tilde{\Theta} & \tilde{\Theta}^\top \bar{Q} \tilde{T}_B \\ \tilde{T}_B^\top \bar{Q} \tilde{\Theta} & \tilde{T}_B^\top \bar{Q} \tilde{T}_B + \bar{R} \end{bmatrix} \xi_k \right\}, \\ \xi_k^\top = &\quad \begin{bmatrix} x_{0|k}^\top & u_{[0,N_p-1]|k}^\top \end{bmatrix}. \end{aligned} \tag{4.5}$$

where $\tilde{\Theta}$, which is the shorthand of $\tilde{\Theta}_{N_p}$, and $\tilde{T}_B$, which is the shorthand of $\tilde{T}_B^{N_p \times N_p}$, are the prediction matrices given by

$$\begin{aligned} \tilde{\Theta}^\top = &\quad \begin{bmatrix} I & A^\top(\delta_0) & A^\top(\delta_0)A^\top(\delta_1) & \dots & \prod_{i=0}^{N_p-1} A^\top(\delta_i) \end{bmatrix}, \\ \tilde{T}_B = &\quad \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ B(\gamma_0) & 0 & \dots & 0 & 0 \\ A(\delta_1)B(\gamma_0) & B(\gamma_1) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \prod_{i=1}^{N_p-2} A(\delta_i)B(\gamma_0) & \prod_{i=2}^{N_p-2} A(\delta_i)B(\gamma_1) & \dots & B(\gamma_{N_p-2}) & 0 \end{bmatrix}, \end{aligned} \tag{4.6}$$

and lastly, $\bar{Q}$ and $\bar{R}$ are the state and input weighting matrices, respectively, which are constructed from the Q and R matrices stacked N_p times to the diagonal via the Kronecker product over matrices, i.e., $\bar{Q} = Q \otimes I_{N_p}$. The prediction matrices in Equation (4.6) are used to parameterize the state predictions within the prediction horizon as,

$$x_{[0,N_p-1]|k} = \tilde{\Theta} x_{0|k} + \tilde{T}_B u_{[0,N_p-1]|k}.$$

Under this reasoning, we state the cost function corresponding to the mean-MPC for linear systems affected by TV multiplicative uncertainties, $J^M_{M,TV}(\xi_k)$. For this purpose we use the following terms in the derivation of Lemma 4.2.1.

$$\begin{aligned} Q^{TV,M}_{N_p-1,M} = &\quad Q, \\ Q^{TV,M}_{N_p-1-j,M} = &\quad Q + \sum_{i=1}^{n_\delta} A_i^\top \Phi_{N_p-1-j} A_i \sigma^2_{[\delta]_i}, \\ j = &\quad 1, 2, \dots, N_p - 1, \end{aligned} \tag{4.7a}$$

$$\begin{aligned} R^{TV,M}_{N_p-1,M} &= R, \\ R^{TV,M}_{N_p-1-j,M} &= R + \sum_{i=1}^{n_\gamma} B_i^\top \Phi_{N_p-1-j} B_i \sigma^2_{[\gamma]_i}, \\ j &= 1, 2, \ldots, N_p - 1, \end{aligned} \tag{4.7b}$$

where

$$\Phi_{N_p-1-j} = \sum_{m=1}^{j} \phi_m(\bar{A}, A_1, \ldots, A_{n_\delta}), \;\; j \geq 1. \tag{4.8}$$

The terms Φ_{N_p-1-j} are summation of the combinatorial terms in Equation (4.8), i.e., $\phi_m(\bar{A}, A_1, \ldots, A_{n_\delta})$, which are defined as the matrix quadratic forms of order $2m$ consisting of all 2^m combinations of $\bar{A}$ and A_i, multiplied with Q and the corresponding variance terms σ_i^2, i.e.,

$$\begin{aligned} \phi_1(\bar{A}, A_1, \ldots, A_{n_\delta}) &= Q, \\ \phi_2(\bar{A}, A_1, \ldots, A_{n_\delta}) &= \bar{A}^\top Q \bar{A} + \sum_{l=1}^{n_\delta} A_l^\top Q A_l \sigma^2_{[\delta]_l}, \\ \phi_3(\bar{A}, A_1, \ldots, A_{n_\delta}) &= \bar{A}^{2^\top} Q \bar{A}^2 + \sum_{l=1}^{n_\delta} \Big(\bar{A}^\top A_l^\top Q A_l \bar{A} \sigma^2_{[\delta]_l} + \ldots \\ &\quad A_l^\top \bar{A}^\top Q \bar{A} A_l \sigma^2_{[\delta]_l} \Big) + \left(\sum_{l_1=1}^{n_\delta} \sum_{l_2=1}^{n_\delta} A_{l_1}^\top A_{l_2}^\top Q A_{l_2} A_{l_1} \sigma^2_{[\delta]_{l_1}} \sigma^2_{[\delta]_{l_2}} \right), \\ &= \sum_{l_2=1}^{n_\delta} \sum_{l_1=1}^{n_\delta} \rho^\top_{\bar{A},A_l,2} \tilde{Q}_2^{TV} \rho_{\bar{A},A_l,2}, \end{aligned} \tag{4.9a}$$

where $\rho^\top_{\bar{A},A_l,2} := \begin{bmatrix} (\bar{A}\bar{A})^\top & (\bar{A}A_{l_1})^\top & (A_{l_2}\bar{A})^\top & (A_{l_2}A_{l_1})^\top \end{bmatrix}$,

$$\tilde{Q}_2^{TV} = \begin{bmatrix} \frac{Q}{n_\delta^2} & 0 & 0 & 0 \\ 0 & \frac{Q\sigma^2_{[\delta]_{l_1}}}{n_\delta} & 0 & 0 \\ 0 & 0 & \frac{Q\sigma^2_{[\delta]_{l_2}}}{n_\delta} & 0 \\ 0 & 0 & 0 & Q\sigma^2_{[\delta]_{l_1}}\sigma^2_{[\delta]_{l_2}} \end{bmatrix}, \tag{4.9b}$$

and finally,

$$\phi_{j+1}(\bar{A}, A_1, \ldots, A_{n_\delta}) = \sum_{(l_1,l_2,\ldots,l_j)}^{n_\delta} \rho^\top_{\bar{A},A_l,j} \tilde{Q}_j^{TV} \rho_{\bar{A},A_l,j}, \tag{4.9c}$$

where $\rho_{\bar{A},A_l,j}$ is the block vector of matrices consisting of all combinations of $\bar{A}$ and A_l[1], and $\tilde{Q}_j^{TV}$ is a diagonal matrix multiplying the corresponding elements of $\rho_{\bar{A},A_l,j}$ with the

[1]Equation (4.9c) consists of n_δ distinct summations over all permutations of l_i.

variances values $\sigma^2_{[\delta]_l}$, i.e.,

$$\rho_{\bar{A},A_l,j} = \begin{bmatrix} \bar{A}^j \\ \bar{A}^{j-1}A_{l_1} \\ \bar{A}^{j-2}A_{l_2}\bar{A} \\ \vdots \\ \prod_{i=1}^{j} A_{l_i} \end{bmatrix}, \quad \tilde{Q}_j^{TV} = \begin{bmatrix} \frac{Q}{n_\delta^j} & 0 & 0 & \cdots & 0 \\ 0 & \frac{Q\sigma^2_{[\delta]_{l_1}}}{n_\delta^{j-1}} & 0 & \cdots & 0 \\ 0 & 0 & \frac{Q\sigma^2_{[\delta]_{l_2}}}{n_\delta^{j-1}} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdot & 0 & Q\prod_{i=1}^{j}\sigma^2_{[\delta]_{l_i}} \end{bmatrix}.$$

Lemma 4.2.1 *Consider an LTI stochastic dynamical system Σ^M_{TV} as in Equation* (4.1c) *and a cost function $J_M(\tilde{\xi}_k)$ as in Equation* (4.4)*. Furthermore assume that δ_k and γ_k are i.i.d. with zero mean and diagonal (uncorrelated) variance-covariance matrices. Then the mean MPC cost function $J^M_{M,TV}(\xi_k)$ is given as*

$$\begin{aligned} J^M_{M,TV}(\xi_k) = &\ \mathbb{E}\{J_M(\xi_k)\}, \\ = &\ \sum_{j=0}^{N_p-1} \bar{x}_{j|k}^\top Q_{j,M}^{TV,M}\bar{x}_{j|k} + u_{j|k}^\top R_{j,M}^{TV,M}u_{j|k}. \end{aligned} \tag{4.10}$$

Proof. For the sake of clarity, we assume that there is only one uncertain element[2] in both $A(\delta_k)$ and $B(\gamma_k)$, i.e., $A(\delta_k) = \bar{A} + A_1\delta_k$ and $B(\gamma_k) = \bar{B} + B_1\gamma_k$, $\delta_k, \gamma_k \in \mathbb{R}$, $k \in \mathbb{Z}_{\geq 0}$. We first express the mean of the cost function in quadratic matrix form, without evaluating the expectation as

$$\mathbb{E}\{J_M(\xi_k)\} = \mathbb{E}\left\{\xi_k^\top \begin{bmatrix} \tilde{\Theta}^\top\bar{Q}\tilde{\Theta} & \tilde{\Theta}^\top\bar{Q}\tilde{T}_B \\ (\star) & \tilde{T}_B^\top\bar{Q}\tilde{T}_B + \bar{R} \end{bmatrix}\xi_k\right\}. \tag{4.11}$$

Since the vector $\xi_k^\top = \begin{bmatrix} x_{0|k}^\top & u_{[0,N_p-1]|k}^\top \end{bmatrix}$ is a deterministic variable, one can express the Equation (4.11) as $\mathbb{E}\{J_M(\xi_k)\} = \xi_k^\top\mathbb{E}\{\tilde{H}\}\xi_k$, hence one only needs to express $\mathbb{E}\{\tilde{H}\}$ in terms of the weights Q and R, the nominal and uncertain system matrices and the variances of uncertainties. Next we construct a nominal MPC problem[3] from the system dynamics Σ^{nom} with stagewise varying weighting matrices $\bar{Q}_M^{TV,M}$, $\bar{R}_M^{TV,M}$ and equate it to the mean of the true cost function $\mathbb{E}\{J_M(\xi)\}$, i.e.,

$$\begin{aligned} \bar{Q}_M^{TV,M} = &\ \text{diag}\left(Q_{0,M}^{TV,M}, Q_{1,M}^{TV,M}, \ldots, Q_{N_p-1,M}^{TV,M}\right), \\ \bar{R}_M^{TV,M} = &\ \text{diag}\left(R_{0,M}^{TV,M}, R_{1,M}^{TV,M}, \ldots, R_{N_p-1,M}^{TV,M}\right), \\ J^M_{M,TV} = &\ \xi_k^\top \begin{bmatrix} \Theta^\top\bar{Q}_M^{TV,M}\Theta & \Theta^\top\bar{Q}_M^{TV,M}T_B \\ (\star) & T_B^\top\bar{Q}_M^{TV,M}T_B + \bar{R}_M^{TV,M} \end{bmatrix}\xi_k. \end{aligned} \tag{4.12}$$

[2]The high dimensional random vectors case follows similarly with the arguments presented below under the assumption of uncorrelated realizations along the time steps.

[3]Here we note that Θ and T_B are constructed from the nominal system dynamics Σ^{nom}, hence similar to the matrices in Equation (4.6) but contains only $\bar{A}$ and $\bar{B}$.

Then evaluate the expectation of the uncertain quadratic form in Equation (4.11) and equate it to the cost function in Equation (4.12) in blocks. We start from $\tilde{\Theta}^\top \bar{Q} \tilde{\Theta}$, i.e.,

$$\Theta^\top \bar{Q}_M^{TV,M} \Theta = \quad \mathbb{E}\{\tilde{\Theta}^\top \bar{Q} \tilde{\Theta}\}, \tag{4.13a}$$

where we have

$$\Theta \bar{Q}_M^{TV,M} \Theta = \quad Q_{0,M}^{TV,M} + \bar{A}^\top Q_{1,M}^{TV,M} \bar{A} + \cdots + \bar{A}^{N_p-1^\top} Q_{N_p-1,M}^{TV,M} \bar{A}^{N_p-1}. \tag{4.13b}$$

and we write the uncertain matrix $\tilde{\Theta}$ as

$$\begin{aligned} \tilde{\Theta} := & \begin{bmatrix} I \\ \bar{A} \\ \bar{A}^2 \\ \vdots \\ \bar{A}^{N_p-2} \\ \bar{A}^{N_p-1} \end{bmatrix} + \begin{bmatrix} 0 \\ A_1\delta_0 \\ A_1\delta_1\bar{A} \\ \vdots \\ A_1\delta_{N_p-3}\bar{A}^{N_p-3} \\ A_1\delta_{N_p-2}\bar{A}^{N_p-2} \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \prod_{i=0}^{N_p-2} A_1\delta_i \end{bmatrix}, \\ =: & \quad \tilde{\Theta}_0 + \tilde{\Theta}_1 + \cdots + \tilde{\Theta}_{2^{N_p-1}-1}, \end{aligned} \tag{4.13c}$$

with obvious definitions of $\tilde{\Theta}_i$. Then observe that

$$\mathbb{E}\{\tilde{\Theta}^\top \bar{Q} \tilde{\Theta}\} = \quad \mathbb{E}\left\{ \sum_{i=0}^{2^{N_p-1}-1} \tilde{\Theta}_i^\top \bar{Q} \tilde{\Theta}_i \right\}, \tag{4.13d}$$

since $\mathbb{E}\{\tilde{\Theta}_i \bar{Q} \tilde{\Theta}_j\} = 0$ for $i \neq j$. Thus only even order and symmetric terms remain after the expectation operator. Next, we observe that

$$\tilde{\Theta}_0^\top \bar{Q} \tilde{\Theta}_0 = Q + \bar{A}^\top Q \bar{A} + \cdots + \bar{A}^{N_p-1^\top} Q \bar{A}^{N_p-1}, \tag{4.14}$$

which leads to Q terms in all $Q_{j,M}^{TV,M}$ stage weight matrices, see Equation (4.7a). We evaluate the expectation for the remaining $\tilde{\Theta}_i$ terms, i.e.,

$$\mathbb{E}\{\tilde{\Theta}_1^\top \bar{Q} \tilde{\Theta}_1\} = A_1^\top Q A_1 \sigma_\delta^2 + \cdots + \bar{A}^{N_p-2^\top} A_1^\top Q A_1 \bar{A}^{N_p-2} \sigma_\delta^2, \tag{4.15}$$

which leads to $\phi_1(\bar{A}, A_1)$ terms in $Q_{j,M}^{TV,M}$ weighting matrices. Since there are combinatorial but finite number of $\tilde{\Theta}_i$ matrices, one can continue computing the expectations of $\tilde{\Theta}_i \bar{Q} \tilde{\Theta}_i$ to fill the $Q_{j,M}^{TV,M}$ matrices. In general, the expression $\phi_j(\bar{A}, A_1)$ is obtained[4] from $\sum_{i=2^{j-1}}^{2^j-1} \mathbb{E}\{\tilde{\Theta}_i \bar{Q} \tilde{\Theta}_i\}$. Now observe that the terms in Equation (4.14) and Equation (4.15) can be used to fill the $Q_{j,M}^{TV,M}$ in the Equation (4.13b). This results in the backwards recursive accumulation as in Equation (4.7a).

[4]The terms remaining after the expectation operations of $\mathbb{E}\{\tilde{\Theta}_2^\top \bar{Q} \tilde{\Theta}_2\}$ and $\mathbb{E}\{\tilde{\Theta}_3^\top \bar{Q} \tilde{\Theta}_3\}$ construct the $\phi_2(\bar{A}, A_1)$ term, for $\phi_3(\bar{A}, A_1)$ we make use of the expectations of the quadratic forms of $\tilde{\Theta}_i, i = 4, 5, 6, 7$.

Now consider the block $\mathbb{E}\{\tilde{T}_B^\top \bar{Q}\tilde{T}_B + \bar{R}\}$, for which we apply a similar reasoning. First we expand all the terms and evaluate the expectations term by term in $\mathbb{E}\{\tilde{T}_B^\top \bar{Q}\tilde{T}_B + \bar{R}\}$ and then group the terms induced from $Q_{j,M}^{TV,M}$, constructed in the previous step, and leave the rest to be expressed via $R_{j,M}^{TV,M}$. Observe that, by definition, $\tilde{T}_B$ is (column) concatenation of sums of $\tilde{\Theta}_i B$ matrices. Hence one can evaluate $\mathbb{E}\{\tilde{T}_B^\top \bar{Q}\tilde{T}_B + \bar{R}\}$ in short by

$$\mathbb{E}\{B(\gamma)^\top \tilde{\Theta}_j^\top \bar{Q}\tilde{\Theta}_i B(\gamma)\} = \bar{B}^\top \mathbb{E}\{\tilde{\Theta}_j^\top \bar{Q}\tilde{\Theta}_i\}\bar{B} + B_1^\top \mathbb{E}\{\tilde{\Theta}_j^\top \bar{Q}\tilde{\Theta}_i\}B_1\sigma_\gamma^2,$$

while the calculation of $\mathbb{E}\{\tilde{\Theta}_j^\top \bar{Q}\tilde{\Theta}_i\}$ follows from the $\mathbb{E}\{\tilde{\Theta}^\top \bar{Q}\tilde{\Theta}\}$. Hence similar to the expression for the stage-wise varying matrix $Q_{j,M}^{TV,M}$, the accumulation of $R_{j,M}^{TV,M}$ is obtained as in Equation (4.7b). Lastly one verifies whether the polynomial obtained from $\mathbb{E}\{\tilde{\Theta}^\top \bar{Q}\tilde{T}_B\}$ is equivalent with the polynomial obtained from $\Theta\bar{Q}_M^{TV,M}T_B$. By observing the fact that all the terms induced from $B_1\gamma_k$ are equal to zero after the expectation operation, there are no terms containing even orders of $B_1\gamma$ in $\mathbb{E}\{\tilde{\Theta}^\top \bar{Q}\tilde{T}_B\}$, thus $\bar{Q}_M^{TV,M}$ captures all the extra terms. □

A direct consequence of Lemma 4.2.1 is stated below.

Theorem 4.2.1 *Consider an LTI stochastic dynamical system Σ_{TV}^M as in Equation* (4.1c)*, a cost function $J_M(\tilde{\xi}_k)$ as in Equation* (4.4) *subject to uncertain dynamics Σ_{TV}^M and the cost function $J_{M,TV}^M(\xi_k)$ as in Equation* (4.10) *subject to nominal dynamics Σ^{nom}. Furthermore assume that δ_k and γ_k are i.i.d. with diagonal (uncorrelated) variance-covariance matrices. Then*

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} \mathbb{E}\{J_M(\tilde{\xi}_k)\} &= \min_{u_{[0,N_p-1]|k}} J_{M,TV}^M(\xi_k), \\ \arg\min_{u_{[0,N_p-1]|k}} \mathbb{E}\{J_M(\tilde{\xi}_k)\} &= \arg\min_{u_{[0,N_p-1]|k}} J_{M,TV}^M(\xi_k) \end{aligned} \tag{4.16}$$

Proof. With the results of Lemma 1, it is trivial to show that the claim is correct. □

According to the result of Theorem 1, the control signal $u_{[0,N_p-1|k]}^*$ defines an optimal solution for both of the MPC problems. The robust counterpart MPC problem with the cost function $J_{M,TV}^M(k)$ in Equation (4.10) and with decision variables $u_{j|k}$ is equivalent to the following MPC problem

$$\begin{aligned} u_{[0,N_p-1]|k}^* = \arg\min_{u_{[0,N_p-1]|k}} & \sum_{j=0}^{N_p-1} \bar{x}_{j|k}^\top Q_{j,M}^{TV,M} \bar{x}_{j|k} + u_{j|k}^\top R_{j,M}^{TV,M} u_{j|k}, \\ \text{subject to} \quad & \bar{x}_{i+1|k} = \bar{A}\bar{x}_{i|k} + \bar{B}u_{i|k}, \\ & \bar{x}_{0|k} = x_k, i \in \mathbb{Z}_{[0,N_p-1]}. \end{aligned} \tag{4.17}$$

Remark 4.2.1 *The most striking property of the moment-based MPC problem in Equation* (4.16) *is the reduced computational complexity of the MPC problem. Solving the Mean-MPC problem is of equivalent complexity as any nominal MPC problem constructed from the uncertainty-free dynamics Σ^{nom}. The cost function $J_{M,TV}^M$ as in Equation* (4.10) *depends on the nominal system matrices $\bar{A}, \bar{B}$ but also on the perturbations in the system*

matrices A_i, B_i and the corresponding variance terms of the uncertainties σ_δ^2 or σ_γ^2 which scale the weighting matrices $\bar{Q}_M^{M,TV}$ or $\bar{R}_M^{M,TV}$.

Remark 4.2.2 *The number of terms that accumulate in each of the $Q_{j,M}^{M,TV}$ diminishes as the prediction step grows. Specifically $Q_{j,M}^{M,TV} \geq Q_{j+1,M}^{M,TV}$ for all $j \geq 0$. This is in direct contrast to the additive disturbance cases, where the weighting terms (in mean-variance or MVS MPC) Q_j^{MV} are accumulating more and more as the prediction step j grows. Furthermore, in the additive disturbance case the input weighting matrix $\bar{R}$ does not change when one incorporates high order moments, whereas in the model mismatch case the diagonal elements of the matrix $\bar{R}_M^{M,TV}$ are also effected by the expectation operator and the uncertainties attached to the model.*

Remark 4.2.3 *For both of the weighting matrices $Q_{j,M}^{M,TV}$ and $R_{j,M}^{M,TV}$, the number of terms due to the uncertainties in the resulting MPC weights exponentially increases. Denote the number of terms in the expression for $\phi_j(\bar{A}, A_1, \ldots, A_{n_\delta})$ with $|\phi_j(\bar{A}, A_1, \ldots, A_{n_\delta})|$, which is, then, equal to $2^{j-1}n_\delta$. This is due to the fact that we only incorporate the symmetric even order combinations resulting from $A_i, i = 1, \ldots, n_\delta$. If one can approximate the terms in*

$$\begin{aligned} &\mathbb{E}\left\{ \prod_{m=N_p-j+1}^{N_p-2} A(\delta_k)^{m^\top} \bar{Q} A(\delta_k)^m \right\}, \\ &= \mathbb{E}\left\{ \sum_{i=1}^{n_\delta} \prod_{m=N_p-j+1}^{N_p-2} (\bar{A} + A_i[\delta]_i)^{m^\top} Q \prod_{m=N_p-j+1}^{N_p-2} (A + A_i[\delta]_i)^m \right\}. \end{aligned} \tag{4.18}$$

with a linearly growing expression, then the complexity can be kept tractable. Two possible approximations of the expression in Equation (4.18) *are given in Equation* (4.19a) *and Equation* (4.20a)*.*

The first approximation is given as

$$\begin{aligned} Q_{N_p-1,M}^{TV,M,Ap1} &= Q, \\ Q_{N_p-1-j,M}^{TV,M,Ap1} &= Q + \sum_{i=1}^{n_\delta} A_i^\top \Phi_{N_p-1-j}^{Ap1} A_i \sigma_{[\delta]_i}^2, \\ j &= 1, 2, \ldots, N_p - 1, \end{aligned} \tag{4.19a}$$

$$\begin{aligned} R_{N_p-1,M}^{TV,M,Ap1} &= R, \\ R_{N_p-1-j,M}^{TV,M,Ap1} &= R + \sum_{i=1}^{n_\gamma} B_i^\top \Phi_{N_p-1-j}^{Ap1} B_i \sigma_\gamma^2, \\ j &= 1, 2, \ldots, N_p - 1, \end{aligned} \tag{4.19b}$$

$$\Phi_{N_p-1-j}^{Ap1} = \sum_{m=1}^{j} \phi_m^{Ap1}(\bar{A}, A_1, \ldots, A_{n_\delta}), \;\; j \geq 1. \tag{4.19c}$$

$$
\begin{aligned}
\phi_1^{Ap1}(\bar{A}, A_1, \ldots, A_{n_\delta}) = &\quad Q,\\
\phi_2^{Ap1}(\bar{A}, A_1, \ldots, A_{n_\delta}) = &\quad \bar{A}^\top \phi_1^{Ap1}(\bar{A}, A_1, \ldots, A_{n_\delta})\bar{A} + \sum_{l=1}^{n_\delta} A_l^\top Q A_l \sigma^2_{[\delta]_l},\\
\phi_j^{Ap1}(\bar{A}, A_1, \ldots, A_{n_\delta}) = &\quad \bar{A}^\top \phi_{j-1}^{Ap1}(\bar{A}, A_1, \ldots, A_{n_\delta})\bar{A} + \sum_{l=1}^{n_\delta} {A_l^{j-1}}^\top Q A_l^{j-1} \sigma^{2j-2}_{[\delta]_l},
\end{aligned}
\tag{4.19d}
$$

while the second approximation amounts to replacing the additional terms at each prediction stage with quadratic forms in $\bar{A}$ instead of quadratic forms in A_i terms is given as,

$$
\begin{aligned}
Q^{TV,M,Ap2}_{N_p-1,M} = &\quad Q,\\
Q^{TV,M,Ap2}_{N_p-1-j,M} = &\quad Q + \sum_{i=1}^{n_\delta} A_i^\top \Phi^{Ap2}_{N_p-1-j} A_i \sigma^2_{[\delta]_i},\\
j = &\quad 1, 2, \ldots, N_p - 1,
\end{aligned}
\tag{4.20a}
$$

$$
\begin{aligned}
R^{TV,M,Ap2}_{N_p-1,M} = &\quad R,\\
R^{TV,M,Ap2}_{N_p-1-j,M} = &\quad R + \sum_{i=1}^{n_\gamma} B_i^\top \Phi^{Ap2}_{N_p-1-j} B_i \sigma^2_{[\gamma]_i},\\
j = &\quad 1, 2, \ldots, N_p - 1,
\end{aligned}
\tag{4.20b}
$$

$$
\Phi^{Ap2}_{N_p-1-j} = \quad \sum_{m=1}^{j} \phi_m^{Ap2}(\bar{A}, A_1, \ldots, A_{n_\delta}), \;\; j \geq 1. \tag{4.20c}
$$

$$
\begin{aligned}
\phi_1^{Ap2}(\bar{A}, A_1, \ldots, A_{n_\delta}) = &\quad Q,\\
\phi_2^{Ap2}(\bar{A}, A_1, \ldots, A_{n_\delta}) = &\quad \sum_{l=1}^{n_\delta} A_l^\top \phi_1^{Ap2}(\bar{A}, A_1, \ldots, A_{n_\delta}) A_l \sigma^2_{[\delta]_l} + \bar{A}^\top Q \bar{A},\\
\phi_j^{Ap2}(\bar{A}, A_1, \ldots, A_{n_\delta}) = &\quad \sum_{l=1}^{n_\delta} A_l^\top \phi_{j-1}^{Ap2}(\bar{A}, A_1, \ldots, A_{n_\delta}) A_l \sigma^2_{[\delta]_l} + {\bar{A}^{j-1}}^\top Q \bar{A}^{j-1}.
\end{aligned}
\tag{4.20d}
$$

Both of these approximations have (computationally) much more simple recursive formulas for calculating the weighting matrices $Q^{TV,M,Ap1}_{N_p-j,M}$ and $Q^{TV,M,Ap2}_{N_p-j,M}$ (equivalently $R^{TV,M,Ap2}_{N_p-j,M}$ and $R^{TV,M,Ap2}_{N_p-j,M}$) in comparison to the true weighting terms from the expectation operation, i.e., $Q^{TV,M}_{N_p-j,M}$. To select one of the approximations, the operator norms of $\bar{A}$ and $A_l, l = 1, 2, \ldots, n_\delta$ can be compared. If the uncertainties are dominant in the dynamical model Σ^M_{TV}, i.e., there exists an $l \in \mathbb{Z}_{[0,n_\delta]}$ such that $\|A_l\|_{2,2} \approx \|\bar{A}\|_{2,2}$, where $\|\bar{A}\|_{2,2}$ is the induced 2-norm of the matrix $\bar{A}$, then one might select the first approximation case, Equations (4.19a)-(4.19b). This case incorporates the highest order term induced from A_l which is, under the above mentioned norm condition, effective in the resulting closed-loop state trajectories. If the control system designer has considerable confidence on the modeled nominal dynamics, yielding $\|A_l\|_{2,2} \ll \|\bar{A}\|_{2,2}$ for any $l = 1, 2, \ldots, n_\delta$, then one can improve the robustness properties of the close-loop system with the second approximation given in Equations (4.20a)-(4.20b).

4.2.2 High order Moment MPC for Time-Varying Multiplicative Uncertainty in Dynamics

Similar to the mean MPC formulation, we construct the mean-variance MPC which evaluates cost functions that include the second order central moment. Hence we re-shape our

cost function as

$$\bar{J}^{MV}_{M,TV}(\xi_k) = \quad \mathbb{E}\{J_M(\xi_k)\} + \lambda_v \mathbb{D}\{J_M(\xi_k)\}, \tag{4.21}$$

The first term in the right hand side of Equation (4.21) is already calculated, see the expression for $J^M_{M,TV}(\xi_k)$ in Equation (4.10). We obtain the $\mathbb{D}\{J_{M,TV}(\xi_k)\}$ term as,

$$\begin{aligned}
\mathbb{D}\{J_{M,TV}(\tilde{\xi}_k)\} = & \quad \mathbb{E}\{(J_M(\xi_k) - \mathbb{E}\{J_M(\xi_k)\})^\top (J_M(\xi_k) - \mathbb{E}\{J_M(\xi_k)\})\}, \\
= & \quad \sum_{j=0}^{N_p-1} \mathbb{E}\{(2\tilde{x}^\top_{j|k} Q \bar{x}_{j|k} + \tilde{x}^\top_{j|k} Q \tilde{x}_{j|k} - Tr(QM_{j|k})^2\}, \\
= & \quad 4\bar{x}^\top_{j|k} Q M_{j|k} Q \bar{x}_{j|k} + 2Tr\left(QM_{j|k} QM_{j|k}\right),
\end{aligned} \tag{4.22}$$

where $M_{j|k} = \mathbb{E}\{\tilde{x}_{j|k}\tilde{x}^\top_{j|k}\}$ and $\tilde{x}_{j|k} = x_{j|k} - \bar{x}_{j|k}$. Through this calculation it is observed that $\mathbb{D}\{J_{M,TV}(\xi_k)\}$ contains 4^{th} order terms of $\bar{x}_{j|k}$, hence leading to a nonlinear optimization problem. This is an undesired increase in the computational complexity for calculating the MPC control action. Hence we left the construction of MV-MPC for model mismatch case as a future research topic.

Remark 4.2.4 *Similar to the previous reasoning, the cost functions for higher order moment-based MPC are also including high order polynomial terms of decision variables in their formulation, i.e.,* $\mathbb{E}\{(J_{M,TV} - \mathbb{E}\{J_{M,TV}\})^m\}, m \geq 2$. *Hence, the nonlinear nature of the cost function of the moment-based MPC problem is preserved.*

4.3 Moment-based MPC for Time-Invariant Multiplicative Uncertainty in the Dynamics

Here we consider the model used for the predictions as

$$\Sigma^M_{TI} : \begin{cases} x_{k+1} = & A(\delta)x_k + B(\gamma)u_k, \\ y_k = & Cx_k, \end{cases} \tag{4.23}$$

where the uncertainty is defined as

$$A(\delta) = \bar{A} + \sum_{i=1}^{n_\delta} A_i[\delta]_i, \qquad B(\gamma) = \bar{B} + \sum_{i=1}^{n_\gamma} B_i[\gamma]_i.$$

The difference in the TI case from the TV case is that the random variables δ and γ are not varying over time, i.e., the δ and γ are independent random vectors, which are distributed with Gaussian characteristics with diagonal covariance matrices, but over time the realizations of δ and γ does not change. Similar to the previous case, we set the cost function for the MPC problem as in Equation (4.4),

$$J_M(\xi_k) = \sum_{j=0}^{N_p-1} x^\top_{j|k} Q x_{j|k} + u^\top_{j|k} R u_{j|k},$$

where $x_{j|k}$ is parameterized from the dynamics Σ^M_{TI}. Next we discuss the mean MPC cost function for the dynamics affected by the time-invariant uncertainties.

4.3.1 Mean MPC for Time-Invariant Multiplicative Uncertainty in Dynamics

Similar to the time-varying case discussed in the previous section, we consider the mean of the cost function as in;

$$J^M_{M,TI}(\xi_k) = \mathbb{E}\left\{\xi_k^\top \begin{bmatrix} \tilde{\Theta}^\top \bar{Q}\tilde{\Theta} & \tilde{\Theta}^\top \bar{Q}\tilde{T}_B \\ \tilde{T}_B^\top \bar{Q}\tilde{\Theta} & \tilde{T}_B^\top \bar{Q}\tilde{T}_B + \bar{R} \end{bmatrix} \xi_k \right\}, \tag{4.24}$$

where $\tilde{\Theta}$ and $\tilde{T}_B$ are the prediction matrices given by

$$\begin{aligned} \tilde{\Theta}_{N_p} = & \begin{bmatrix} I & A^\top(\delta) & A^{2^\top}(\delta) & \dots & A^{N_p-1^\top}(\delta) \end{bmatrix}^\top, \\ \tilde{T}_B = & \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ B(\gamma) & 0 & \dots & 0 & 0 \\ A(\delta)B(\gamma) & B(\gamma) & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ A^{N_p-2}(\delta)B(\gamma) & A^{N_p-3}(\delta)B(\gamma) & \dots & B(\gamma) & 0 \end{bmatrix}, \end{aligned} \tag{4.25}$$

These prediction matrices can be used to parameterize the state predictions within the prediction horizon as,

$$x_{[0,N_p-1]|k} = \tilde{\Theta} x_{0|k} + \tilde{T}_B u_{[0,N_p-1]|k}.$$

Similar to the TV case, we first define the matrices that will be used in the construction of $J^M_{M,TI}(\xi_k)$ before demonstrating the equivalence between moment-based MPC and nominal MPC.

$$\begin{aligned} Q^{TI,M}_{N_p-1,M} = & \quad Q, \\ Q^{TI,M}_{N_p-1-j,M} = & \quad Q + \sum_{i=1}^{n_\delta} A_i^\top \Psi^Q_{N_p-1-j} A_i, \\ j = & \quad 1, \dots N_p - 1, \end{aligned} \tag{4.26a}$$

$$\begin{aligned} R^{TI,M}_{N_p-1,M} = & \quad R, \\ R^{TI,M}_{N_p-1-j,M} = & \quad R + \sum_{i=1}^{n_\gamma} B_i^\top \Psi^R_{N_p-1-j} B_i \sigma^2_{[\gamma]_i}, \\ j = & \quad 1, \dots N_p - 1, \end{aligned} \tag{4.26b}$$

$$\begin{aligned} Q^{TI,M}_{N_p-1-j,N_p-1-l,M} = & \quad \sum_{i=1}^{n_\delta} A_i^\top \Psi^Q_{N_p-1-j,N_p-1-l} A_i, \\ R^{TI,M}_{N_p-1-j,N_p-1-l,M} = & \quad \sum_{i=1}^{n_\delta} B_i^\top \Psi^R_{N_p-1-j,N_p-1-l} B_i, \\ 1 \leq & \quad i < j \leq N_p - 1, \end{aligned} \tag{4.26c}$$

where

$$\begin{aligned} \Psi^Q_{N_p-1-j} = & \quad \sum_{m=1}^{j} \psi_m(\bar{A}, A_1, \dots, A_{n_\delta}, Q), \\ \Psi^R_{N_p-1-j} = & \quad \sum_{m=1}^{j} \psi_m(\bar{A}, A_1, \dots, A_{n_\delta}, R), \\ j = & \quad 1, \dots N_p - 1, \end{aligned} \tag{4.26d}$$

$$\begin{aligned}\Psi^{Q}_{N_p-1-j,N_p-1-l} = &\sum_{m=1}^{l}\psi_{j,m}(\bar{A},A_1,\ldots,A_{n_\delta},Q),\\ \Psi^{R}_{N_p-1-j,N_p-1-l} = &\sum_{m=1}^{l}\psi_{j,m}(\bar{A},A_1,\ldots,A_{n_\delta},R),\\ 1\leq\ & i<j\leq N_p-1,\end{aligned} \tag{4.26e}$$

and lastly the permutation forms of $\bar{A}$ and A_i expressed via the variable $\psi_m(\bar{A},A_1,\ldots,A_{n_\delta})$ and $\psi_{j,m}(\bar{A},A_1,\ldots,A_{n_\delta})$, are defined as the matrix quadratic functions of order $2m$ (in matrix variable $A(\delta)$) consisting of all even order permutations of A_l, multiplied with Q and the corresponding raw moment terms $\mathbb{E}\{[\delta]_i^m\}$ where m is the number of A_i terms in the term under consideration, i.e.,

$$\begin{aligned}\psi_1(\bar{A},A_1,\ldots,A_{n_\delta},Q) = &\sum_{l=1}^{n_\delta}Q\sigma^2_{[\delta]_l}=\rho^\top_{\bar{A},A_l,0}\tilde{Q}^{TI,Q}_{0,M}\rho_{\bar{A},A_l,0},\\ \psi_1(\bar{A},A_1,\ldots,A_{n_\delta},R) = &\ Q=\rho^\top_{\bar{A},A_l,0}\tilde{Q}^{TI,R}_{0,M}\rho_{\bar{A},A_l,0},\end{aligned} \tag{4.27a}$$

$$\begin{aligned}\psi_2(\bar{A},A_1,\ldots,A_{n_\delta},Q) = &\sum_{l=1}^{n_\delta}\bar{A}^\top Q\bar{A}\sigma^2_{[\delta]_l}+A_l^\top QA_l\mathbb{E}\{[\delta]_l^4\}=\rho^\top_{\bar{A},A_l,1}\tilde{Q}^{TI,Q}_{1,M}\rho_{\bar{A},A_l,1},\\ \psi_2(\bar{A},A_1,\ldots,A_{n_\delta},R) = &\sum_{l=1}^{n_\delta}\bar{A}^\top Q\bar{A}+A_l^\top QA_l\sigma^2_{[\delta]_l}=\rho^{TI^\top}_{\bar{A},A_l,1}\tilde{Q}^{TI,R}_{1}\rho^{TI}_{\bar{A},A_l,1},\end{aligned} \tag{4.27b}$$

$$\begin{aligned}\psi_3(\bar{A},A_1,\ldots,A_{n_\delta},Q) = &\sum_{l=1}^{n_\delta}\Big(\bar{A}^{2^\top}Q\bar{A}^2\sigma^2_{[\delta]_l}+\bar{A}^{2^\top}QA_l^2\mathbb{E}\{[\delta]_l^4\}+\ldots\\ &\bar{A}^\top A_l^\top Q\bar{A}A_l\mathbb{E}\{[\delta]_l^4\}+A_l^\top\bar{A}^\top Q\bar{A}A_l\mathbb{E}\{[\delta]_l^4\}+\ldots\\ &A_l^\top\bar{A}^\top QA_l\bar{A}\mathbb{E}\{[\delta]_l^4\}+\bar{A}^\top A_l^\top QA_l\bar{A}\mathbb{E}\{[\delta]_l^4\}+\ldots\\ &A_l^{2^\top}Q\bar{A}^2\mathbb{E}\{[\delta]_l^4\}+A_l^{2^\top}QA_l^2\mathbb{E}\{[\delta]_l^6\},\\ = &\sum_{l=1}^{n_\delta}\rho^\top_{\bar{A},A_l,2}\tilde{Q}^{TI,Q}_{2,M}\rho_{\bar{A},A_l,2},\end{aligned} \tag{4.27c}$$

$$\begin{aligned}\psi_j(\bar{A},A_1,\ldots,A_{n_\delta},Q) = &\sum_{l=1}^{n_\delta}\rho^\top_{\bar{A},A_l,j-1}\tilde{Q}^{TI,Q}_{j-1,M}\rho_{\bar{A},A_l,j-1},\\ \psi_j(\bar{A},A_1,\ldots,A_{n_\delta},R) = &\sum_{l=1}^{n_\delta}\rho^\top_{\bar{A},A_l,j-1}\tilde{Q}^{TI,R}_{j-1,M}\rho_{\bar{A},A_l,j-1},\end{aligned} \tag{4.27d}$$

and

$$\begin{aligned}\psi_{2,1}(\bar{A},A_1,\ldots,A_{n_\delta},Q) = &\sum_{l=1}^{n_\delta}\bar{A}^\top Q\sigma^2_{[\delta]_l}=\rho^\top_{\bar{A},A_l,1}\tilde{Q}^{TI,Q}_{1,0,M}\rho_{\bar{A},A_l,0},\\ \psi_{2,1}(\bar{A},A_1,\ldots,A_{n_\delta},R) = &\ \bar{A}^\top Q=\rho^\top_{\bar{A},A_l,1}\tilde{Q}^{TI,R}_{1,0,M}\rho_{\bar{A},A_l,0},\\ \psi_{3,1}(\bar{A},A_1,\ldots,A_{n_\delta},Q) = &\sum_{l=1}^{n_\delta}\bar{A}^{2^\top}Q\sigma^2_{[\delta]_l}+A_l^{2^\top}Q\mathbb{E}\{[\delta]_l^4\}=\rho^\top_{\bar{A},A_l,2}\tilde{Q}^{TI,Q}_{2,0,M}\rho_{\bar{A},A_l,0},\end{aligned} \tag{4.27e}$$

$$\begin{aligned}\psi_{3,2}(\bar{A},A_1,\ldots,A_{n_\delta},Q) = &\sum_{l=1}^{n_\delta}\bar{A}^{2^\top}Q\bar{A}\sigma^2_{[\delta]_l}+A_l^{2^\top}Q\bar{A}\mathbb{E}\{[\delta]_l^4\}+\ldots\\ &\bar{A}^\top A_l^\top QA_l^\top\mathbb{E}\{[\delta]_l^4\}+A_l^\top\bar{A}^\top QA_l^\top\mathbb{E}\{[\delta]_l^4\}\\ = &\ \rho^\top_{\bar{A},A_l,2}\tilde{Q}^{TI,Q}_{2,1,M}\rho_{\bar{A},A_l,1},\\ \psi_{j,i}(\bar{A},A_1,\ldots,A_{n_\delta},Q) = &\ \rho^\top_{\bar{A},A_l,j-1}\tilde{Q}^{TI,Q}_{j-1,i-1,M}\rho_{\bar{A},A_l,i-1}.\end{aligned} \tag{4.27f}$$

With these definitions, we state the optimization problem corresponding to the mean-MPC for TI multiplicative uncertainty case, $J^M_{M,TI}(\xi_k)$.

Lemma 4.3.1 *Consider an LTI stochastic dynamical system Σ^M_{TI} as in Equation* (4.1b) *and a cost function $J_M(\tilde{\xi}_k)$ as in Equation* (4.4)*. Furthermore assume that δ and γ are independent random variables with diagonal (uncorrelated) variance-covariance matrices. Then the mean MPC cost function $J^M_{M,TI}(\xi_k)$ is given as*

$$\begin{aligned} J^M_{M,TI}(\xi_k) = \quad & \mathbb{E}\left\{J_M(\xi_k)\right\} = \mathbb{E}\left\{\sum_{j=0}^{N_p-1} x_{j|k}^\top Q x_{j|k} + u_{j|k}^\top R u_{j|k}\right\}, \\ = \quad & \sum_{j=0}^{N_p-1} \bar{x}_{j|k}^\top Q^{TI,M}_{j,M} \bar{x}_{j|k} + \sum_{j=0}^{N_p-2}\sum_{i=0}^{N_p-1} \bar{x}_{j|k}^\top Q^{TI,M}_{j,i,M} \bar{x}_{i|k} + \dots \\ & \sum_{j=0}^{N_p-1} u_{j|k}^\top R^{TI,M}_{j,M} u_{j|k} + \sum_{j=0}^{N_p-2}\sum_{i=0}^{N_p-1} u_{j|k}^\top R^{TI,M}_{j,i,M} u_{i|k} \end{aligned} \tag{4.28}$$

where the stage-wise varying weighting matrices $Q^{TI,M}_{j,M}$, $Q^{TI,M}_{j,i,M}$, $R^{TI,M}_{j,M}$ and $R^{TI,M}_{j,i,M}$ are found from the backwards recursion defined via permutations of terms over the uncertainty matrices A_i and central moments of random variables δ and γ as in Equations (4.26)-(4.27)*.*

Proof. The proof is structurally the same as the proof given in Lemma 1. The straightforward way of obtaining the cost terms is explicit evaluation of the expectation operator, since the true state can be decomposed as

$$x_{j|k} = \quad \bar{x}_{j|k} + \sum_{l=1}^{n_\delta}\sum_{i=1}^{j} \rho_{\bar{A},A_l,i-1} A_l[\delta]_l \bar{x}_{j-i|k} + \sum_{l=1}^{n_\gamma}\sum_{i=1}^{j} \rho_{\bar{A},A_l,i-1} B_l[\gamma]_l u_{j-i|k}. \tag{4.29}$$

The expectation operator vanishes the odd order matrix terms and then we rewrite all the terms as a nominal MPC problem with the given stage-wise varying state and input weights[5]. □

A direct consequence of Lemma 4.3.1 is stated below.

Theorem 4.3.1 *Consider an LTI stochastic dynamical system Σ^M_{TI} as in Equation* (4.1b)*, a cost function $J_M(\tilde{\xi}_k)$ as in Equation* (4.4) *subject to uncertain dynamics Σ^M_{TI} and the cost function $J^M_{M,TI}$ as in Equation* (4.28) *subject to nominal dynamics Σ^{nom}. Furthermore assume that δ and γ are independent from each other and described by zero mean and diagonal (uncorrelated) variance-covariance matrices. Then*

$$\begin{aligned} \min_{u_{[0,N_p-1]|k}} \mathbb{E}\{J_M(\tilde{\xi}_k)\} = \quad & \min_{u_{[0,N_p-1]|k}} J^M_{M,TI}(\xi_k), \\ \arg\min_{u_{[0,N_p-1]|k}} \mathbb{E}\{J_M(\tilde{\xi}_k)\} = \quad & \arg\min_{u_{[0,N_p-1]|k}} J^M_{M,TI}(\xi_k) \end{aligned} \tag{4.30}$$

[5] Instead of the constructive proof given for Lemma 4.2.1, one can use Equation (4.29) for showing the claimed result with a reduced number of steps.

Proof. With the results of Lemma 2, it is trivial to show that the claim is correct. □

Hence $u^*_{[0,N_p-1|k]}$ defines an optimal solution for both of the MPC problems. The nominal counterpart MPC problem with the cost function $J^M_{M,TI}(\xi_k)$ in Equation (4.24) and with decision variables $u_{j|k}$ is equivalent to the following MPC problem

$$\begin{aligned} u^*_{[0,N_p-1]|k} = \quad \arg\min_{u_{[0,N_p-1]|k}} \quad & \sum_{j=0}^{N_p-1} \bar{x}^\top_{j|k} Q^{TI,M}_{j,M} \bar{x}_{j|k} + \sum_{j=0}^{N_p-2} \sum_{i=j+1}^{N_p-1} 2\bar{x}^\top_{j|k} Q^{TI,M}_{j,i,M} \bar{x}_{i|k} + \dots \\ & \sum_{j=0}^{N_p-1} u^\top_{j|k} R^{TI,M}_{j,M} u_{j|k} + \sum_{j=0}^{N_p-2} \sum_{i=j+1}^{N_p-1} 2u^\top_{j|k} R^{TI,M}_{j,i,M} u_{i|k} \\ \text{subject to} \quad & \bar{x}_{i+1|k} = \bar{A}\bar{x}_{i|k} + \bar{B}u_{i|k}, \\ & \bar{x}_{0|k} = x_k, i \in \mathbb{Z}_{[0,N_p-1]}. \end{aligned} \tag{4.31}$$

Remark 4.3.1 *Deductions similar to the MV MPC for the TV uncertainty case are also valid for the MV MPC for TI uncertainty case. The nonlinear nature of the resulting MPC problem is further deteriorating the computational complexity, since the upper bounds on the evaluation of the optimizers in quadratic programs with uncertain Hessian matrices are difficult to calculate, see the discussion on the uncertain quadratic programs for (determibnsitic) robust optimization problems in [43]. Future research is directed towards incorporating the variance information of states, due to multiplicative uncertainties, into the robust counterpart cost function.*

4.4 Closed-loop Analysis of Moment-based MPC Formulations for Multiplicative Uncertainty in Dynamics

4.4.1 Stability Analysis of Moment-based MPC for Multiplicative Uncertainty in Dynamics

In this section we present the stability conditions for systems controlled with moment-based MPC. We use the cost functions given in Equation (4.10) and (4.28) to establish mean-square stability (MS-stability) of closed-loop stochastic systems.

By using the result of Proposition 3.1.1, we state the robust stability conditions on the closed-loop system controlled via mean MPC. Here we provide the stability conditions for the TV multiplicative uncertainty case, since the reasoning for the case of the weighting matrices $\bar{Q}^{TI,M}_M$ and $\bar{R}^{TI,M}_M$ also follows with similar deductions[6]. The following variables are used in the derivation of the stability conditions. The MPC cost function $J^{TV,M}_M(\xi_k)$ is taken as the candidate Lyapunov function, while for establishing stability, we follow the standard approach by assuming existence of a stabilizing state feedback controller K which, virtually, becomes active after the calculated optimal control actions, i.e., $u_{N_p|k} = Kx_{N_p|k}$. We construct a shifted cost function $\hat{V}(\xi_{k+1})$ which takes

$$\hat{u}_{[0,N_p-1]|k+1} = \begin{bmatrix} u^*_{[1,N_p-1]|k} & K\bar{x}_{N_p|k} \end{bmatrix}$$

[6]The operations in the TI case is much more tedious since we incorporate the cross terms.

as a feasible solution. Furthermore,

$$
\begin{aligned}
& \mathbb{E}\{\hat{V}(\xi_{k+1})\} - V(\xi_k) =: \xi_k \bar{M}_M^{M} \xi_k, \\
\bar{M}_M^{TV,M} := & \begin{bmatrix} \Theta^\top \bar{Q}_\Delta \Theta & \Theta^\top \bar{Q}_\Delta T_B^{N_p} \\ (\star) & T_B^{N_p^\top} \bar{Q}_\Delta T_B^{N_p} + \bar{R}_\Delta \end{bmatrix}, \\
\bar{Q}_\Delta^{TV} := & \operatorname{diag}\left(Q_{ex,M}^{TV,M} - Q_{0,M}^{TV,M}, Q_{0,M}^{TV,M} - Q_{1,M}^{TV,M}, \dots, Q_{N_p-1,M}^{TV,M} + K^\top R_{N_p-1,M}^{TV,M} K \right), \\
\bar{R}_\Delta^{TV} := & \operatorname{diag}\left(R_{ex,M}^{TV,M} - R_{0,M}^{TV,M}, R_{0,M}^{TV,M} - R_{1,M}^{TV,M}, \dots, R_{N_p-2,M}^{TV,M} - R_{N_p-1,M}^{TV,M} \right), \\
Q_{ex,M}^{TV,M} = & \sum_{l=1}^{n_\delta} A_{[\delta]_l}^\top \sum_{i=0}^{N_p-1} \left(\bar{A}^{i^\top} Q_{i,M}^{TV,M} \bar{A}^i + \bar{A}^{N_p-1^\top} K^\top R_{N_p-1,M}^{TV,M} K \bar{A}^{N_p-1} \right) A_{[\delta]_l} \sigma_{[\delta]_l}^2, \\
R_{ex,M}^{TV,M} = & \sum_{l=1}^{n_\gamma} B_{[\gamma]_l}^\top \sum_{i=0}^{N_p-1} \left(\bar{A}^{i^\top} Q_{i,M}^{TV,M} \bar{A}^i + \bar{A}^{N_p-1^\top} K^\top R_{N_p-1,M}^{TV,M} K \bar{A}^{N_p-1} \right) B_{[\gamma]_l} \sigma_{[\gamma]_l}^2.
\end{aligned}
\tag{4.32}
$$

Theorem 4.4.1 *Consider an uncertain LTI dynamical system Σ_{TV}^M as in Equation* (4.1c)*. If the mean MPC problem in Equation* (4.16) *is feasible at time k and $\bar{M}_M^{TV,M} \leq 0$, then the closed-loop system is MS-stable.*

Proof. For brevity, we consider the scalar uncertainty case, $\delta_k, \gamma_k \in \mathbb{R}$, while the proof follows with same reasoning and tedious algebra for multidimensional uncertainty case. We start our deduction by evaluating the difference in Lyapunov functions for current time instant and the next time instant, i.e.,

$$
\begin{aligned}
\mathbb{E}\{\hat{V}(\xi_{k+1})\} - V(\xi_k) = \; & -(\|\bar{x}_{0|k}^2\|_{Q_{0,M}^{TV,M}} + \|u_{0|k}\|_{R_{0,M}^{TV,M}}^2) + g_{\delta,\gamma}^{TV,M} + \\
& \cdots \sum_{i=1}^{N_p-1} \bar{x}_{i|k}^\top (Q_{i-1,M}^{TV,M} - Q_{i,M}^{TV,M}) \bar{x}_{i|k} + \\
& u_{i|k}^\top (R_{i-1,M}^{TV,M} - R_{i,M}^{TV,M}) u_{i|k},
\end{aligned}
\tag{4.33}
$$

where

$$
g_{\delta,\gamma}^{TV,M} = \bar{x}_{0|k}^\top Q_{ex,M}^{TV,M} \bar{x}_{0|k} + u_{0|k}^\top R_{ex,M}^{TV,M} u_{0|k} + \bar{x}_{N_p-1|k}^\top (Q_{N_p-1,M}^{TV,M} + K^\top R_{N_p-1,M}^{TV,M}) \bar{x}_{N_p-1|k}.
$$

Then adjust the right hand side of Equation (4.33) in matrix form such that $\mathbb{E}\{\hat{V}(\xi_{k+1})\} - V(\xi_k) = \xi_k \bar{M}_M^{TV,M} \xi_k$. If this expression is negative for all time instants k, equivalent to $\bar{M}_M^{TV,M} \leq 0$, then the closed-loop system is MS-stable. □

4.4.2 Closed-loop Characteristics and Bandwidth Analysis of Moment-based MPC for Multiplicative Uncertainty in Dynamics

In the mean MPC for model mismatch case, both of the weighting matrices $Q_{j,M}^{TI,M}$ (or $Q_{j,M}^{TV,M}$) and $R_{j,M}^{(\cdot),M}$ are effected by extra terms induced from the uncertainty effecting the system predictions, hence the resulting controller gain from the unconstrained MPC problem is also changing. The accumulation of the state weighting matrix is in general leading

to higher gain controllers in the classical control sense. However this increase in the weighting term is compensated with the accumulation of the input weighting matrix $R_{j,M}^{(\cdot),M}$. The increase of $R_{j,M}^{(\cdot),M}$ terms, which is due to the variable γ, decreases the aggressiveness of the resulting controller. This, in turn, is expected to yield a lower gain controller, thus, improving, in the classical control sense, the robustness aspect of the closed-loop system against the plant-model mismatches.

We demonstrate the effects of weighting matrices on the closed-loop behaviour on a simple simulation example which consists of two integrators, that is defined as

$$x_{k+1} = Ax_k + Bu_k, A = \begin{bmatrix} 1 & 1+\delta \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1+\gamma \end{bmatrix},$$

for which we construct four different cases; case (i) and case (ii) considers the TV mean MPC and TI mean MPC formulations, respectively, with the uncertain variables assumed as $\delta \sim \mathcal{N}(0, 0.01),\ \gamma \sim \mathcal{N}(0, 0.01)$, while for the case (iii) we assume that the A matrix is a deterministic matrix in prediction model, i.e., $\delta = 0$, and for the case (iv) we assume that the B is deterministic in prediction model, i.e., $\gamma = 0$, . The 'true' plant contains uncertain effects, with the mentioned uncertainty levels, for all cases and for both of the cases (iii) and (iv), we assume the TV mean MPC formulation to calculate the state and input weighting matrices. Lastly as comparison purposes we also show the nominal MPC result, which ignores the effect of uncertain elements. We report the closed-loop trajectories and the sensitivity function which are visualized in Figures 4.1-4.2. As can be observed from Figure 4.2, for both of the cases (i) and (ii), the bandwidth decreases in comparison to the nominal case. The case (ii) (TI mean) MPC casts the closed-loop system substantially more robust (in the classical sense) compared to the TV mean MPC and nominal MPC cases. As can be observed by comparing the cases (iii) and (iv) with the previous ones, once the uncertainties are dismissed from the MPC cost function the close-loop responses imitate the nominal MPC, while the effect of not incorporating input uncertainties, case (iv) (where γ is set to zero in the prediction model), is deteriorating the closed-loop performance more than the case (iii), since the random variable δ effects both state and input weighting matrices.

4.4.3 Generalization of Moment-based MPC with Multiplicative Uncertainty in Dynamics Towards Uncertainties with Even Distribution Functions

Till now we have only considered the case where the random variables are assumed to be described via Gaussian distribution characteristics. However, one does not necessarily need this strict assumption. It can be shown that the discussion in this chapter can also be extended towards different type of distribution characteristics, such as uncertainties that are distributed with uniform, triangular or Laplacian pdfs. As long as the distribution that is effecting the system matrices is an even function (symmetric with respect to zero) then the formulations given for both TV and TI mean MPC cases are similar to each other for different distribution functions. For this purpose, one needs to change the variance and higher order moment values in the cost function of mean MPC (TI or TV case) for non-Gaussian random variables γ and δ, that occurs in the backwards recursions during the calculation of $Q_{j,M}^{(\cdot),M}$ or $R_{j,M}^{(\cdot),M}$. We visualize the effects of using different distribution function families in a simulation example which is used in bandwidth analysis section. The

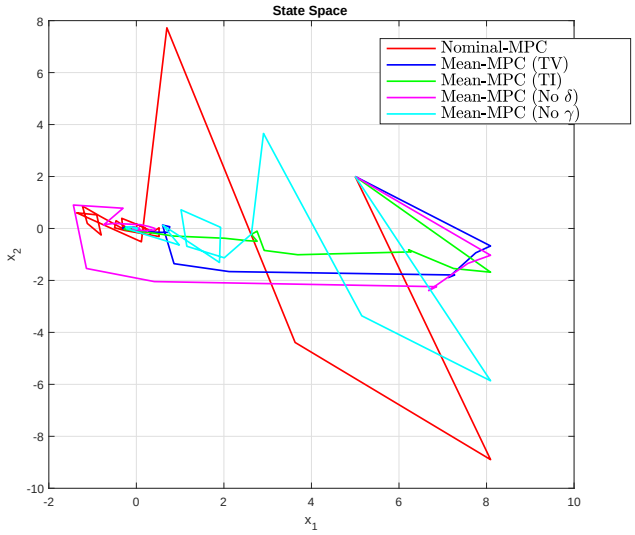


Figure 4.1: The state trajectories (phase-space) with different moment-based MPC formulations.

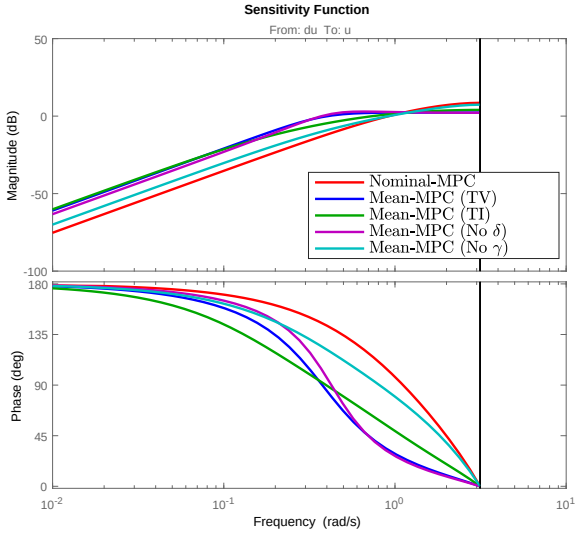


Figure 4.2: The sensitivity function of closed-loop system with different moment-based MPC formulations.

simulation results are shown in Figures 4.3-4.4.

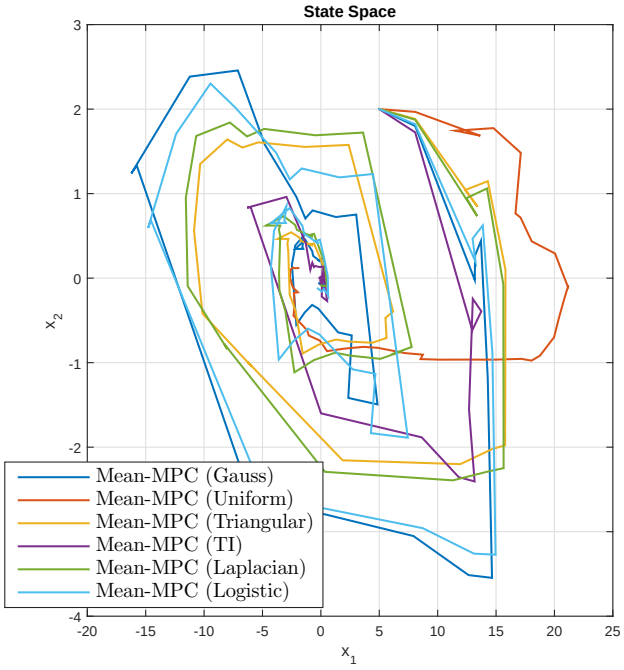


Figure 4.3: The state trajectories of moment-based MPC formulations with different even pdfs.

4.5 Simulation Example for Moment-based MPC for Multiplicative Uncertainty in Dynamics

In this section, we make use of a mass-spring system, similar to the case in [185], which consists of two masses connected with a spring. We control this system with various mean MPC formulations that are formulated in this chapter. The control action is effective on one of the masses and the control goal is to effectively stabilize the system. The dynamic model of the process is taken as follows;

$$x_{k+1} = A(\delta_k)x_k + B(\gamma_k)u_k,$$
$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -\frac{K+\delta}{m_1} & \frac{K+\delta}{m_1} & 1 & 0 \\ \frac{K+\delta}{m_2} & \frac{-K+\delta}{m_2} & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{m_1} \\ \frac{\gamma}{m_2} \end{bmatrix}, \tag{4.34}$$

where the parameters, the spring constant K and the masses m_1 and m_2 are taken as, $K = 1,\ m_1 = 0.5,\ m_2 = 2$ and the random variables' pdfs are assumed to be given as $\delta \in$

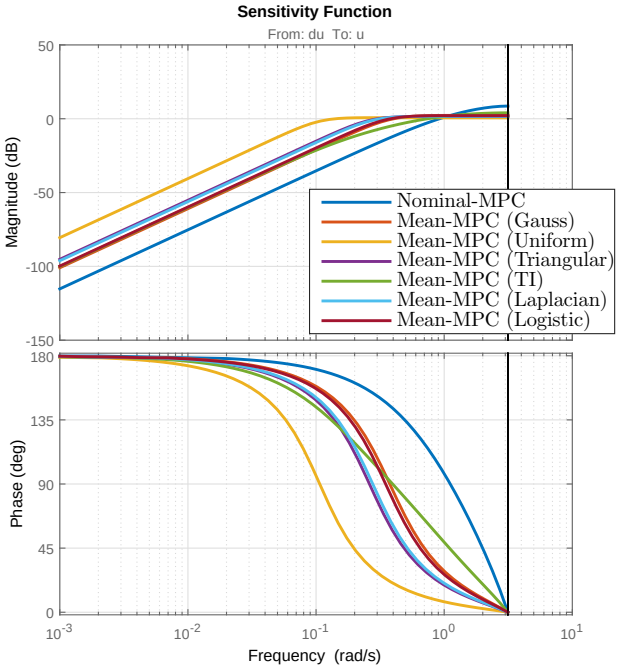


Figure 4.4: The sensitivity function of closed-loop systems with moment-based MPC formulations formulations with different even pdfs.

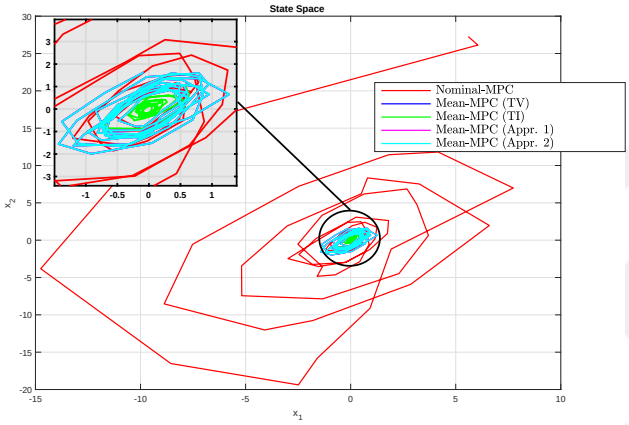
$\mathcal{N}(0, 0.01)$ and $\gamma \in \mathcal{N}(0, 0.01)$.

We present two sets of simulations, the first simulation example demonstrates the TV uncertainty case, i.e., the true disturbances are varying over time, while the MPC controllers are constructed from the cost functions found as the TI mean MPC, the TV mean MPC, the approximating case 1 and 2 for mean MPC and the nominal MPC. In the second simulation example we consider TI random variables in the true system dynamics and show the responses for five different MPC controllers mentioned above. In both of the simulation studies we have used the same (base) state and input weighting matrices, such as

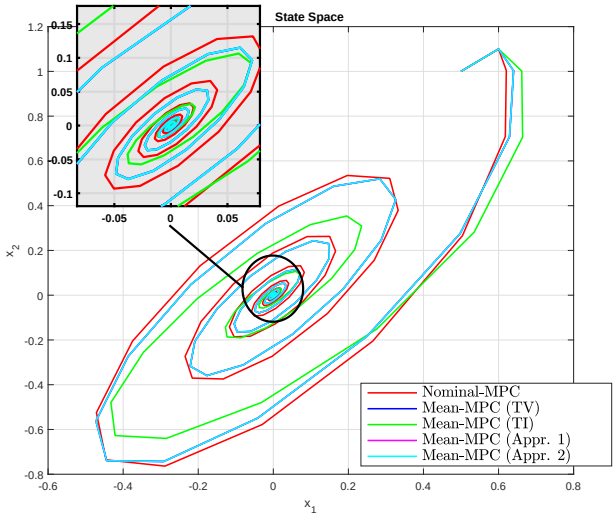
$$Q = 5I_4, \quad R = I_1, \quad Q_f = 0,$$

while the prediction horizon is taken as 4. We initialize the state at $x_0^\top = \begin{bmatrix} 0.5 & 1 & -0.1 & 0.1 \end{bmatrix}$ for both of the simulations.

The resulting closed-loop simulation trajectories are visualized in Figure 4.5a for the case of TV random variables effecting the true process and Figure 4.5b for the time invariant random variables. Most striking aspect about the simulations is that in the TV uncertainties case, the nominal controller which does not incorporate the information about the uncertainty can not stabilize the system, while all other MPC controllers improve the



(a) The position trajectories with different moment-based MPC formulations for TV plant-model mismatch case.



(b) The position trajectories with different moment-based MPC formulations for TI plant-model mismatch case.

Figure 4.5: MSD System with moment-based MPC.

controlled system's behaviour and eventually steer the states to origin.

4.6 Conclusions on Moment-based MPC Problems for Linear Systems with Multiplicative Uncertainty

In this chapter, we have investigated the moment-based MPC problems which are based on the statistics (centralized moments) of the cost functionals for linear stochastic systems with multiplicative uncertainties. We have evaluated the mean of the MPC cost function for time-varying and time-invariant uncertainties affecting the system dynamics in an affine fashion. We present explicit formulae for the reformulation of the state and input weighting matrices which are different for these two cases. Since computation of the weighting matrices is tedious for long prediction horizons, we present two different approximation possibilities for realistic processes or MPC parameters. Lastly, we provide the conditions on the robust stability of the closed-loop system and presented simulation results and discussions on the closed-loop performance.

- The moment-based MPC problems can be expressed in terms of nominal dynamics with stage-wise varying state and input weighting matrices over the prediction horizon.
- The computational complexity of moment-based MPC problems is reduced drastically, in comparison to the other robust MPC techniques.
- The non-Gaussian random variables case demonstrates the possibility of generalizing the moment-based MPC towards generic random variables which have finite moments.
- The structure of uncertain effects within the prediction model changes the reformulation of the predictive control problem, i.e., controllers tend to become
    1. More aggressive if the prediction model is assumed to be correct yet the measurements are corrupted.
    2. More timid if it is known that the model is leading to wrong predictions.

Chapter 5

Constraints in Moment-based Model Predictive Control Problems

A method is a trick I use twice.

George Pólya - How to solve it

The goal of this chapter is to present constraint tightening techniques for MPC problems cast for uncertain dynamical systems by using statistical information (centralized moments) of the uncertain effects. We make use of expectation and variance operators to describe the uncertain state's future spread information. By manipulating the variance, we show that one can effectively control the constraint violation levels. We start this chapter by first formulating the constraint tightening techniques for the additive uncertainty case while considering the mean and mean-variance (MV) formulations for perturbations that are distributed with Gaussian pdf. We present several independent results to reduce the complexity associated with the mean variance treatment of the constraint functions. One of the crucial problems of uncertainties with unbounded domains is the fact that the recursive feasibility (over time) can not be guaranteed in deterministic sense. We present a method to provide probabilistic guarantees on recursive feasibility by using the variance of state trajectories. Secondly, we extend our results for disturbances that are characterized with zero mean even pdfs that are not Gaussian functions. The even nature of pdf function allows us to formulate explicit constraint tightening methods for bounded and even pdf uncertainties. Finally, we consider the multiplicative uncertainty case, where uncertainties are present in the prediction model. We present results on constraint treatment for time varying or invariant perturbations effecting dynamical system parameters.

This chapter is structured as follows. We conclude the introduction section with introducing the notation used in the next sections. In Section 5.2 we present the results for constrained MPC problems with additive Gaussian pdf uncertainties. We first treat bound constraints and formulate a non-conservative and a conservative robust counterpart problem.

[0]Substantial content of this chapter is also published or presented in 'M.B. Saltik, L. Özkan, S. Weiland. Constraint Tightening in Moment-based Model Predictive Control.'

Then we extend the results for polytopic constraints, which are a generalization of bound constraints. In Section 5.2.4, we extend our deductions to systems which are perturbed with random variables that are distributed with even and bounded pdfs. We formulate constraint tightening techniques under the setting of moment-based MPC problems for bounded distributions. In Section 5.3 we discuss the case where the prediction model contains statistical information on the uncertain effects on the system dynamics.

5.1 Problem Setting for Robust Counterpart Constraint Formulations in Moment-based MPC Problems

Similar to the previous two chapters, here we consider several possible descriptions for linear systems under the effect of uncertainty. We define a nominal system, denoted with Σ^{nom} given in state-space representation as in Equation (5.1a), a system with additive disturbances effecting the state and output equations, denoted with Σ^{add} as given in Equation (5.1b), and two systems effected by multiplicative uncertainty, for the time-invariant (TI) and the time-varying (TV) uncertainty cases, denoted with Σ^{mul}_{TI} and Σ^{mul}_{TV} as in Equations (5.1c) and (5.1d), respectively;

$$\Sigma^{nom}: \quad \bar{x}_{k+1} = \quad A\bar{x}_k + Bu_k, \tag{5.1a}$$

$$\Sigma^{add}: \quad x_{k+1} = \quad Ax_k + Bu_k + Fw_k, \tag{5.1b}$$

$$\Sigma^{mul}_{TI}: \quad x_{k+1} = \quad A(\delta)x_k + B(\gamma)u_k, \tag{5.1c}$$

$$\Sigma^{mul}_{TV}: \quad x_{k+1} = \quad A(\delta_k)x_k + B(\gamma_k)u_k, \tag{5.1d}$$

These plants are subject to three types of uncertainties which are denoted by w_k, δ_k (or δ) and γ_k (or γ) for $k \in \mathbb{Z}_{\geq 0}$. For these systems, the future predictions of state $\bar{x}_{j|k} := \bar{x}_{k+j}(\bar{x}_k, u_{[k,k+j-1]})$ can be expressed, respectively, as

$$\bar{x}_{j|k} = A^j \bar{x}_{0|k} + \sum_{i=0}^{j-1} A^{j-1-i} B u_{i|k}, \tag{5.2a}$$

$$x_{j|k} = A^j x_{0|k} + \sum_{i=0}^{j-1} A^{j-1-i} B u_{i|k} + \sum_{i=0}^{j-1} A^{j-1-i} F w_{i|k}, \tag{5.2b}$$

while for the multiplicative cases the state predictions can be expressed as

$$\begin{aligned} x_{0|k} &= \quad x_k, \\ x_{j|k} &= \quad \prod_{i=0}^{j-1} A(\delta_{i|k}) x_{0|k} + \sum_{i=0}^{j-1} \prod_{l=i+1}^{j-1} A(\delta_l) B(\gamma_i) u_{i|k}, \qquad j \geq 1, \end{aligned} \tag{5.2c}$$

where $\prod_{i=l_1}^{l_2} = I$ if $l_1 > l_2$.

The rest of notation used in this chapter and the nature of uncertainties are assumed to be as in the specifications described in Chapter 3 and Chapter 4.

5.2 Constraint Reformulation in Moment-based MPC Problems with Additive Gaussian Perturbations

In this section, we consider the state constraints for the case of moment-based MPC problem cast on linear systems with additive disturbances, Σ^{add} in Equation (5.1b). Here we do not work with control laws, i.e., $u_k = \kappa(x_k)$, hence satisfying constraints on control actions is straightforward for the MPC problems discussed here. On the contrary, guaranteeing (hard) state (equivalently output) constraint satisfaction is impossible, if the disturbances $w_k,\ k \in \mathbb{Z}_{\geq 0}$, assume values from unbounded domains, such as the Gaussian pdfs and with a non-trivial F matrix. For this case, confidence bounds on the constraint satisfaction can be provided. Next, we discuss the bound constraints case cast on the evolution of dynamics Σ^{add}.

5.2.1 Bound Constraints in Moment-based MPC for Additive Gaussian Uncertainty

We first consider the bound constraints on the entries of inputs and states (equivalently outputs[1]), i.e.,

$$\begin{aligned} [-b^x]_i \leq [x_k]_i \leq [b^x]_i,\ \forall k \in \mathbb{Z}_{\geq 0}, \\ [-b^u]_i \leq [u_k]_i \leq [b^u]_i,\ \forall k \in \mathbb{Z}_{\geq 0}, \end{aligned} \tag{5.3}$$

where b^x and $-b^x$ are the constraint vectors, of which the entries are constraining the state evolution.

First we state a trivial, yet core, observation for linear systems with additive uncertainty. This result states that the mean of the constraint distribution is equivalent to the constraints on the nominal predicted state (generated from system Σ^{nom}) for system Σ^{add} in Equation (5.1b).

Theorem 5.2.1 *Consider a stochastic LTI dynamical system Σ^{add} in Equation* (5.1b) *affected by Gaussian perturbations $w_k,\ k \in \mathbb{Z}_{\geq 0}$ and the bound constraints as in Equation* (5.3)*. Then, for every $i \in \mathbb{Z}_{[1,n]}$,*

$$[-b^x]_i \leq \mathbb{E}\{[x_{j|k}]_i\} \leq [b^x]_i \iff [-b^x]_i \leq [\bar{x}_{j|k}]_i \leq [b^x]_i, \tag{5.4}$$

where $\bar{x}_{j|k}$ is the j^{th}-step prediction of Σ^{nom} with initial condition $\bar{x}_{0|k} = x_k$ and inputs $u_{l|k},\ l \in \mathbb{Z}_{[0,j-1]}$.

Proof. For the true state predictions $x_{j|k}$ given in Equation (5.2b), the disturbances $w_{l|k}$ are occurring in an affine way. By using the linearity property of the expectation operator, the disturbance terms $\sum_{i=0}^{j-1} A^{j-1-i} F w_{i|k}$ evaluate to zero thus nominal predictions $\bar{x}_{j|k}$ represent $\mathbb{E}\{x_{j|k}\}$. □

Theorem 5.2.1 shows that the mean based reformulation of constraints for linear systems with additive disturbances does not improve the constraint satisfaction properties[2]. One can improve the constraint satisfaction probability, which we denote with α_j for j^{th} prediction

[1] Since, in general, the states and outputs are algebraically coupled to each other, i.e. $y_k = Cx_k$, all of the results presented here are valid also for the outputs, with required changes in the constraint functions.

[2] Yet on average the constraints are expected to be satisfied, under the ergodicity of the uncertain effects.

stage at time instant k, by incorporating the covariance of the state predictions $\Sigma_{x_{j|k}}$ into the robust counterparts of the constraint functions as;

$$\begin{array}{ll} [\mathbb{E}\{x_{j|k}\} + \lambda_j \mathbb{D}\{x_{j|k}\}]_i \leq & [b^x]_i, \\ [\mathbb{E}\{x_{j|k}\} - \lambda_j \mathbb{D}\{x_{j|k}\}]_i \geq & [-b^x]_i. \end{array} \tag{5.5}$$

We approximate the MV reformulation of the constraints in Equation (5.5) by using the Mahalanobis distance ([109]) and the state covariance matrix $\Sigma_{x_{j|k}}$ which can be parameterized with the characteristics of disturbances w_k. First, observe that the state predictions can be decomposed into $x_{j|k} = \bar{x}_{j|k} + \tilde{x}_{j|k}$, with $x_{j|k} \in \mathcal{N}(\bar{x}_{j|k}, \Sigma_{x_{j|k}})$ and $\Sigma_{x_{j|k}} = \mathbb{E}\{\tilde{x}_{j|k}\tilde{x}_{j|k}^\top\}$ while $\Sigma_{x_{0|k}} = 0$, if there is no initial condition mismatch[3]. The state (prediction) covariance matrix, for each prediction stage $j \in \mathbb{Z}_{[0,N_p-1]}$, is calculated as[4],

$$\begin{array}{rcl} \Sigma_{x_{j|k}} = & \mathbb{E}\{(x_{j|k} - \bar{x}_{j|k})(x_{j|k} - \bar{x}_{j|k})^\top\} = \mathbb{E}\{\tilde{x}_{j|k}\tilde{x}_{j|k}^\top\} \\ = & \sum_{i=0}^{j-1} A^{j-1-i} F \Sigma_w F^\top A^{j-1-i^\top}, \; j \geq 1. \end{array} \tag{5.6}$$

Essentially, the state covariance matrix $\Sigma_{x_{j|k}}$ is used in Equation (5.5) for backing off the predicted (and nominal) operating point $\bar{x}_{j|k}$ from the (bound) constraints $\mathcal{P}([-b^x], [b^x])$, similar to the case in [355] and $\mathcal{P}([-b^x], [b^x])$ denotes the polytope defined by the vertices $[-b^x]$ and $[b^x]$ ([54]). We make use of Mahalanobis distance ([109]) to tighten the constraints, which evaluates the distance of an observation ρ to a Gaussian distribution $x \in \mathcal{N}(c_x, \Sigma_x)$.

Definition 5.2.1 *The Mahalanobis distance (M-distance), denoted with $d(\rho, x)$, of a point $\rho \in \mathbb{R}^{n_x}$ to a random variable $x \in \mathcal{N}(c_x, \Sigma_x)$ is given by*

$$d(\rho, x)^2 := (\rho - c_x)^\top \Sigma_x^{-1} (\rho - c_x). \tag{5.7}$$

The definition of M-distance takes into account the mean c_x and the covariance matrix Σ_x of random variable x, hence the directions with relatively small variance values are evaluated with larger M-distance for an Euclidean level set (hyper-spheres in the vector space). Furthermore, for a random realization of x, say ρ, the M-distance $d(\rho, x)$ not only provides confidence intervals but itself is also a random variable distributed with $\chi^2(n_x)$ distribution, the χ^2 distribution with n_x degrees of freedom, [355], where n_x is the dimension of x. Hence if one sets a desired probability level for satisfying the constraint at prediction stage j as $\alpha_j \in \mathbb{R}_{[0,1]}$ while denoting the cumulative distribution function of $\chi^2(n_x)$ with $\Omega_{\chi^2} : \mathbb{R} \to \mathbb{R}_{[0,1]}$, then the minimum M-distance for guaranteeing the probability level α_j is given by $D^{2^*}(\alpha_j) := \Omega^{-1}_{\chi^2(n_x)}(\alpha_j)$, where $\Omega^{-1}_{\chi^2(n_x)} : \mathbb{R}_{[0,1]} \to \mathbb{R}^{n_x}$ is the inverse

[3] For the case where the initial condition is not known exactly at time instant k, such as the output feedback MPC configuration, the derivations presented here should also incorporate initial condition mismatch covariance matrix as an linearly additive term, due to independence of uncertainties and linearity of dynamics. Here, for brevity, we assume that initial condition of state is known exactly by the MPC controller.

[4] This matrix can be calculated and stored in MPC algorithm during the offline (design) stage.

cumulative distribution function of $\chi^2(n_x)$. This claim can be seen from;

$$\begin{aligned} \mathbb{P}\{d(\rho, x_{j|k})^2 \leq D^{2^*}(\alpha_j)\} = &\quad \alpha_j, \\ \mathbb{P}\{(\rho - c_{x_{j|k}})^\top \Sigma_{x_{j|k}}^{-1}(\rho - c_{x_{j|k}}) \leq D^{2^*}(\alpha_j)\} = &\quad \alpha_j, \\ \mathbb{P}\{\rho \in \mathcal{E}(c_{x_{j|k}}, \Sigma_{x_{j|k}}, D^*(\alpha_j))\} = &\quad \alpha_j, \end{aligned} \tag{5.8}$$

where the equation reads as the probability of a random realization of $x_{j|k}$, i.e., ρ, being an element of the ellipsoid $\mathcal{E}$ defined with the center $c_{x_{j|k}}$, the shape matrix $\Sigma_{x_{j|k}}$ and the radius $D^*(\alpha_j)$ is α_j. With this line of reasoning, we transform the bound constraints on the states, i.e., Equation (5.5), to chance constraints in Lemma 5.2.1.

Lemma 5.2.1 *Consider a linear dynamical system Σ^{add} as in Equation* (5.1b) *affected with Gaussian perturbations $w_k,\ k \in \mathbb{Z}_{\geq 0}$ and bound constraints as in Equation* (5.3)*. Given an initial condition $x_{0|k}$ and instantaneous constraint satisfaction levels $\alpha_j,\ j \in \mathbb{Z}_{[0,N_p-1]}$ and let $\Sigma_{x_{j|k}} = L_j L_j^\top, ||\beta||_2^2 = 1, \beta \in \mathbb{R}^{n_x}$. Then*

$$[-b^x]_i \leq [\bar{x}_{j|k} + D^*(\alpha_j) L_j \beta]_i \leq [b^x]_i,\ i = 1, \ldots, n_x, \tag{5.9}$$

implies

$$\mathbb{P}\{[-b^x]_i \leq [x_{j|k}]_i \leq [b^x]_i\} \geq \alpha_j,\ i = 1, \ldots, n_x. \tag{5.10}$$

Proof. From Equation (5.8), we know that the ellipsoid defined by $\bar{x}_{j|k}, \Sigma_{x_{j|k}}$ and $D^*(\alpha_j)$ guarantees the probability level

$$\mathbb{P}\{x_{j|k} \in \mathcal{E}(\bar{x}_{j|k}, \Sigma_{x_{j|k}}, D^*(\alpha_j))\} = \alpha_j.$$

Then we show that the ellipsoid $\mathcal{E}(\bar{x}_{j|k}, \Sigma_{x_{j|k}}, D^*(\alpha_j))$ is a subset of the constraint region defined by the bounds, i.e., $\mathcal{P}(-b^x, b^x)$. By Definition 5.2.1 the confidence ellipsoid loci is given by $z = \bar{x}_{j|k} + D^*(\alpha_j) L_j \beta$. With this loci, if Equation (5.9) holds, then the minimum M-distance between $\bar{x}_{j|k}$ and any point on the constraint set is larger than $D^*(\alpha_j)$, hence $\mathcal{E}(\bar{x}_{j|k}, \Sigma_{x_{j|k}}, D^*(\alpha_j)) \subset \mathcal{P}(-b^x, b^x)$. □

Remark 5.2.1 *Checking whether an ellipsoid lies inside a polytope, i.e., $\mathcal{E} \subset \mathcal{P}$, is a second order conic program (SOCP). This increases the complexity class of the MPC problem from a linearly constrained optimization problem to an optimization problem with conic constraints.*

Ignoring the trivial cases, with system dynamics Σ^{add} one generally ends up with a nondiagonal covariance matrix $\Sigma_{x_{j|k}}$. This matrix can be approximated by a diagonal (uncorrelated) matrix, such as $\text{Tr}(\Sigma_{x_{j|k}}) I_{n_x}$ or any other matrix which satisfies $\hat{\Sigma}_{x_{j|k}} \geq \Sigma_{x_{j|k}}$ and $\hat{\Sigma}_{x_{j|k}}$ diagonal. A diagonal covariance matrix allows the user to treat the state space components individually, since the tip points of the ellipsoid $\mathcal{E}(\bar{x}_{j|k}, \hat{\Sigma}_{x_{j|k}}, D^*(\alpha_j))$ are aligned with the bases of the state space, i.e., the eigenvectors of $\hat{\Sigma}_{x_{j|k}}$ are parallel to basis directions of the state space. Then one can use the Gaussian error functions, denoted with $\text{erf}(\cdot)$, for

allocating the uncertainties that exceed the constraint level set. For a random variable with diagonal covariance matrix with standard deviations $\hat{\sigma}_{[x]_i}$, $a_i \in \mathbb{R}_{\geq 0}$ and $i \in \mathbb{Z}_{[1,n_x]}$ we have,

$$\mathbb{P}\{-a_i \leq [\hat{x}]_i \leq a_i\} = \text{erf}\left(\frac{a_i}{\sigma_{[\hat{x}]_i}\sqrt{2}}\right),$$

This approximation reduces the complexity of conic constraints (Equation (5.9)) back to linear ones, see Lemma 5.2.2 and also [51].

Theorem 5.2.2 *Consider a linear dynamical system Σ^{add} as in Equation* (5.1b) *affected with Gaussian perturbations $w_k,\ k \in \mathbb{Z}_{\geq 0}$ and bound constraints as in Equation* (5.3)*. Given a diagonal over-approximation of the state covariance matrix, $\hat{\Sigma}_{x_{j|k}} \geq \Sigma_{x_{j|k}}$, then the tightened set of bound constraints guaranteeing probabilistic confidence levels in Equation* (5.10) *are given as*

$$[-b^x]_i + [a_{j|k}]_i \leq [\bar{x}_{j|k}]_i \leq [b^x]_i - [a_{j|k}]_i, \tag{5.11}$$

where $a_{j|k}$ satisfies

$$\prod_{i=1}^{n_x} \text{erf}\left(\frac{[a_{j|k}]_i}{\hat{\sigma}_{[x_{j|k}]_i}\sqrt{2}}\right) \geq \alpha_j, \tag{5.12}$$

and $\hat{\Sigma}_{x_{j|k}} = \text{diag}(\hat{\sigma}_{[x_{j|k}]_1}, \ldots, \hat{\sigma}_{[x_{j|k}]_{n_x}})$.

Proof. First we make use of the new random variable $\hat{x}_{j|k}$ with $\hat{x}_{j|k} \in \mathcal{N}(0, \hat{\Sigma}_{x_{j|k}})$. Then, observe that $\hat{\Sigma}_{x_{j|k}} \geq \Sigma_{x_{j|k}}$,

$$\mathbb{P}\{x_{j|k} \in \mathcal{P}(-b^x, b^x)\} \geq \mathbb{P}\{\bar{x}_{j|k} + \hat{x}_{j|k} \in \mathcal{P}(-b^x, b^x)\}.$$

Then we guarantee $\mathbb{P}\{\bar{x}_{j|k} + \hat{x}_{j|k} \in \mathcal{P}(-b^x, b^x)\} \geq \alpha_j$. By using the tightening distances $a_{j|k}$ as design parameters and the diagonal structure of $\hat{\Sigma}_{x_{j|k}}$, i.e., for $i = 1, \ldots, n_x$,

$$\mathbb{P}\{[\bar{x}_{j|k}]_i - [a_{j|k}]_i \leq [\bar{x}_{j|k}]_i + [\hat{x}_{j|k}]_i \leq [\bar{x}_{j|k}]_i + [a_{j|k}]_i\} = \prod_{i=1}^{n_x} \text{erf}\left(\frac{[a_{j|k}]_i}{\hat{\sigma}_{[x_{j|k}]_i}\sqrt{2}}\right),$$

Then we set the tightened bounds as in Equation (5.11) by selecting each component of $a_{j|k}$ which satisfies Equation (5.12). □

By employing the methods discussed in Theorem 5.2.2 we end up with linear constraints in the nominal state predictions generated from Σ^{nom} in Equation (5.1a) and a new constraint of tightening distances a_i as in Equation (5.12). However there are two sources inducing conservatism in this constraint reformulation. The first source is the diagonal over-approximation of the covariance matrix $\Sigma_{x_{j|k}}$ by $\hat{\Sigma}_{x_{j|k}}$. This means that we construct the (smallest) axis-aligned ellipsoidal region which covers the ellipsoidal region described by $\Sigma_{x_{j|k}}$ in the M-distance sense. The second source of conservatism is the $a_{j|k}$ values. We know that $\text{erf}(\zeta) \leq 1, \zeta \in \mathbb{R}_{\geq 0}$, thus multiplication of error functions quickly diminishes for large state space dimensions, hence it might be inconvenient to pre-allocate the tightening distances $[a_{j|k}]_i$ without reasonable physical interpretation.

5.2.2 Polytopic Constraints in Moment-based MPC for Additive Gaussian Uncertainty

In the following part polytopic constraints on the true state $x_{j|k}$ are transformed into conic constraints on nominal state $\bar{x}_{j|k}$, where the confidence region depends on the confidence levels α_j and covariance matrix $\Sigma_{x_{j|k}}$, similar to the case for the bound constraints. In this case, we define the constraint functions as

$$A_j x_{j|k} \leq b_j, \tag{5.13}$$

and provide a statistical approach to satisfy

$$\mathbb{P}\{A_j x_{j|k} \leq b_j\} \geq \alpha_j$$

where α_j are the chance levels of satisfying the constraints at j^{th} prediction stage. In Lemma 5.2.2 we summarize the probabilistic confidence guarantees on the polytopic constraints.

Lemma 5.2.2 *Consider a linear dynamical system Σ^{add} as in Equation* (5.1b) *affected with Gaussian perturbations $w_k,\ k \in \mathbb{Z}_{\geq 0}$ and polytopic constraints as in Equation* (5.13)*. Given an initial condition $x_{0|k}$ and instantaneous constraint satisfaction levels $\alpha_j,\ j \in \mathbb{Z}_{0,N_p-1}$ and let $\tilde{A}_j := D^*(\alpha_j)A_j L_j,\ L_j L_j^\top := \Sigma_{x_{j|k}},\ \|\beta\|^2 = 1$. Then,*

$$A_j \bar{x}_{j|k} + \tilde{A}_j \beta \leq b_j, \tag{5.14}$$

implies

$$\mathbb{P}\{A_j x_{j|k} \leq b_j\} \geq \alpha_j. \tag{5.15}$$

Proof. First observe the fact that $\zeta_{j|k} = A_j x_{j|k}$ is a random variable with characteristics given as $\zeta_{j|k} \in \mathcal{N}(A_j \bar{x}_{j|k}, A_j \Sigma_{x_{j|k}} A_j^\top)$. Then the deductions are essentially same with the proof given in Lemma 5.2.1. □

Similar to the bound constraints case, one can approximate the conic constraints in Equation (5.14) with linear constraints and over-approximating diagonal covariance matrix $\hat{\Sigma}_{x_{j|k}}$.

Theorem 5.2.3 *Consider a linear dynamical system Σ^{add} as in Equation* (5.1b) *affected with Gaussian perturbations $w_k,\ k \in \mathbb{Z}_{\geq 0}$ and polytopic constraints as in Equation* (5.13)*. Given an initial condition $x_{0|k}$, instantaneous constraint satisfaction levels $\alpha_j,\ j \in \mathbb{Z}_{[0,N_p-1]}$, and a diagonal over-approximation of the covariance matrix $A_j \Sigma_{x_{j|k}} A_j^\top$, i.e., $\hat{\Sigma}_{x_{j|k}} \geq A_j \Sigma_{x_{j|k}} A_j^\top$ and $\hat{\Sigma}_{x_{j|k}} = \text{diag}(\hat{\sigma}_{[x_{j|k}]_1}, \ldots, [\hat{\sigma}[x_{j|k}]_{n_x}})$. If*

$$A_j \bar{x}_{j|k} + a_{j|k} \leq \quad b_j, \tag{5.16}$$

where $a_{j|k}$ satisfies Equation (5.12)*, then*

$$\mathbb{P}\{A_j x_{j|k} \leq b_j\} \geq \alpha_j.$$

Proof. First we restate that, if $\hat{x}_{j|k} \in \mathcal{N}(0, \hat{\Sigma}_{x_{j|k}})$ and $\mathbb{P}\{-a_i \leq [\hat{x}]_i \leq a_i\} = \text{erf}\left(\frac{a_i}{\hat{\sigma}_{[x_{j|k}]_i}\sqrt{2}}\right)$, then $\mathbb{P}\{A_j\bar{x}_{j|k} - a_{j|k} \leq A_j\bar{x}_{j|k} + \hat{x}_{j|k} \leq A_j\bar{x}_{j|k} + a_{j|k}\} \geq \prod_{i=1}^{n_x} \text{erf}\left(\frac{[a_{j|k}]_i}{\hat{\sigma}_{[x_{j|k}]_i}\sqrt{2}}\right)$, with the equality satisfied only if A_j has full image rank. Due to hypothesis, $\hat{\Sigma}_{x_{j|k}} \geq A_j\Sigma_{x_{j|k}}A_j^\top$, we have $\mathbb{P}\{A_jx_{j|k} \leq b_j\} \geq \mathbb{P}\{A_j\bar{x}_{j|k} + \hat{x}_{j|k} \leq b_j\}$. Then if one selects the $a_{j|k}$ vector such that the condition

$$\prod_{i=1}^{n_x} \text{erf}\left(\frac{[a_{j|k}]_i}{\hat{\sigma}_{[x_{j|k}]_i}\sqrt{2}}\right) \geq \alpha_j$$

is satisfied, then the confidence bounds are guaranteed. □

Remark 5.2.2 *Using the reasoning given above, one can also formulate ellipsoidal constraints on the true state, such as*

$$x_{j|k}^\top H_j x_{j|k} - 2b_j^\top x_{j|k} - c_j \leq 0. \tag{5.17}$$

Since we can decompose the true state as $x_{j|k} = \bar{x}_{j|k} + \tilde{x}_{j|k}$, the constraints in Equation (5.17) *can be expressed as, assuming H_j symmetric,*

$$\bar{x}_{j|k}^\top H_j \bar{x}_{j|k} - 2\tilde{b}_j^\top \bar{x}_{j|k} - \tilde{c}_j \leq 0, \tag{5.18}$$

where $\tilde{b}_j = b_j - \tilde{x}_{j|k}^\top H_j$ and $\tilde{c}_j = c_j - \tilde{x}_{j|k}^\top H_j \tilde{x}_{j|k} + 2\tilde{b}_j^\top \tilde{x}_{j|k}$. Satisfying this constraint in its general form with guaranteed probability levels is difficult, see [43]. This is due to the fact that the coefficients of each term should be bounded by an ellipsoid to cast the constraint satisfaction problem tractable. For the term $\tilde{b}_j$, we can express the uncertainty region with an ellipsoid, since $\tilde{x}_{j|k}$ is Gaussian. However the term $\tilde{c}_j$ is not distributed with a Gaussian pdf, due to the quadratic term $\tilde{x}_{j|k}^\top H_j \tilde{x}_{j|k}$. Here we provide an approximate result which assumes that the distribution of coefficients $\tilde{b}_j$ and $\tilde{c}_j$ are known to be Gaussian beforehand.

Proposition 5.2.1 *[43] Consider a linear dynamical system Σ^{nom} as in Equation* (5.1a) *and quadratic constraints as in Equation* (5.18) *where*

$$\tilde{b}_j = \bar{b}_j + \sum_{i=1}^{n_Q} [\beta_j]_i \tilde{b}_{ji}, \qquad \tilde{c}_j = \bar{c}_j + \sum_{i=1}^{n_Q} [\beta_j]_i \tilde{c}_{ji},$$

where $\beta_j^\top \beta_j = 1$. *If*

$$\begin{bmatrix} \bar{c}_j + 2\bar{x}_{j|k}^\top \bar{b}_j - \tau_j & (\star) & (\star) & \dots & (\star) & (\star) & \bar{x}_j^\top H_j \\ 1/2\tilde{c}_{j1} + \bar{x}_{j|k}^\top \tilde{b}_{j1} & \tau_j & 0 & \dots & 0 & 0 & 0 \\ 1/2\tilde{c}_{j2} + \bar{x}_{j|k}^\top \tilde{b}_{j2} & 0 & \tau_j & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 1/2\tilde{c}_{jn_Q-1} + \bar{x}_{j|k}^\top \tilde{b}_{jn_Q-1} & 0 & \ddots & \ddots & \tau_j & 0 & 0 \\ 1/2\tilde{c}_{jn_Q} + \bar{x}_{j|k}^\top \tilde{b}_{jn_Q} & 0 & 0 & \dots & 0 & \tau_j & 0 \\ H_j \bar{x}_{j|k} & 0 & 0 & \dots & 0 & 0 & I_{n_x} \end{bmatrix} \geq 0, \tag{5.19}$$

then $\bar{x}_{j|k}^\top H_j \bar{x}_{j|k} - 2\tilde{b}_j \bar{x}_{j|k} - \tilde{c}_j \leq 0$ *is guaranteed to be satisfied.*

5.2.3 Improving the Probability of Recursive Feasibility in Moment-based MPC

In this section we discuss the probabilistic guarantees for recursive feasibility in moment-based MPC for systems with additive Gaussian perturbations. Here we show an approach which, for a nominally feasible initial point x_0, decreases the chance of having an infeasible MPC problem over time iterations. For this purpose we pose a terminal constraint set, denoted with $\mathcal{X}_{ter}$, and for brevity we define the terminal constraint set from bound constraints as,

$$\mathcal{X}_{ter} := \{x \in \mathbb{R}^{n_x} | [-b_{ter}^x]_i \leq [x]_i \leq [b_{ter}^x]_i, i = 1, \dots, n_x\}. \tag{5.20}$$

We also define the steady state covariance matrix Σ_{ss} which is the largest covariance matrix[5] and can be found from the solution of the matrix equality

$$\Sigma_{ss} = A\Sigma_{ss}A^\top + F\Sigma_w F^\top.$$

Furthermore we make use of a terminal controller K_f, which casts the nominal system Σ^{nom} to be positively invariant inside a set $\hat{\mathcal{X}}_{ter}$, i.e., $\bar{x}_{N_p+i} := (A + BK_f)^i \bar{x}_{N_p|k} \in \hat{\mathcal{X}}_{ter}$. Finally, we define a tightened terminal constraint set $\bar{\mathcal{X}}_{ter}$ from the steady state covariance matrix as,

$$\bar{\mathcal{X}}_{ter} = \mathcal{X}_{ter} \ominus \mathcal{E}(0, \Sigma_{ss}, D^*(\alpha_{ter})) = \mathcal{X}_{ter} \ominus \{D^*(\alpha_{ter}) L_{ss} \beta_{ter}\}, \tag{5.21}$$

where $\Sigma_{ss} = L_{ss}L_{ss}^\top$, $\beta_{ter}^\top \beta_{ter} = 1$ and the Pontryagin difference operation between two sets $\mathcal{X}, \mathcal{Y}$, assuming $\mathcal{Y} \subseteq \mathcal{X}$, is defined as

$$\mathcal{X} \ominus \mathcal{Y} := \{x | x + y \in \mathcal{X}, \forall y \in \mathcal{Y}\}.$$

Next we state the statistical recursive feasibility properties of closed-loop systems with moment-based MPC.

[5] The monotonic increase of the covariance matrices can be seen from

$$\Sigma_{x_{j|k}} = \sum_{i=0}^{j-1} A^i F \Sigma_w F^\top {A^i}^\top = A\Sigma_{x_{j-1|k}} A^\top + F\Sigma_w F^\top.$$

Lemma 5.2.3 *Consider a linear dynamical system Σ^{add} as in Equation* (5.1b) *with terminal constraints $\mathcal{X}_{ter}$ as in Equation* (5.20)*, given the initial state x_k, the chance levels α_{ter} and a terminal controller K_f. If there exists a tightened constraint set $\hat{\mathcal{X}}_{ter}$ such that (i) $\hat{\mathcal{X}}_{ter} \subset \bar{\mathcal{X}}_{ter}$, (ii) the nominal state $\bar{x}_{N_p|k}$ is positively invariant with controller K_f within the set $\hat{\mathcal{X}}_{ter}$, then satisfying the constraint*

$$\bar{x}_{N_p|k} \in \hat{\mathcal{X}}_{ter} \tag{5.22}$$

guarantees

$$\mathbb{P}\{[-b^x_{ter}]_i \leq [x_{N_p|k+1}]_i \leq [b^x_{ter}]_i\} \geq \alpha_{ter}.$$

Proof. By hypothesis, $\bar{\mathcal{X}}_{ter}$ in Equation (5.21) and Lemma 5.2.1, we know that

$$\mathbb{P}\{[-b^x_{ter}]_i \leq [x_{N_p|k|}]_i \leq [b^x_{ter}]_i\} \geq \alpha_{ter}.$$

Then, we guarantee probabilistic feasibility at time $k+1$. Consider the terminal controller as being active in the subsequent time iterations as the suboptimal but feasible control action. Due to the positive invariance of nominal states inside $\hat{\mathcal{X}}_{ter}$, i.e., $\bar{x}_{N_p+j|k} \in \hat{\mathcal{X}}_{ter} \subset \bar{\mathcal{X}}_{ter}$. Furthermore, since $\Sigma_{x_{N_p+j|k}} \leq \Sigma_{ss}$ we have $\mathbb{P}\{x_{N_p+j|k} \in \mathcal{X}_{ter}\} \geq \alpha_{ter}$, for all $j \geq 0$. Assuming the control action with controller K_f is still feasible in the time instant $k+1$[6], then $K_f x_{N_p|k}$ is a feasible control action for $u_{[0,N_p-1]|k+1}$ which leads to the recursive feasibility. □

5.2.4 Constraint Reformulations in Moment-based MPC for Additive Perturbations with Even and Bounded Distributions

In this section, we consider the linear dynamical system Σ^{add} which is affected in the state dynamics by additive perturbations which are not Gaussians. Instead we assume that the disturbances are distributed with even (symmetric with respect to zero) and bounded pdf characteristics. For brevity, we only consider bound type of constraints, while similar approaches can be derived for polytopic or terminal constraints also.

The difference between the Gaussian disturbances case and non-Gaussian but even and bounded distribution case is due to the fact that the Mahalanobis distance concept (which is developed for Gaussian distributions) does not carry for other pdfs. The weighted sum of Gaussian random variables construct another Gaussian random variable, however this is not generally true for weighted sums of generic random variables. Consider again the decomposition of the true state $x_{j|k}$ as $x_{j|k} = \bar{x}_{j|k} + \tilde{x}_{j|k}$. If the distribution of perturbations w_k is not Gaussian, then the uncertain state vector $\tilde{x}_{j|k}$ has a distribution which is difficult to parametrize in terms of system matrices or the statistics of the perturbations. However

[6] If the uncertainty realization is severe (the tails of Gaussian) then the controller can not guarantee constraint satisfaction.

the covariance matrix $\Sigma_{x_{j|k}}$ is easy to express, since

$$\begin{aligned} \tilde{x}_{j|k} &= \sum_{i=0}^{j-1} A^{j-1-i} F w_{i|k}, \;\; j \geq 1, \quad \tilde{x}_{0|k} = 0, \\ \Sigma_{x_{j|k}} &= \mathbb{E}\{ \sum_{i=0}^{j-1} A^{j-1-i} F w_{i|k} w_{i|k}^{\top} F^{\top} {A^{j-1-i}}^{\top} \}, \\ &= \sum_{i=0}^{j-1} A^{j-1-i} F \Sigma_{w_{i|k}} F^{\top} {A^{j-1-i}}^{\top}, \\ &= N_j \bar{\Sigma}_w N_j^{\top}, \end{aligned} \tag{5.23}$$

where $N_j := \begin{bmatrix} A^{j-1}F & \dots & AF & F \end{bmatrix}$ and $\bar{\Sigma}_w = \Sigma_w \otimes I_j$ in which the Kronecker product is denoted with $\otimes$. Next, we bound the spread of $\tilde{x}_{j|k}$ to provide deterministic guarantees on the constraint satisfaction properties by using the covariance matrix. For this purpose, one can still use the approaches discussed in the previous section by using the Mahalanobis distance, since as $j \gg 1$ and by the Central Limit Theorem ([16]), the distribution of $\tilde{x}_{j|k}$ converges to a Gaussian distribution with zero mean and covariance matrix $\Sigma_{x_{j|k}}$. However for exact results a covariance dependent distance should be explicitly constructed which further depends on the pdf characteristics of the random variables and their weighted convolutions with each other as the prediction stages increases, such as the uniform or triangular pdfs which construct Irwin-Hall distribution ([151, 166]). Instead, here we provide an approximation approach for satisfying the constraints with deterministic guarantees, in the case of generic even and bounded distributions.

We start with approximating the domain of random vector $x_{j|k}$. The domain of the random vector $\tilde{x}_{j|k}$ is given by the row absolute sum of the matrix N_j, assuming that the components of disturbances $[w_{j|k}]_i$ are distributed in the domain $[-1, 1]$[7]. For the case where $n_w = 1$, the bounds on the true state are given by

$$[\mathcal{B}_j]_l := \sum_{i=0}^{j-1} |[A^i F]_l|.$$

By using the components of $\mathcal{B}_j$ we present two different constraint handling methods for deterministic guarantees. The first approach, highlighted in Lemma 5.2.4, is using the (vector) norm equivalence relations, hence prone to high conservatism. The second proposed approach, presented in Lemma 5.2.5, is bounding the uncertain state trajectories domain through least squares problems.

Lemma 5.2.4 *Consider a stochastic dynamical system Σ^{add} in Equation* (5.1b) *affected with random variables $w_k,\ k \in \mathbb{Z}_{\geq 0}$ that are distributed within $[-1, 1]$ and bound constraints as in Equation* (5.3)*. If*

$$[-b^x]_i + \left[\sqrt{\frac{1}{\sigma_w^2} j n_w \text{diag}(\Sigma_{x_{j|k}})} \right]_i \leq [\bar{x}_{j|k}]_i \leq [b^x]_i - \left[\sqrt{\frac{1}{\sigma_w^2} j n_w \text{diag}(\Sigma_{x_{j|k}})} \right]_i, \tag{5.24}$$

[7]The domain condition can be satisfied by scaling the components of F matrix.

is satisfied for $i = 1, \ldots, n_x$*, then*

$$[-b^x]_i \leq [x_{j|k}]_i \leq [b^x]_i$$

is guaranteed to be satisfied.

Proof. First we state $\Sigma_{x_{j|k}} = N_j \bar{\Sigma}_w^j N_j^\top$. If, for brevity, we assume $\bar{\Sigma}_w^j = \sigma_w^2 I_{jn_w}$, then $\Sigma_{x_{j|k}} = \sigma_w^2 N_j I_{jn_w} N_j^\top$. Now denote the row vectors of matrix N_j as $n_{jl}^\top,\ l = 1, \ldots, n_x$, i.e.,

$$N_j = \begin{bmatrix} n_{j1} & n_{j2} & \ldots & n_{jn_x} \end{bmatrix}^\top.$$

Then the realizations of $\tilde{x}_{j|k}$ satisfy $\tilde{x}_{j|k} \in [-\mathcal{B}_j, \mathcal{B}_j]$, while

$$[\mathcal{B}_j]_l = ||n_{jl}||_1,\ l = 1, 2, \ldots, n_x.$$

Next observe that $||n_{jl}||_2^2$ is given by the l^{th} element of the diagonal of the covariance matrix $\Sigma_{x_{j|k}}$, i.e.,

$$||n_{jl}||_2^2 = \frac{1}{\sigma_w^2}[\text{diag}(\Sigma_{x_{j|k}})]_l.$$

Then we use the (vector) norm equivalence between the 1- and 2-norms, which expresses the bounds in vector form as

$$\begin{array}{rl} \|n_{jl}\|_1 \leq & \sqrt{jn_w}\|n_{jl}\|_2 \\ [\mathcal{B}_j]_l \leq & \sqrt{\frac{1}{\sigma_w^2} jn_w\ [diag(\Sigma_{x_{j|k}})]_l}. \end{array} \tag{5.25}$$

Using this relation it is easy to show that

$$\left[\bar{x}_j - \sqrt{\frac{1}{\sigma_w^2} jn_w\ diag(\Sigma_{x_j})}\right]_l \leq [x_j]_l \leq \left[\bar{x}_j + \sqrt{\frac{1}{\sigma_w^2} jn_w\ diag(\Sigma_{x_{j|k}})}\right]_l,$$

which guarantees the constraint satisfaction for any realization of $w_k \in [-1, 1]^{n_w}$ and linear in the optimization variables. □

Remark 5.2.3 *By using Lemma 5.2.4 and different distributions one can guarantee the constraint satisfaction for bounded disturbances without explicitly calculating the reachable intervals in the expense of induced conservatism due to norm equivalence relations.*

One can make use of the variance values of known even and bounded distributions, such as the uniform ($\sigma_w^2 = 1/3$) or triangular ($\sigma_w^2 = 1/18$) pdfs. Furthermore one can also apply the discussed technique to the cases where the pdf is unbounded but in practice it is safe to assume that there is an upper bound on the realizations, such as the truncated Gaussian pdfs.

The second technique is structurally similar with the previous approach, whereas the conservatism is reduced by a least squares (LS) problem.

Lemma 5.2.5 *Consider a stochastic dynamical system Σ^{add} in Equation* (5.1b) *affected with random variables $w_k,\ k \in \mathbb{Z}_{\geq 0}$ that are distributed within $[-1, 1]$ and bound constraints as in Equation* (5.3)*. Given v_j^* as the solution of least squares problem*

$$\begin{aligned} v_j^* = \arg\min_{v_j \in \mathbb{R}^{n_x}} \quad & \|\Sigma_{x_{j|k}} v_j - \mathcal{B}_j\|_2, \\ \text{subject to} \quad & [\Sigma_{x_{j|k}} v_j]_i \geq [\mathcal{B}_j]_i, i = 1, \ldots, n_x. \end{aligned} \tag{5.26}$$

Then,

$$[-b^x]_i \leq [x_{j|k}]_i \leq [b^x]_i$$

is guaranteed to be satisfied, if

$$[-b^x]_i + [\Sigma_{x_{j|k}} v_j^*]_i \leq [\bar{x}_{j|k}]_i \leq [b^x]_i - [\Sigma_{x_{j|k}} v_j^*]_i, \tag{5.27}$$

is satisfied.

Proof. With the feasible solution of the LS problem given in Equation (5.26), we have

$$[\bar{x}_{j|k} - \Sigma_{x_{j|k}} v_j^*]_i \leq [\bar{x}_{j|k} - \mathcal{B}_j]_i \leq [x_{j|k}]_i \leq [\bar{x}_{j|k} + \mathcal{B}_j]_i \leq [\bar{x}_{j|k} + \Sigma_{x_{j|k}} v_j^*]_i,$$

which yields the tightened constraints

$$[-b^x]_i + [\Sigma_{x_{j|k}} v_j^*]_i \leq [\bar{x}_{j|k}]_i \leq [b^x]_i -]\Sigma_{x_{j|k}} v_j^*]_i.$$

□

5.3 Constraint Reformulation in Moment-based MPC for Linear Systems with Multiplicative Uncertainty

In this section, we consider the constraint satisfaction properties for the plant-model mismatch case. Here we assume that the dynamical evolution of the process is modelled as in $\Sigma_{(\cdot)}^{mul}$ of Equations (5.1c) and (5.1d). First we present the results for the Σ_{TV}^{mul} case, where the uncertainties affecting the system dynamics are known to be varying over time.

5.3.1 Time-Varying Multiplicative Uncertainty and Constraint Handling in Moment-based MPC

We start our discussion with the bound constraints as in Equation (5.3), which is in the form of

$$[-b_j^x]_i \leq [x_{j|k}]_i \leq [b_j^x]_i.$$

Similar to the dynamics with additive perturbations case, i.e., Lemma 5.2.1, first we consider only the mean of the constraint relations for Σ_{TV}^{mul}. If the predicted state vector $x_{j|k}$ is expressed in terms of $x_{0|k}, u_{[0,j-1]|k}, \delta_{[0,j-1]|k}, \gamma_{[0,j-1]|k}$ as in Equation (5.2c) and the expectation operator is evaluated, then the resulting predictions are found as

$$\mathbb{E}\left\{x_{j|k}\right\} = \mathbb{E}\left\{\prod_{i=0}^{j-1}(A(\delta_{i|k}))x_0 + \sum_{i=0}^{j-1}\prod_{l=i+1}^{j-1}(A(\delta_{l|k}))(B(\gamma_{i|k}))u_i\right\} = \bar{x}_{j|k},$$

assuming that there is no correlation between $\delta_{j|k}$ and $\gamma_{i|k}$ for all $i \in \mathbb{Z}_{[0,N_{p-1}]}$ and $j \in \mathbb{Z}_{[0,N_{p-1}]}$. Thus in the case of utilizing only the expectation operator, we end up with nominal state predictions to be used in the bound constraints. This reasoning is also valid for the polytopic constraint functions, since the uncertainties occur in affine form. If one considers quadratic type of constraints, i.e.,

$$x_{j|k}^{\top} H_j x_{j|k} \leq 0, \tag{5.28}$$

then the expectation of the constraint contains extra terms. We express the robust counterpart of quadratic constraints next.

Lemma 5.3.1 *Consider a stochastic dynamical system Σ_{TV}^{mul} in Equation* (5.1d) *affected with random variables δ_k and γ_k, $k \in \mathbb{Z}_{\geq 0}$. Given the quadratic constraints as in Equation* (5.28)*. Then*

$$\mathbb{E}\{x_{j|k}^{\top} H_j x_{j|k}\} = \bar{x}_{j|k}^{\top} H \bar{x}_{j|k} + \sum_{i=0}^{j-1} \sum_{l=1}^{n_\delta} \bar{x}_{i|k}^{\top} A_l^{\top} H_{j,i}^{TV} A_l \bar{x}_{i|k} \sigma_{[\delta]_l}^2 + \sum_{l=1}^{n_\gamma} u_{i|k}^{\top} B_l^{\top} H_{j,i}^{TV} B_l u_{i|k} \sigma_{[\gamma]_l}^2, \tag{5.29}$$

where

$$\begin{aligned} H_{j,j-1}^{TV} &= H, \\ H_{j,j-2}^{TV} &= \sum_{i=1}^{n_\delta} (\bar{A}^{\top} H \bar{A} + A_i^{\top} H A_i \sigma_{[\delta]_i}^2), \\ &= \sum_{i=1}^{n_\delta} (\rho_{\bar{A},A_i,1}^{\top} \tilde{H}_{j,1}^{TV} \rho_{\bar{A},A_i,1}), \\ H_{j,j-m}^{TV} &= \sum_{i=1}^{n_\delta} \rho_{\bar{A},A_i,m-1}^{\top} \tilde{H}_{j,m-1}^{TV} \rho_{\bar{A},A_i,2}, \end{aligned} \tag{5.30}$$

where

$$\begin{aligned} \rho_{\bar{A},A_i,j} &= \begin{bmatrix} \bar{A}^{j^{\top}} & A_i^{\top} \bar{A}^{j-1^{\top}} & \dots & A_i^{j^{\top}} \end{bmatrix}^{\top}, \\ \tilde{H}_{j,m}^{TV} &= (I_{2^m} \otimes H_j). \end{aligned}$$

Proof. The proof follows from the expressing $x_{j|k}$ in terms of the nominal predictions and input actions as

$$x_{j|k} = \bar{x}_{j|k} + \sum_{i=1}^{j} \sum_{l=1}^{n_\delta} \rho_{\bar{A},A_l,i-1}^{+} A_l [\delta]_l \bar{x}_{j-i|k} + \sum_{l=1}^{n_\gamma} \rho_{\bar{A},A_l,i-1}^{+} B_l [\gamma]_l u_{j-i|k}, \tag{5.31}$$

where

$$\begin{aligned} \rho_{\bar{A},A_l,0}^{+} &= I, \\ \rho_{\bar{A},A_l,1}^{+} &= (\bar{A} + A_l [\delta]_l), \\ \rho_{\bar{A},A_l,2}^{+} &= (\bar{A}\bar{A} + \bar{A} A_{l_1} [\delta_k]_{l_1} + A_{l_2} [\delta_k]_{l_2} \bar{A} + A_{l_2} [\delta_k]_{l_2} A_{l_1} [\delta_k]_{l_1}), \\ \rho_{\bar{A},A_l,j}^{+} &= \sum_{i=1}^{2^j} \left[\rho_{\bar{A},A_l,j}\right]_i \bar{\delta}, \end{aligned}$$

where the summation the last expression is conducted over $n_x \times n_x$ block elements with obvious definitions $[\rho_{\bar{A},A_{l_i},j}]_i$ multiplied with $[\delta]_{l_i}$ elements which depends on the number of A_{l_i} terms in the expression that constructs the $\bar{\delta}$ vector. Next, we express the expectation of the quadratic form, which ends up as in Equation (5.29). □

Similar to the additive case (Equation (5.5)), to back-off the predicted true state from the constraint region, and to include the effects of model uncertainty into the constraint satisfaction guarantees, we incorporate the second order terms of state in the constraint equations for the affine constraint functions of the state, i.e.,

$$\begin{aligned}[\bar{x}_{j|k} + \lambda_j \mathbb{D}\{x_{j|k}\}]_i \leq \quad [b_j^x]_i,\\ [\bar{x}_{j|k} - \lambda_j \mathbb{D}\{x_{j|k}\}]_i \geq \quad [-b_j^x]_i.\end{aligned} \tag{5.32}$$

If one expresses the variance of the predicted state as $\mathbb{D}\{x_{j|k}\} = \mathbb{E}\{(x_{j|k} - \mathbb{E}\{x_{j|k}\})(x_{j|k} - \mathbb{E}\{x_{j|k}\})^\top\}$, then we observe that the covariance values of different components of $\bar{x}_{j|k}$, i.e., $[\bar{x}_{j|k}]_i$, are effecting the magnitude of constraint tightening. Instead here we consider a conservative approximation of the variances, by incorporating the sum of variances in the constraint tightening step, i.e.,

$$\begin{aligned}[\bar{x}_{j|k} + \lambda_j \tilde{\mathbb{D}}\{x_{j|k}\}]_i \leq \quad [b_j^x]_i,\\ [\bar{x}_{j|k} - \lambda_j \tilde{\mathbb{D}}\{x_{j|k}\}]_i \geq \quad [-b_j^x]_i,\end{aligned}$$

where $\tilde{\mathbb{D}}\{x_{j|k}\} := \mathbb{E}\{\tilde{x}_{j|k}^\top \tilde{x}_{j|k}\}$ which transforms the bound constraints in Equation (5.32) into a tractable form. We provide explicit expression for $\tilde{\mathbb{D}}\{x_{j|k}\}$ next.

Lemma 5.3.2 *Consider a stochastic dynamical system Σ_{TV}^{mul} in Equation* (5.1d) *affected with random variables δ_k and γ_k, $k \in \mathbb{Z}_{\geq 0}$. Then*

$$\tilde{\mathbb{D}}\{x_{j|k}\} = \sum_{i=0}^{j-1} \sum_{l=1}^{n_\delta} \bar{x}_{i|k}^\top A_l^\top I_{j,i}^{TV} A_l \bar{x}_{i|k} \sigma_{[\delta]_l}^2 + \sum_{l=1}^{n_\gamma} u_{i|k}^\top B_l^\top I_{j,i}^{TV} B_l u_{i|k} \sigma_{[\gamma]_{l_1}}^2, \tag{5.33}$$

where $I_{j,i}^{TV}$ are found from Equation (5.30) *by replacing H_j with I_{n_x}.*

Proof. The claim follows from Lemma 5.3.1 as a special instance. □

5.3.2 Time-Invariant Multiplicative Uncertainty and Constraint Handling in Moment-based MPC

In this section, we present the results for the Σ_{TI}^{mul} case, where the uncertainties affecting the system dynamics are known to be invariant over time. First we present the results for the polytopic constraints[8], which is in the form of

$$A_j x_{j|k} \leq b_j^x. \tag{5.34}$$

[8]Here we skip the bound constraints case, since bound constraints are a special case of polytopic constraints

Similar to the dynamics with additive perturbations case, i.e., Lemma 5.2.2, first we consider only the mean of the constraints in Equation (5.34) for the dynamics Σ_{TI}^{mul}. In this case the predicted state vector $x_{j|k}$ is expressed in terms of $x_{0|k}, u_{[0,j-1]|k}, \delta, \gamma$ as in Equation (5.2c) and we evaluate the expectation operator, which results in

$$\mathbb{E}\left\{x_{j|k}\right\} = \bar{x}_{j|k} + \sum_{i=2}^{j}\sum_{l=1}^{n_\delta} \rho^{+,TI}_{\bar{A},A_l,i-1} A_l \bar{x}_{j-i},$$

where $\rho^{+,TI}_{\bar{A},A_l,i}$ is constructed from the expression in Equation (5.31), but in this case the predictions of Σ_{TI}^{mul} are used and the summed terms consist of even multiples of A_l, so that the expectation operation does not vanishes, i.e.,

$$\begin{aligned}\mathbb{E}\left\{x_{j|k}\right\} = &\quad \bar{x}_{j|k} + \sum_{i=2}^{j}\sum_{l=1}^{n_\delta} \rho^{+,TI}_{\bar{A},A_l,i-1} A_l \bar{x}_{j-i},\\ = &\quad \bar{x}_{j|k} + \sum_{l=1}^{n_\delta} A_l A_l \sigma^2_{[\delta]_l} \bar{x}_{j-2} + \sum_{l=1}^{n_\delta} \left(A_l \bar{A} A_l + \bar{A} A_l A_l\right) \sigma^2_{[\delta]_l} \bar{x}_{j-3} + \ldots,\end{aligned}$$

Thus in the case of utilizing only the expectation operator, we end up with nominal state predictions over multiple stages to be used in the polytopic constraints. Then the robust counterpart of the polytopic constraint in Equation (5.34) is obtained as

$$\begin{aligned} \mathbb{E}\{A_j x_{j|k}\} &\leq b^x_j,\\ A_j\left(\bar{x}_{j|k} + \sum_{i=2}^{j}\sum_{l=1}^{n_\delta} \rho^{+,TI}_{\bar{A},A_l,i-1} A_l \bar{x}_{j-i}\right) &\leq b^x_j. \end{aligned} \tag{5.35}$$

Equation (5.35) formulates the mean reformulation of the polytopic constraints[9].

Next, we present the robust counterparts of quadratic constraints on the state, i.e.,

$$\mathbb{E}\left\{x^\top_{j|k} H_j x_{j|k}\right\} \leq 0. \tag{5.36}$$

We provide the explicit form of the mean reformulation of quadratic constraints under TI uncertainty in Lemma 5.3.3.

Lemma 5.3.3 *Consider a stochastic dynamical system Σ_{TI}^{mul} as in Equation* (5.1c) *affected with random variables δ and γ. Given the quadratic constraints as in Equation* (5.36)*. Then*

$$\begin{aligned}\mathbb{E}\{x^\top_{j|k} H_j x_{j|k}\} = &\quad \bar{x}^\top_{j|k} H_j \bar{x}_{j|k} + \sum_{i=0}^{j-2}\sum_{l=1}^{n_\delta} 2\bar{x}^\top_{j|k} H^{TI}_{j,j,i} A_l \bar{x}_{i|k} +\\ &\sum_{i=0}^{j-1}\sum_{l=1}^{n_\delta}\Big(\bar{x}^\top_{i|k} A_l^\top H^{TI}_{j,i} A_l \bar{x}_{i|k} +\\ &\ldots \sum_{m=0}^{i-1} 2\bar{x}^\top_{i|k} A_l^\top H^{TI}_{j,i,m} A_l \bar{x}_{m|k}\Big) +\\ &\sum_{i=0}^{j-1}\sum_{l_1=1}^{n_\gamma}\sum_{l_2=1}^{n_\delta}\Big(u^\top_{i|k} B_{l_1}^\top H^{TI}_{j,i} B_{l_1} u_{i|k} \sigma^2_{[\gamma]_{l_2}} +\\ &\ldots \sum_{m=0}^{i-1} 2u^\top_{i|k} B_l^\top H^{TI}_{j,i,m} B_l u_{m|k} \sigma^2_{[\gamma]_{l_2}}\Big),\end{aligned} \tag{5.37}$$

[9] By selecting $A_j^\top = \begin{bmatrix} I_{n_x} & -I_{n_x}\end{bmatrix}$, one can express the bound constraints case.

where

$$\begin{aligned} H^{TI}_{j,j-1} &= H_j, \\ H^{TI}_{j,j-2} &= \sum_{l=1}^{n_\delta} (\bar{A}^\top H_j \bar{A} + A_l^\top H_j A_l \sigma^2_{[\delta]_i}), \\ &= \sum_{l=1}^{n_\delta} \rho^\top_{\bar{A},A_l,1} \tilde{H}^{TI}_{j,1} \rho_{\bar{A},A_l,1} \sigma^2_{[\delta]_i}, \\ H^{TI}_{j,j-i} &= \sum_{l=1}^{n_\delta} \rho^\top_{\bar{A},A_l,i-1} \tilde{H}^{TI}_{j,i-1} \rho_{\bar{A},A_l,i-1}, \end{aligned} \tag{5.38a}$$

$$\begin{aligned} H^{TI}_{j,j,j-2} &= \sum_{l=1}^{n_\delta} H_j A_l \sigma^2_{[\delta]_m} = \rho^\top_{\bar{A},A_l,0} \tilde{H}^{TI}_{j,0,1} \rho_{\bar{A},A_l,1}, \\ H^{TI}_{j,j,j-3} &= \sum_{l=1}^{n_\delta} H_j (\bar{A} A_l + A_l \bar{A}) \sigma^2_{[\delta]_i}), \\ &= \sum_{i=1}^{n_\delta} \rho^\top_{\bar{A},A_l,0} \tilde{H}^{TI}_{j,0,2} \rho_{\bar{A},A_l,2}, \\ H^{TI}_{j,j,j-m} &= \sum_{l=1}^{n_\delta} \rho^\top_{\bar{A},A_l,0} \tilde{H}^{TI}_{j,0,m-1} \rho_{\bar{A},A_l,m-1} \end{aligned} \tag{5.38b}$$

$$\begin{aligned} H^{TI}_{j,j-1,j-m} &= \sum_{l=1}^{n_\delta} \rho^\top_{\bar{A},A_l,0} \tilde{H}^{TI}_{j,0,m-1} \rho_{\bar{A},A_l,m-1}, \\ H^{TI}_{j,j-t,j-m} &= \sum_{l=1}^{n_\delta} \rho^\top_{\bar{A},A_l,t-1} \tilde{H}^{TI}_{j,t-1,m-1} \rho_{\bar{A},A_l,m-1}, \end{aligned} \tag{5.38c}$$

where

$$\rho_{\bar{A},A_l,i} = \begin{bmatrix} \bar{A}^{i^\top} & A_l^\top \bar{A}^{i-1^\top} & \dots & A_l^{i^\top} \end{bmatrix}^\top,$$

and $\tilde{H}^{TI}_{j,i}$ and $\tilde{H}^{TI}_{j,i,k}$ are symmetric matrices constructed from concatenation of H_j and zero matrices multiplied with the high order moments of the uncertainty $[\delta]_l$ for the even order combinations of A_l.

Proof. The proof follows from expressing $x_{j|k}$ in terms of the nominal predictions and input actions as

$$x_{j|k} = \bar{x}_{j|k} + \sum_{i=1}^{j} \sum_{l=1}^{n_\delta} \rho^+_{\bar{A},A_l,i-1} A_l [\delta]_l \bar{x}_{j-i|k} + \sum_{l_2=1}^{n_\delta} \sum_{l_1=1}^{n_\gamma} \rho^+_{\bar{A},A_{l_2},i-1} B_{l_1} [\gamma]_{l_1} u_{j-i|k},$$

similar to the proof of Lemma 5.3.1. Then we evaluate the expectation of the quadratic form which ends up as in Equation (5.37). □

For the dynamics Σ^{mul}_{TI} case, the expectation operator introduces extra quadratic elements for the robust counterpart problem. One can also implement the mean-variance formulation of constraint function, i.e., assuming the constraint is a bound constraint,

$$\begin{aligned} [\mathbb{E}\{x_{j|k}\} + \lambda_j \mathbb{D}\{x_{j|k}\}]_i \leq & \quad [b^x_j]_i, \\ [\mathbb{E}\{x_{j|k}\} - \lambda_j \mathbb{D}\{x_{j|k}\}]_i \geq & \quad [-b^x_j]_i. \end{aligned} \tag{5.39}$$

We already have the expected state, which is given as

$$\mathbb{E}\{x_{j|k}\} = \bar{x}_{j|k} + \sum_{i=2}^{j}\sum_{l=1}^{n_\delta} \rho^{+,\mathbb{E}}_{\bar{A},A_l,i-1} A_l \bar{x}_{j-i} + \sum_{i=2}^{j}\sum_{l_1=1}^{n_\gamma}\sum_{l_2=1}^{n_\delta} \rho^{+,\mathbb{E}}_{\bar{A},A_{l_2},i-1} B_{l_1} u_{j-i}.$$

This leads to computational difficulties if one evaluates $\tilde{\mathbb{D}}\{x_{j|k}\} := \mathbb{E}\{\tilde{x}^\top_{j|k}\tilde{x}_{j|k}\}$. Thus, here we transform the bound constraints in Equation (5.39) into a tractable form by considering $\mathbb{E}\{x^\top_{j|k}x_{j|k}\}$ and provide an explicit expression for this term in Lemma 5.3.4.

Lemma 5.3.4 *Consider a stochastic dynamical system Σ^{mul}_{TI} in Equation* (5.1c) *which is affected with random variables δ and γ. Then*

$$\begin{aligned}\mathbb{E}\{x^\top_{j|k}x_{j|k}\} = \;& \bar{x}^\top_{j|k}\bar{x}_{j|k} + \sum_{i=0}^{j-2} 2\bar{x}^\top_{j|k} I^{TI}_{j,j,j-i} A_l \bar{x}_{i|k} + \\ & \sum_{i=0}^{j-1}\sum_{l=1}^{n_\delta}\left(\bar{x}^\top_{i|k}A_l^\top I^{TI}_{j,i} A_l \bar{x}_{i|k} + 2\sum_{m=0}^{i-1}\bar{x}^\top_{i|k}A_l^\top I^{TI}_{j,i,l} A_l \bar{x}_{m|k}\right) + \\ & \sum_{i=0}^{j-1}\sum_{l=1}^{n_\gamma}\left(u^\top_{i|k}B_l^\top I^{TI}_{j,i} B_l u_{i|k} + 2\sum_{m=0}^{i-1} u^\top_{i|k}B_l^\top I_{j,i,m} u_{m|k}\right),\end{aligned} \tag{5.40}$$

where $I^{TI}_{j,i}$ and $I^{TI}_{j,i,m}$ are found from Equations (5.38a)*,* (5.38b)*,* (5.38c) *by replacing H_j with I_{n_x}.*

5.4 Conclusions on Robust Reformulation of Constraints in Moment-based MPC Problems

In this chapter, we have investigated the constraint satisfaction properties of a novel robust MPC algorithm which is based on the statistics (centralized moments) of the constraint functionals for linear stochastic systems with additive or multiplicative uncertainties. We have evaluated the mean and mean-variance of the state constraints in MPC problems. Furthermore, we extend our discussion towards the non-Gaussian disturbances and report two state covariance matrix based robust counterpart formulation of constraints for the mean and MV MPC problems for additive disturbances with even and bounded probability density functions. Lastly, we evaluate the mean of the MPC constraint functions, which are taken as bound, polytope and quadratic constraint functions, that incorporate time-varying and time-invariant uncertainties affecting the system dynamics in an affine fashion.

- All of the constraints in moment-based MPC problems can be expressed in terms of nominal dynamics with stage-wise varying weighting matrices over the prediction horizon.
- Nominal MPC problem decreases the computational complexity drastically in comparison to the other robust MPC techniques where an explicit maximization step or chance integrals need to be calculated for robust counterpart problem.

Chapter 6

Model-based Applications on Whey Protein Separation Process and Ultrafiltration Membranes

> The real price of everything is the toil and trouble of acquiring it.
>
> Adam Smith - The Wealth of Nations

In this chapter, and the rest of the dissertation, we present the research activities conducted on model-based applications towards an industrial whey protein separation process. First, we discuss the subprocesses, and their operation, in the whey protein separation plant. The ultrafiltration (UF) membrane unit operation, a subprocess in the separation process, receives greater emphasis in the discussion, due to the dynamical nature of its operation. We present a control relevant dynamical model of ultrafiltration membrane unit operation which contains both the fundamental balance laws and empirical relations among physical variables. Having such goals for the use of the rigorous model, we choose lumped models instead of a distributed parameter system formalism. Furthermore, we make use of physical laws instead of only black box components to model the separation phenomena and to describe physical variables in the model description explicitly. We test three different models, of different complexity in describing the membrane fouling, with respect to several statistical significance tests based on industrial data sets. Furthermore to test the predictive

[0] Substantial content of this chapter is also published or presented in:

- M.B. Saltik, L. Özkan, M. Jacobs, A. van der Padt. Dynamic modeling of ultrafiltration membranes for whey separation processes.
- M.B. Saltik, L. Özkan, M. Jacobs, A. van der Padt. Optimal Start-Up and Operation Policy for an Ultrafiltration Membrane Unit in Whey Separation.
- M.B. Saltik, L. Özkan, S. Weiland and P.M.J. Van den Hof. Optimal Sensor Selection Problem for Membrane Separation Systems.

capabilities, we construct various operation scenarios (constant inlet flow or pressure, inclusion of a fresh membrane etc.) and comment on the responses in the flows, concentrations and resistance development to demonstrate that the developed model has the capability of mimicking an UF membrane process and therefore can be used in online control applications. Furthermore, a method to address which signal to measure in a simplified UF membrane process is presented. Lastly operational aspects of UF membranes are considered and we compare a novel operating strategy with the current way of operation in the industrial process.

6.1 Whey Protein Separation Process

6.1.1 Motivation for Model-based Operations towards Whey Protein Separation Process

In Chapter 1, recent developments on OMBA strategies such as RTO or MPC have been discussed. Both of these operation strategies, see Figure 1.2a, have become a standard in the oil and bulk chemicals process industries ([103]). However food industry is yet to make use of process models as frequent as these industries. In the light of this observation, incorporation of economic and process models which are cast as constraints in the decision making process is expected to lead to improve the operation in various aspects such as efficiency or reliability. Since model-based applications exploit the nonlinear and/or time-varying structure of the processes in the plant, better predictions are expected to be generated. These predictions are then leading to better decisions before or during the operation of plant. If the predictions match up with the true process in a satisfactory level, then savings and efficiency can be improved by diverse set of tools, such as monitoring or advanced process control.

Steady state rigorous models are used in offline or economic applications, but incorporating rigorous models into online monitoring and/or control applications remains as a challenging task. Firstly, dynamic rigorous model development and validation is challenging and costly, even if the end goal is to use it for offline purposes. Furthermore a modeling campaign requires many iterations and a cooperative effort across many domains of 'expertise' in the case of industrial plants.

The challenge of using dynamic rigorous models in real-time has some other aspects. One of the important aspects is the computational complexity of the proposed application. A rigorous model mimics the true process with high resolution power under the condition that there is enough computational power at hand to simulate or manipulate model variables. The reason for numerical difficulties is due to the combination of slow and fast dynamics. The intersection of dynamical (slow) evolution of process variables with algebraic (fast) coupling between these variables lead to models with differential algebraic equations (DAEs). These type of models are exceptionally difficult to simulate reliably and quickly. This leads to extra constraints on the possible monitoring or control methods that can be implemented. With the current tools for model-based operation, many highly efficient methods can only work with small scale linear ordinary differential equation models, while these simple models can not provide a satisfactory level of realism.

Hence the first step in model development for process industries is to decide on the end goal of the model. For this purpose, in the next section, we present the whey protein separation process and the UOs on which it is built.

6.1.2 *Description and Operation of Whey Protein Separation Process*

In this dissertation we consider a specific whey protein separation setup. This plant consists of a series of unit operations (UOs) namely, reverse osmosis membrane units, pasteurization unit, the ultrafiltration membrane units, evaporator unit and lastly spray drying unit, as visualized in Figure 6.1. The whey process receives (thin) whey as the inlet. Whey is a water

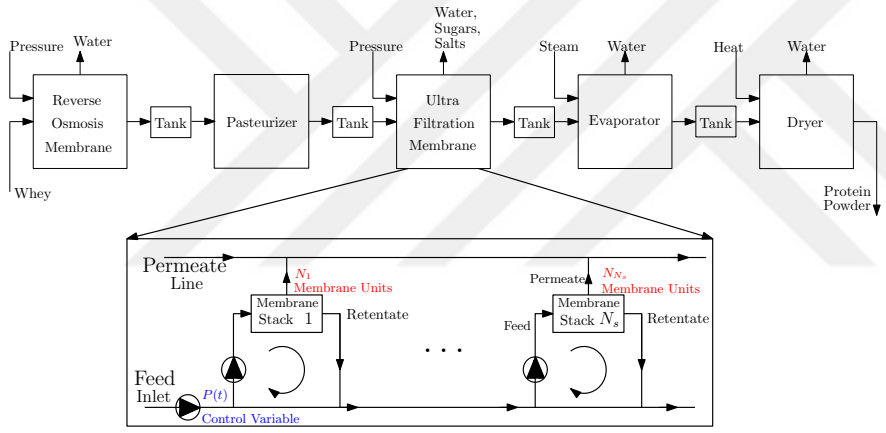


Figure 6.1: Block diagrams indicating UO inlet/outlets in whey separation process and a close look to ultrafilration membranes.

based mixture of organic molecules such as whey proteins, various type of sugars and fats with residual other components. Although small variations in the component concentrations of inlet whey are expected over time, these variations are assumed to be small process disturbances instead of high and effective fluctuations. The inlet is first processed in reverse osmosis (RO) membranes and pasteurizer UOs. RO membranes are membrane (separation) units with small sized pores and operating with high transmembrane pressures (TMPs), allowing small molecules, mainly only water, to reach to one of the two outlets, the permeate stream. With RO UO, the protein concentration in thin whey is increased by a factor of two. The pasteurizer UO gathers the outlet of ROs and buffers the concentrated whey till the desired levels of pasteurization is reached. No dynamical effect is expected with regard to whey components within this UO. After the pasteurization of the whey, which now contains about $10 - 12\%$ protein concentration on dry matter, the outlet is sent to the ultrafiltration (UF) membrane unit. A series of UF membrane stacks are used to separate whey components; the proteins in whey are separated from the sugars, fats, salts. Furthermore, the water concentration of the whey is also regulated in this unit operation. A schematic of a (spiral) membrane unit is shown in Figure 6.2. After the organic components are separated in UF membranes, the percentage of whey protein on the dry matter increases to the levels around $35 - 40\%$.

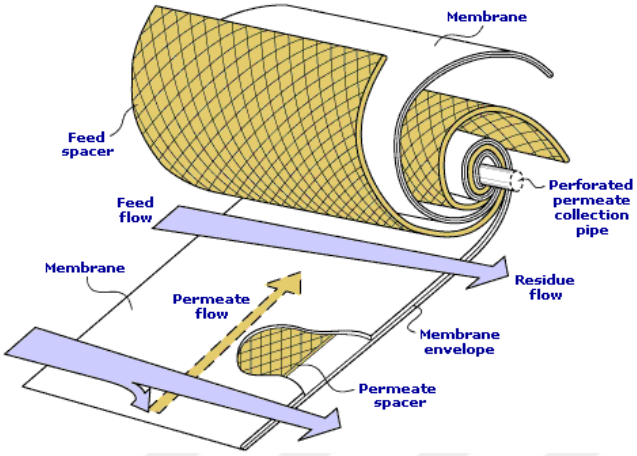


Figure 6.2: Membrane composition and product flows, [353].

The whey filtration operation continues with evaporator and dryer UOs. These UOs are large-scale units which are used to separate only the water to obtain protein powder products. To decrease the water concentration in the processed whey, the retentate (protein rich outlet) and permeate (outlet with other organic components) streams of UF membranes are processed in the evaporator UO. The evaporated permeate, now mainly consists of sugars of different types, is one of the outputs of the plant, whereas the retentate of UF membrane is used in both evaporator and spray drier UOs that turn thick whey into protein powder.

6.1.3 Ultrafiltration Membrane Process and Its Operation

Within the whey protein separation process, we focus on the UF membrane unit operation, since the dominant (at the final product) and dynamical responses are observed in these units within the protein separation process. Membranes are pressure driven systems consisting of three ports, the feed stream, also called as the feed inlet, the retentate stream, the performance output, and the permeate stream, the controlled output. In each membrane stack, consisting of multiple membrane units, the material is circulated with high speed tangential to membrane surface, where separation between small molecule size and large molecule size components occur. Due to high circulation speed, the concentrations of components are assumed to be uniform over the membrane units in a membrane stack. The pressure difference between the feed and permeate ports creates a mass flow that is mainly consisting of water (with other small molecule components) while the large molecule components, such as whey proteins, can not pass through the membrane pores and hence either circulate again or leave the membrane from the retentate port.

Membrane filtration processes have gained significant popularity in various applications and industries over the last couple of decades. They operate in mild operating temperatures leading to energy savings, low heat load and do not require additional chemicals to separate the desired product. In the food and beverages industry alone, the total membrane market has been estimated to be worth 1.182 billion $ in 2008 ([344]). In the dairy industry, it is estimated that 75% of total ultrafiltration membranes are used to fractionate the whey

proteins ([118, 349]).

The performance requirement of a membrane separation process is controlling the concentrations and total mass flow of the retentate port. This is achieved by controlling the mass flux, or its integral, permeate mass flow. Despite the increasing number of applications, membrane units have a major bottleneck; the decline of permeate flux due to the membrane fouling. This phenomenon causes performance drop which is at first instance compensated by the increase of the baseline transmembrane pressure (TMP) and later by the addition of fresh (or clean) membranes. Eventually, the operation of membranes are interrupted completely for chemical cleaning in order to restore the original performance. Huge savings and high level of productivity are expected if the control and operational policy of these units are improved. To this end, modeling the dynamic behaviour of the membranes is essential.

The type of dynamic models describing membranes can range from very detailed distributed parameter models to lumped and data-driven models ([87, 99, 104, 130, 146, 172, 231, 244, 269]) with the former being the preferred choice. Mathematical models in the form of distributed parameters systems provide detailed knowledge which could be very useful especially in the design of new membrane processes. On the other hand, data driven models are easy to develop once the required experimental data is gathered. However, the validity of these type of models is limited and they lack information on the internal variables which provide an indication on the process performance.

An important decision in any modeling campaign is the end use of the model. Here, we are interested in developing a dynamical model for an ultrafiltration (UF) membrane unit which can be utilised in online model-based applications such as soft sensing and model-based control. Furthermore this model will allow the possibility to

- track fouling or membrane's performance deterioration over time that is due to several difficult-to-measure phenomena, such as the gel layer formation or the dynamic concentration polarization;
- regulate the concentration and mass flow of the retentate stream;
- make decisions when to add a fresh membrane or when to start chemical cleaning.

We classify currently available membrane models in the literature based on i) the end use and ii) how the fouling phenomenon is described. The first class consists of several model structures and sizes ranging from static or black box models, such as [269], to medium sized lumped models, [146], and finally to large-scale models considering the spatial dependency of pressure distribution in membranes units ([354]). For the case of data driven models ([130, 269]) in which the relations between the internal variables are formulated with regression functions, the model complexity is drastically decreased. At the other end of the spectrum are the large-scale models represented by partial differential equations, such as [55, 87, 305, 358]. A comparative study between the black box models and first principle models is presented in [99].

Membranes are used in various separation processes with different geometric structures (pore size etc.). This fact leads to quite different membrane models utilised for design or optimal operation purposes. The reverse osmosis (RO) membranes are used generally, but not only, for wastewater treatment processes, for which membrane model examples can

be found in [59, 60, 172]. For slightly larger pore size membranes (nanofiltration (NF)) the number of possible applications and hence the number of reported uses of membrane models is high. Here, we only mention [57, 58], for nanofiltraton membrane modeling campaigns. The next class of membranes are called as UF membranes. We observe that there are many case-specific membrane models as presented in[47, 104, 146, 156, 244], while only the publication [231] discusses a general purpose modelling approach.

The second classification topic is the fouling and how it is formulated in the model. There are multiple distinct phenomena such as gel layer formation ([254]) and/or concentration polarization ([48]) contributing to the membrane fouling and, hence, performance degradation. These phenomena are studied extensively in several publications ([87, 99]), while these mainly focus on one specific (or leading) phenomenon while leaving the other contributing factors out. Another fouling modeling approach, the membrane resistance concept which is lumping the previously mentioned effects, is also discussed in literature ([146]). In many cases multiple resistances (in series configuration) are modelling pore cloaking, concentration polarization or cake formation as distinct functions and eventually summed up to construct the total membrane resistance to couple the TMP to the membrane flux. More specifically, for RO membranes the fouling models are provided in [15, 48, 312] and for the membranes used in dairy product processes fouling is discussed in [157, 161, 167, 312]. The works [147] or [332] can be given as general surveys on membrane fouling and its modeling.

Offline operation analysis and optimization based application are also reported in literature. The closed or open loop performance assessment of membrane process operation presented in [372, 373, 374] or in [375]. Furthermore, optimal operation for membrane (batch) processes is studied extensively in [271, 272, 273] but see also [270] for a thorough discussion on membrane based operations. The cleaning aspect of membrane operation, occurring after the accumulation of (irreversible) fouling, is investigated in [257, 381]. Detailed surveys on membrane operation and the membrane models utilized for fouling can also be found in [63, 105, 283, 312, 320, 349].

6.2 A Grey-box Model for Ultrafiltration Membrane Processes

The objective of UF membrane modelling is to develop a dynamic model, with smallest possible complexity with sufficient predictive capabilities, that describes the input-output behaviour and the evolution of the performance variables. For this purpose, we have identified the following tasks:

• **The input-state-output structure:** Each membrane port can be described by the concentration of components (fats, sugars, proteins and water) in the whey flows, the energy hold-up and the pressure. Furthermore, we make use of the membrane resistance concept to model the evolution and the effect of fouling. The input variables of the membrane system are defined as the mass flow at the feed inlet, the operating pressure and the mass concentrations of components in the feed inlet, while the measured/measurable outputs are considered as the mass flow at the permeate port, the concentration and the mass flow of retentate port.

• **Performance variables and efficient control of these variables:** The main purpose of controlling UF membrane separation processes is i) to achieve desired concentration levels at the retentate port while ii) increasing the time window between two consecutive

(chemical or mechanical) cleaning actions, that are required to undo the effects of fouling. For this purpose a regulatory controller should be designed to keep the retentate mass flow at a constant level through adjusting the permeate mass flow while keeping the membrane resistance, the indicator of fouling, with a minimum increase over time. This control is achieved by designing the feed mass flow and pressure profiles.

• **Computational complexity:** In literature membranes are generally modelled as distributed parameters systems whose solution can be computationally intensive. The current control and estimation techniques are, generally, not applicable to this class of models and in most cases require additional model approximation or model reduction steps. In order to achieve a considerably simpler model in comparison to the distributed models, a lumped model structure is selected. On the other hand, for high performance operation the internal or difficult-to-measure variables should be accessible to the controller, which eliminates the use of black box models. Thus a physics based modelling approach is taken with several simplifying assumptions, such as lumping the fouling using the membrane resistance concept. With these motivations in mind, we introduce a mathematical model consisting of the mass balance relations with the Darcy's Law, including a variable resistance term indicating the accumulation of fouling and its effect on the mass flux through the membrane. We use two membrane resistance variables. One of the variables indicates the membrane's intrinsic resistance to permeate flux, while the other one accumulates over operating window. The proposed model uses (simple) black-box components to distinguish individual component flows through the retention factors. Some in-depth physical aspects of membranes are left out of the model, since low level investigation (of polarization or cake formation etc.) is leading to more complex model, possibly with unidentifiable set of parameters from the measured data and lacking practical applicability for control purposes. Hence, a grey-box approach is taken to model the physical aspects dominating the operational (or control-relevant) behaviour while keeping the required computational complexity at an acceptable level for monitoring and control purposes.

6.2.1 Ultrafiltration Membrane Model Based on Physical Laws

The physical model consists of conservation laws, constitutive relations and the empirical relations between permeate mass flow as a function of applied pressure and the protein concentration at the retentate port.

Assumptions and Constitutive Relations

The pressure values for the ports of j^{th} membrane stack are assumed to be in the form of;

$$\begin{aligned} P_f^j(t) =& P^j(t), \\ P_r^j(t) =& P_f^j(t) - P_{drop}, \\ P_p^j(t) =& P_f^j(t) - \frac{P_{drop}}{2} - P_{TMP}^j(t) = P_{atm}, \end{aligned} \tag{6.1}$$

where P_f^j is the feed pressure which is equal to applied (base-line) pressure P^j, P_r^j is the retentate stream's pressure where P_{drop}^j is the pressure drop that is occurring over the j^{th} membrane stack that needs to be compensated over the membrane stacks. The pressure

drop over a membrane is assumed to be equal in all membranes for brevity. P_p^j is the permeate pressure and P_{TMP}^j is the TMP which is the driving force of mass flow in membrane separation systems. We assume that the permeate line has constant (atmospheric) pressure, P_{atm}. In the rest of the document we skip denoting the subscripts f, r, p with $(\cdot)$ wherever there is no dependency on the ports.

The densities of each stream, $\rho_{(\cdot)}^j$ (kg/m^3), are expressed as

$$\rho_{(\cdot)}^j(t) = g_\rho(T^j, P_{(\cdot)}^j, x_{(\cdot)}^{j,i}), \tag{6.2}$$

where $g_\rho(T, P, x)$ is a empirical (regression) function that is determined by the membrane and the whey properties. This empirical function is parametrized with the temperature, the operating pressure and the mass fractions of components at each port and can be constructed from data gathered during the operation.

Lastly, we assume that the temperature values throughout the system are equal to each other and set to $T_{(\cdot)}^j = 300$.

Balance Relations

The balance equations are derived from the physical laws, specifically mass balance relation. In the following we express these relations in the order of, the component mass balance, the total mass hold up and the coupling between the volume flow relations, respectively

$$\begin{aligned}
\dot{m}^{j,i}(t) =& F_f^j(t)x_f^{j,i}(t) - F_r^j(t)x_r^{j,i}(t) - F_p^j(t)x_p^{j,i}(t), \\
& i \in \mathcal{I}^{comp} := \{\text{Water, Fat, Protein}, \dots\}, \\
m_{Total}^j(t) =& \sum_{i \in \mathcal{I}^{comp}} m^{j,i}(t), \\
x_r^{j,i}(t) =& \frac{m^{j,i}(t)}{m_{Total}^j(t)}, \quad i \in \mathcal{I}^{comp}, \\
\frac{F_f^j(t)}{\rho_f^j(t)} =& \frac{F_r^j(t)}{\rho_r^j(t)} + \frac{F_p^j(t)}{\rho_p^j(t)},
\end{aligned} \tag{6.3}$$

where $m^{j,i}(t)$ (kg) is the mass of component i at the j^{th} membrane stack at time $t \in \mathbb{R}_{\geq 0}$, F_f^j, F_r^j, F_p^j (kg/s) are the mass flows of the feed, retentate and permeate ports, and $x_f^{j,i}, x_r^{j,i}, x_p^{j,i}$ (kg/kg) are the mass fractions of component i at these ports of j^{th} membrane stack, respectively. The variable m_{Total}^j denotes the total mass holdup in the j^{th} membrane stack. Last relation couples the volume flows of each port by assuming constant volume of material inside the membrane units.

Lastly, we define the performance indicator or so-called volume-reduction-factor (VRF) as,

$$VRF^j = \frac{\frac{F_f^j(t)}{\rho_f^j(t)}}{\frac{F_r^j(t)}{\rho_r^j(t)}} = 1 + \frac{F_p^j}{F_r^j}\frac{\rho_r^j}{\rho_p^j} \approx 1 + \frac{F_p^j}{F_r^j}. \tag{6.4}$$

Membrane Flux and Retention Factor Descriptions

An essential part of the membrane modeling is to express the membrane mass flux (to permeate port) in terms of operating conditions and the fouling development. In the case of UF, it has been observed that:

- For a wide range of pressure, the mass flux is linearly dependent on the TMP ($\frac{\partial J}{\partial P} = c$, for low P and c is constant).
- Protein concentration introduces an asymptotic limiting flux, as a pressure independent maximum value ($\frac{\partial J}{\partial P} \to 0$, as P grows).
- At low fluxes, or low pressures, the mass flux limits to the clean water flux.
- As the operation time increases, fouling dominates the process, i.e., the mass flux decreases monotonically, eventually requiring a cleaning action.

Here we make use of the membrane resistance concept to couple the operating TMP to the membrane mass flux through Darcy's Law. The total membrane resistance is assumed to consist of two parts, a(n) static (or empirical) resistance and a dynamic resistance. The static membrane resistance, which does not refer to static over time but to the static properties of the membrane unit itself, captures the following points, due to the physical structure of membranes:

- The flux is initially linear with the TMP and has an initial value of zero for zero pressure.
- The flux approaches a maximum value as the pressure increases and has zero slope at this value. This maximum flux value is dependent on the operating value of the retentate protein mass fraction.

A typical relation for static membrane resistance is visualised in Figure 6.3 ([13]). This behaviour captures the membrane specific flux-TMP-component concentration relation, which is occurring much faster than the fouling development. In the literature this relation is modeled via the assumptions on the operating point (either the linear region or the saturated region), while, to the best of our knowledge, there is not an equation describing this nonlinear relation in full complexity. In order to include this behaviour we characterize the flux-TMP-protein concentration relations with nonlinear empirical functions.

Furthermore, the dynamic resistance equation is derived considering the following observations:

- In order to incorporate concentration polarization effects, as the retentate mass fraction of proteins increases, the dynamic resistance increases;
- The increased transmembrane flux increases the so-called pore blocking and hence the fouling resistance increases with the flux;
- Lastly, the transmembrane pressure is also assumed to be increasing the dynamic resistance over time, due to cake formation phenomenon.

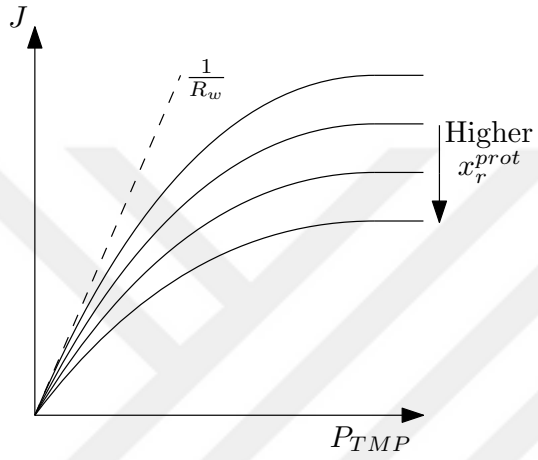


Figure 6.3: Static membrane resistance and its dependence on the transmembrane pressure and retentate protein mass fraction.

These items lead to the following construction for permeate mass flow, as a function of membrane flux, operating pressure and protein mass fraction of retentate stream:

$$\begin{aligned} F_p^j(t) =& \sum_{s=1}^{N_s^j} A_s^j J_s^j(t), \\ J_s^j =& \frac{P_{TMP}^j(t)}{R^j(t)}, \\ R^j(t) =& R_{(\cdot)}^{j,emp}(t) + R^{j,dyn}(t), \\ R_{(\cdot)}^{j,emp}(t) =& f_{emp}(x_r^j(t), P_{TMP}^j), \\ \dot{R}^{j,dyn}(t) =& \alpha^j P_{TMP}^j(t)^\beta J^j(t) x_r^{j,prot}(t), \end{aligned} \tag{6.5}$$

where N_s^j is the number of membranes in the membrane stack j, A_s is the membrane area of s^{th} membrane in stack j, J_s^j is the membrane flux of s^{th} membrane in stack j, which is assumed to be equal over the s^{th} membrane, R^j is the total (or overall) membrane resistance of j^{th} membrane that is the summation of two resistances, where, $R^{j,dyn}$ is the dynamic resistance which is standing for fouling, while the $R^{j,emp}$ is the static membrane resistance, membrane specific flux-pressure-mass fraction relation of the j^{th} membrane stack. We use the relations defined in [146] to construct the $R^{j,dyn}$ with the initial condition for the j^{th} membrane being R_0^j. The parameters α^j, β^j are membrane specific parameters that shape the long-term behaviour of membrane while $f_{emp}(\cdot)$ is a function specific to each membrane, mapping the operating conditions to membrane flux.

First we comment on the so-called long term resistance $R^{j,dyn}$ that is introduced to model the deterioration of membrane permeate flux over time. The dynamic membrane

resistance is a function of the TMP, the permeate flux and the retentate protein mass fraction as indicated in [146]. By the used convention, these variables are always positive, assuming α positive, hence the fouling resistance is always increasing. However the membranes show a fast change in the flux behaviour depending on the feed concentration or TMP. This fast change in flux can not be captured only with $R^{j,dyn}$, which necessitates the modeling of static membrane resistance, through nonlinear empirical functions. If there would be no fouling effect, one could track static membrane resistance function by applying different TMP and feed streams with different concentrations.

We have introduced several different functions that are empirically fitted to the theoretically expected membrane flux versus pressure and concentration relation. These different functions are named as;

1. Polynomial function,
2. Logarithmic-exponential function,
3. Exponential function.

The *polynomial empirical resistance* term $R^{j,emp}_{poly}$ is constructed from a bias term and depends on the mass fraction of retentate and the transmembrane pressure, which is expressed as;

$$R^{j,emp}_{poly} := \frac{P^j_{TMP}}{a^j_0 + a^j_1 x^j_r(t) + a^j_2 P^j_{TMP}(t)}. \tag{6.6}$$

This function can be also seen as the first order approximation of membrane flux characteristic function as visualized in Figure 6.3.

The second empirical function is the *logarithmic-exponential static resistance* function, which is parameterized as;

$$R^{j,emp}_{log} := \frac{P^j_{TMP}}{\mathcal{J}^j}, \quad \mathcal{J}^j(t) = a^j \left(1 - \frac{\log\left(\exp\left(\frac{b^j - P^j_{TMP}}{c^j}\right) + 1\right)}{\log\left(\exp\left(\frac{b^j}{c^j}\right) + 1\right)} \right), \tag{6.7}$$

where the parameters a^j, b^j, c^j summarizes the membrane flux characteristics. The parameters $a^j,\ b^j$ are further decomposed to;

$$a^j := \frac{1}{a^j_0 + a^j_1 x^{j,prot}_r}, \quad b^j := \frac{a^j}{R^j_w}. \tag{6.8}$$

These parameters are constructed from the physical properties of membranes, i.e.,

- The parameter a^j is the maximum permeate flux for a steady state retentate protein mass fraction, under no fouling assumption.
- The parameter b^j is the so-called transition pressure, where the membrane switches from linear dependence on TMP to saturated flux region.
- The parameter R^j_w is approximately equal to clean water resistance.

- The parameter c^j is the transition width, that is the pressure band where the membrane flux is switching from linear region (or operating mode) to saturated region (see Figure 6.3).

The last empirical function is the *exponential empirical resistance*, denoted with $R_{exp}^{j,emp}$. This function is defined as;

$$R_{exp}^{j,emp} := \frac{P_{TMP}^j}{a^j - a^j \exp\left(\frac{-P_{TMP}^j}{b^j}\right)}, \tag{6.9}$$

where a^j and b^j are the membrane specific parameters that are defining the saturation level and the initial slope, respectively. The parameter a^j is defined similar to the logarithmic-exponential case, as in Equation (6.8), hence still standing for the saturated permeate mass flux. However, the parameter b^j does not directly correspond to the transition pressure or the transition width in comparison to logarithmic-exponential case. Although the physical connection is lost, the parameter b^j still determines the slope (more rigorously decay factor) of the mass flux versus TMP relation. We remind that in all of the proposed methods the static resistance, and hence permeate mass flux, depends on TMP and retentate protein mass fraction.

The second coupling relation is the retention factors. Retention factors, denoted with $\mathcal{R}^{j,i}$, are determining what percentage of component $i \in \mathcal{I}^{comp}$ passes from feed side to permeate side. More rigorously, the definition is taken as,

$$\mathcal{R}^{j,i} := 1 - \frac{x_p^{j,i}}{x_r^{j,i}}.$$

According to the instantaneous operating conditions, the retention factors increase or decrease. In this model we make use of a first order linear fit for calculating the retention factors for each and every component in $\mathcal{I}^{comp}$ set, except water. The retention factors are expressed as a function of protein mass fraction in the retentate $x_r^{j,prot}(t)$ and the remaining two parameters of linear function are component specific constants, i.e.,

$$\begin{aligned} \mathcal{R}^{j,i}(t) =& a_{\mathcal{R}}^{j,i} x_r^{j,prot} + b_{\mathcal{R}}^{j,i}, \\ & \text{where } i \in \mathcal{I}^{comp}, \end{aligned} \tag{6.10}$$

where $a_{\mathcal{R}}^{j,i}$ is the (negative) parameter specific to i^{th}-component for the j^{th} membrane stack, which indicates that as the protein concentration increases in retentate, the retention factor decreases for all components. The second parameter $b_{\mathcal{R}}^{j,i}$ is the i^{th}-component's infinite dilution retention factor. Now we are able to express the permeate mass fraction $x_p^{j,i}$ as a function of retention factors, i.e.;

$$\begin{aligned} x_p^{j,i}(t) =& (1 - \mathcal{R}^{j,i}(t)) x_r^{j,i}(t), \quad i \in \tilde{\mathcal{I}}^{comp}, \\ x_p^{j,Water} =& 1 - \sum_{i \in \tilde{\mathcal{I}}^{comp}} x_p^{j,i}, \quad \tilde{\mathcal{I}}^{comp} := \mathcal{I}^{comp} \backslash \{'Water'\}, \end{aligned} \tag{6.11}$$

where the retention relation for water is separated from others to satisfy algebraic constraint regarding the mass fractions.

6.2.2 Estimation of Parameters in the Ultrafiltration Membrane Model

In this section, we estimate the parameters and validate the model through an industrial data set, obtained from a whey filtration facility. We first construct the experiment and then state the parameter estimation problem, and lastly provide the results for three different empirical resistance functions $R^{j,emp}_{(\cdot)}$.

Experimental Data

The experimental data consists of 2 different data sets of 20 hours of measurements of a whey separation process. The process contains 8 membrane stages, connected in series. Six of the membranes are operational at the start of the experiment while the remaining two membranes are switched on when the performance levels are not maintainable over the course of operation. The sampling rate of experiments is 0.08 hour per sample. The feed stream is assumed to be consisting of five components, water, protein, fat, sugar, salt. Furthermore, we assume that the experiment is conducted at 300° Kelvin. The pressure drop over one membrane and permeate pressure are assumed to be equal to

$$P_{drop} = 1.5\ bar, \qquad P_p = P_{atm} \approx 1\ bar.$$

The inputs to the membrane process are the base-line pressure and feed mass flow with known component concentrations. The measured outputs are the total permeate mass flow and retentate mass flow of the last membrane of the serially connected membrane stacks. There is a redundancy in the selection of output, hence we utilize the total permeate mass flow as the only measured output. Since there is no concentration or mass fraction measurement, the retention factors are not included in the analysis. Thus we assume the retention factor parameters to be equal to

$$b_{\mathcal{R}} = \{0, 1, 1, 0, 0\},$$

for water, protein, fat, sugar, salt, in the respective order, for all membranes. Furthermore, we assume that the parameter $a^{j,i}_{\mathcal{R}}$ is equal to 0 for all components, that is the retention factors are not effected by the concentration levels of retentate. These assumptions on the retention factors are required to be compensated through the membrane resistance relations, lack of introducing $a^{j,i}_{\mathcal{R}}$ is compensated with fouling resistance $R^{j,dyn}$ and $b^{j,i}_{\mathcal{R}}$ is compensated with $R^{j,emp}_{(\cdot)}$. Due to the lack of measurement data from individual membranes, we assume all of the internal parameters of membrane stacks are equal to each other.

Parameter Estimation Problem

The membranes are known to be full of water at the start of the experiment, hence for each membrane we consider the initial conditions of differential equation (6.3) as;

$$m^j(0) = V^j \rho^j_w x^j_r(0),\ j \in \{1, 2, \ldots, 8\},$$

where ρ^j_w denotes the water density, V^j is the volume of the membrane and

$$x^j_r(0) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix},\ j \in \{1, 2, \ldots, 8\}.$$

Lastly, we set the initial conditions for membrane resistance. We assume that the membranes are cleaned before the operation, hence at the initial time there is no effective fouling, i.e.,

$$R^{j,dyn}(0) = 0,\ j \in \{1, 2, \ldots, 8\}.$$

The parameters to be estimated are grouped into two. The first group is the set of parameters of the empirical resistance functions $R^{j,emp}_{(\cdot)}$. For each of the empirical functions defined in the previous section, we estimate a different set of parameters, such as;

- For $R^{j,emp}_{poly}$, the estimation variables are $a^j_0,\ a^j_1,\ a^j_2$.
- For $R^{j,emp}_{log}$, the estimation variables are $a^j_0,\ a^j_1,\ R^j_w,\ c^j$.
- For $R^{j,emp}_{exp}$, the estimation variables are $a^j_0,\ a^j_1,\ b^j$.

The second group of estimation parameters are the ones related to the dynamic resistance. The dynamic resistance $R^{j,dyn}$ is parametrized with two important parameters, α^j and β^j. However, the structure of the differential equation cast these parameters to be correlated with each other. Hence, we set β^j as 0.5, as in [146], to improve the accuracy of the optimization problem for all membranes. Thus, we only estimate the parameter α^j in the dynamic membrane resistance equation.

We use gPROMS modeling and optimization software to construct the differential algebraic equations model and to solve the parameter estimation problem, see [140][1].The (only) cost function that we considered is provided as,

$$(z(\theta)) = \frac{N}{2}\log(2\pi) + \frac{1}{2}\min_{\theta\in\Theta}\left\{\sum_{i=1}^{Ne}\sum_{i=1}^{Nv_i}\sum_{k=1}^{Nm_{ij}}\log(\sigma^2_{i,j,k}) + \frac{(\bar{z}_{i,j,k} - z_{i,j,k})^2}{\sigma^2_{i,j,k}}\right\},$$

where, N is the total number of measurements taken during all the experiments, θ is the set of model parameters to be estimated and Θ is the set where the parameters can take values. Furthermore Ne, Nv_i, and Nm_{ij} denotes the number of experiments, number of variables measured in the i^{th} experiment and Nm_{ij} is the measurement index of the j^{th} variable in the i^{th} experiment. Lastly, $\sigma_{i,j,k}$ denotes the variance of k^{th} measurement of variable j in experiment i, $z_{i,j,k}$ is the model prediction of variable j in experiment i and $\bar{z}$ is the k^{th} measured value of variable j of experiment i. With this cost function, we present the results of the parameter estimation algorithm in the next section.

Estimation Results and Statistical Analysis

In this section, we provide the parameter estimation results for three different models. We start with the so-called polynomial empirical membrane resistance $R^{j,emp}_{poly}$.

For the first type of empirical resistance function we have four different estimation parameters, three of them are a_0, a_1, a_2 which are describing $R^{j,emp}_{poly}$, given in Equation (6.6) without the stack dependency over the index j, i.e.,

$$R^{j,emp}_{poly} = \frac{P_{TMP}}{a_0 + a_1 x^j_r + a_2 P_{TMP}}.$$

[1]Since the optimization problem is constrained to the developed model, the DAE system needs to be integrated over time, which is the reason why we selected the mentioned software.

The remaining estimation parameter is the fouling coefficient, i.e., α. The assumption of having the same parameter values over membrane stacks is rarely true in practice, however the measurement data is not informative to distinguish each and every membrane's parameters. To be able to distinguish the parameters of the membranes one requires measurements from individual membranes, see also [318]. Hence in the rest of the dissertation, due to the stated assumption, the index j indicating the dependency on each membrane stack is not used for the equivalent parameters.

With the above stated construction, the results of the parameter estimation problem are:

$$\alpha = 5.44 * 10^{-5}, \;\; a_0 = 3.44 * 10^{-2}, \;\; a_1 = 3.11 * 10^{2}, \;\; a_2 = 3.17 * 10^{-2}, \qquad (6.12)$$

while the parametric uncertainty is described by the covariance matrix

$$\mathcal{C}^{poly} = \begin{bmatrix} 1.36 * 10^{-12} & * & * & * \\ -1.45 * 10^{-7} & 1.34 & * & * \\ -5.95 * 10^{-6} & -5.69 * 10^{1} & 2.97 * 10^{3} & * \\ 4.44 * 10^{-11} & -1.79 * 10^{-4} & 1.11 * 10^{-3} & 1.18 * 10^{-7} \end{bmatrix},$$

which yield the results for parameters with 95% confidence interval, $\alpha = (5.44 \pm 0.23) * 10^{-5}$, $a_0 = (3.44 \pm 227.8) * 10^{-2}$, $a_1 = (3.11 \pm 1.07) * 10^{2}$, $a_2 = (3.17 \pm 0.067) * 10^{-2}$.

In Figures 6.4-6.5 we visualize the experimental data, the feed and permeate mass flow measurements, and the predicted trajectory with the optimal parameter values for the three different models. The resulting statistical properties of the parameter estimation problem are tabulated next. In Equation (6.13) the correlation matrix of estimated parameters is given as

$$\mathbb{C}^{poly} = \begin{bmatrix} 1 & * & * & * \\ -0.108 & 1 & * & * \\ -0.0938 & -0.902 & 1 & * \\ 0.111 & -0.448 & 0.0594 & 1 \end{bmatrix}, \qquad (6.13)$$

The 95% t-values (with respect to 95% confidence intervals) of parameters are found as (for a reference t-value of 1.65) $\{23.74, 0.015, 2.912, 46.99\}$ indicating statistical significance problems associated with parameter a_0. The parity plot for this parameter estimation problem is visualized in Figure 6.6a while the lack-of-fit test result (for the optimal parameter values and resulting residuals) indicate an acceptable fit, with the weighted residual and 95%-χ^2 values being 447 and 543, respectively, where the weighted residual should be lower than the 95%-χ^2 value for the regression to be acceptable. Lastly the R^2 value is found as

$$R^2_{poly} = 1 - \frac{SSR}{SST} = 1 - \frac{223.5}{4259.2} = 0.95.$$

These statistical results indicate two major problems; i) the available data is not informative to estimate parameter a_0; ii) relatively high variance levels of parameters, mainly the ones related to R^{emp}_{poly}. This is expected since the polynomial empirical resistance is a first order approximation of the membrane flux characteristics.

The second model for which we estimate its parameters is the membrane resistance consisting of Equation (6.7), i.e., the logarithmic-exponential empirical resistance and the dynamical membrane resistance. The parameters which need to be estimated are; a_0, a_1, R_w, c

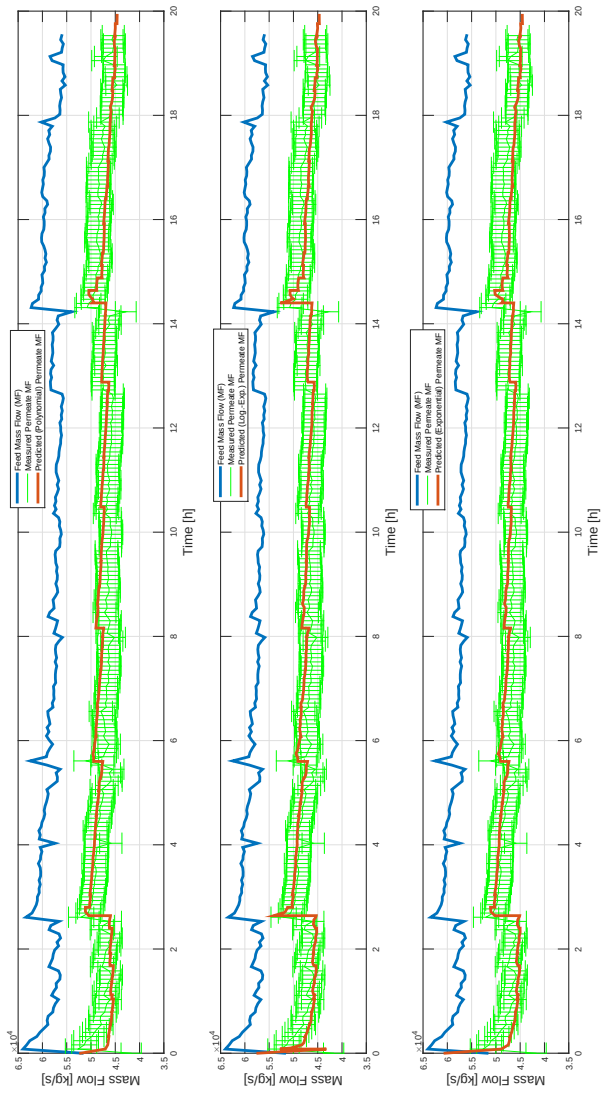


Figure 6.4: The first experimental and simulation data for three different models, upper, middle and lower, visualizing the feed mass flow and the permeate mass flow with the predicted permeate mass flow profiles for three different models.

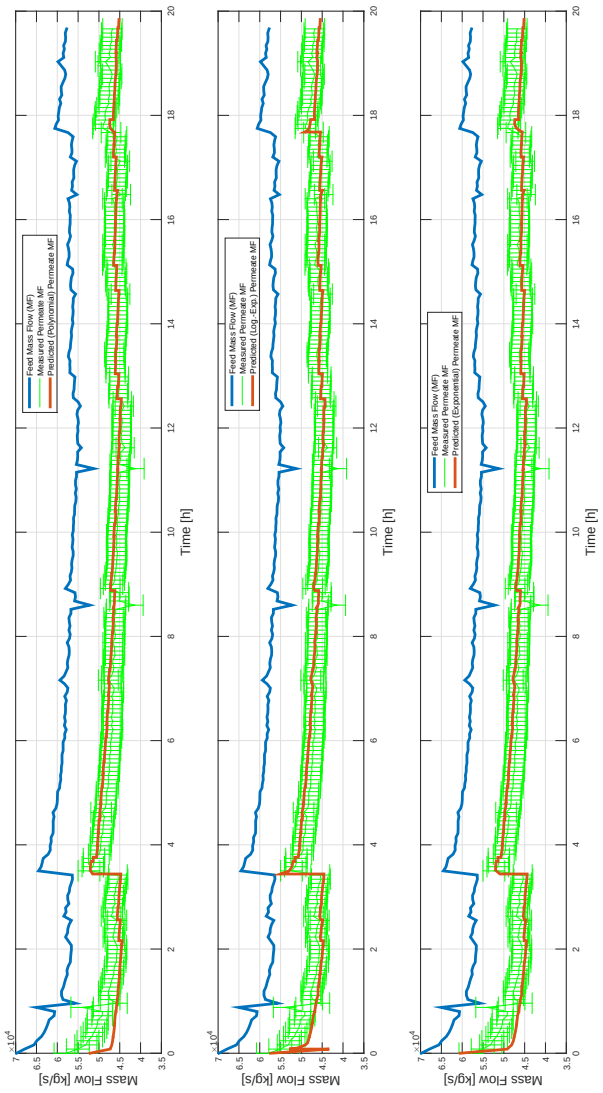


Figure 6.5: The output data and the simulation generated from the optimal parameter values for the second data set, visualizing the feed mass flow and the permeate mass flow with the predicted permeate mass flow profiles for three different models.

for $R_{log}^{j,emp}$ and α for $R^{j,dyn}$. However, the initial results indicate that there is a strong dependency between R_w and c, see Equation (6.7), hence we set c to be equal to our initial guess value 0.34, which is selected by making use of its physical interpretation and discussions with available membrane experts from *FrieslandCampina*. Conducting the parameter estimation algorithm under this assumption we obtain the following optimal values of parameters

$$\alpha = 5.1402 * 10^{-5}, \;\; a_0 = 2.35769 * 10^{-3}, \;\; a_1 = 4.92 * 10^{-1}, \;\; R_w = 99.96, \quad (6.14)$$

while the covariance matrix is found as

$$\mathcal{C}^{log} = \begin{bmatrix} 2.02 * 10^{-12} & * & * & * \\ 3.84 * 10^{-12} & 1.66 * 10^{-9} & * & * \\ -7.93 * 10^{-8} & -6.96 * 10^{-6} & 3.59 * 10^{-2} & * \\ 3.65 & 4.92 * 10^{1} & -1.04 * 10^{4} & 2.55 * 10^{12} \end{bmatrix},$$

which yield the results for parameters with 95% confidence interval as, $\alpha = (5.1402 \pm 0.28) * 10^{-5}$, $a_0 = (2.36 \pm 0.08) * 10^{-3}$, $a_1 = (4.92 \pm 0.37) * 10^{-1}$, $R_w = (0.0009 \pm 9.9) * 10^{6}$. The predictions generated from the optimal parameters are given in Figure 6.5. The statistical results of the parameter estimation problem are as follows, the correlation matrix is given in Equation (6.15), while the lack of fit test results indicate an acceptable fit, weighted residual and 95%-χ^2 values are reported as 262 vs. 543. The 95% t-values are found as $\{18.41, 29.42, 13.22, 1 * 10^{-5}\}$ (for the reference t-value of 1.65), which means, apart from the last parameter, the confidence to estimates is high. The estimate of the last parameter, R_w, has high variance attached to its optimal estimate, since the input-output data is gathered in the saturated operating region of membranes, thus not providing information about the linear operating region of the flux vs. TMP characteristics, see Figure 6.3. Lastly, the parity plot is given in Figure 6.6b and the R^2 test results are found as

$$R_{log}^2 = 1 - \frac{130.9}{4166.5} = 0.97.$$

$$\mathbb{C}^{log} = \begin{bmatrix} 1 & * & * & * \\ 0.0663 & 1 & * & * \\ -0.295 & -0.9 & 1 & * \\ 0.508 & 0.239 & -0.108 & 1 \end{bmatrix}. \quad (6.15)$$

Thirdly, we solve the parameter estimation problem for the exponential membrane model introduced as in Equation (6.9). The optimal estimates are found as

$$\alpha = 4.58 * 10^{-5}, \;\; a_0 = 2.70 * 10^{-3}, \;\; a_1 = 3.50 * 10^{-1}, \;\; b = 3.55 * 10^{-1}, \quad (6.16)$$

while the covariance matrix is found as

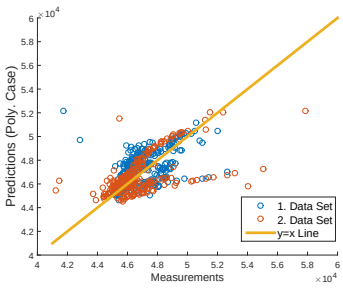
$$\mathcal{C}^{exp} = \begin{bmatrix} 1.4 * 10^{-12} & * & * & * \\ 3.66 * 10^{-11} & 1.53 * 10^{-9} & * & * \\ -5.09 * 10^{-9} & -6.15 * 10^{-6} & 3.18 * 10^{-2} & * \\ -1.08 * 10^{-9} & -1.39 * 10^{-8} & 1.84 * 10^{-4} & 3.51 * 10^{-5} \end{bmatrix}. \quad (6.17)$$

which yield the results for parameters with 95% confidence interval as, $\alpha = (4.58 \pm 0.23) * 10^{-5}$, $a_0 = (2.70 \pm 0.08) * 10^{-3}$, $a_1 = (3.50 \pm 0.35) * 10^{-1}$, $b = (3.55 \pm 0.37) * 10^{-1}$. Similarly, the predictions are visualized in Figure 6.5, the Equation (6.18) tabulates the correlation matrix results of the estimation problem, while the lack of fit test results (371 vs. 543) indicate that the predictions fit well to the measurements. The 95% t-values are found as $\{19.7, 35.13, 9.988, 9.638\}$ (in comparison to the reference t-value 1.65), the parity plot is given in Figure 6.6c and the R^2 test result is found as

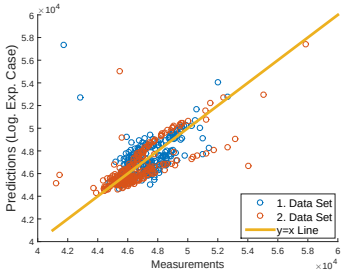
$$R^2_{exp} = 1 - \frac{185.4}{4221} = 0.96.$$

$$\mathbb{C}^{exp} = \begin{bmatrix} 1 & * & * & * \\ 0.0791 & 1 & * & * \\ -0.241 & -0.883 & 1 & * \\ -0.488 & -0.19 & -0.0551 & 1 \end{bmatrix}. \tag{6.18}$$

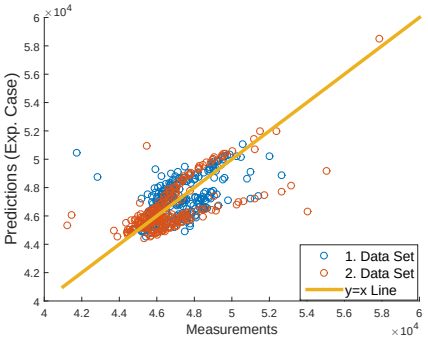
In order to compare the results obtained from these three different models, we calculate the



(a) Parity plot for polynomial case.



(b) Parity plot for logarithmic-exponential case.



(c) Parity plot for exponential case.

Figure 6.6: The parity plots visualizing the predicted and measured permeate mass flows for three different membrane models.

membrane flux versus the TMP and retentate mass fraction characteristics. These results are provided in Figure 6.7, where the concentration retentate mass fraction is decreasing from 0.024 to 0 (pure water case) in four equal steps. The curves that have larger flux values stand for smaller feed mass fractions. Although the flux vs. TMP curves do vary for the different models, the results indicate that the physical phenomenon is modeled similarly in each of the empirical resistance case and has close resemblance with Figure 6.3. The statistical results indicate that the best model out of the three proposed empirical resistance models is the exponential empirical membrane resistance case. The correlation values (Equation (6.18)) and the t-values are better than the other two cases, casting the uncertainty attached to the parameters to be low as presented in Equation (6.17). This result can be explained as both the polynomial empirical resistance, which does not reflect any physical prior knowledge about the system, and the logarithmic empirical resistance, which is parameterized according to the physical aspects of the membranes, has identifiability issues. The first case, polynomial empirical resistance, is not complex enough to describe the membrane resistance behaviour, while the highly coupled, complex relation induced from logarithmic empirical resistance case casts the parameters to be correlated with each other, hence increasing the uncertainty attached to optimal parameters (Section 6.2.2). However all of the identified models have high regression statistics, indicating that the input-output mapping is statistically significant. The exponential empirical resistance model outperforms the other two cases, with less and uncorrelated parameters and superior statistical properties, hence we make use of the membrane model with exponential empirical membrane resistance in the rest of the dissertation, unless explicitly stated otherwise.

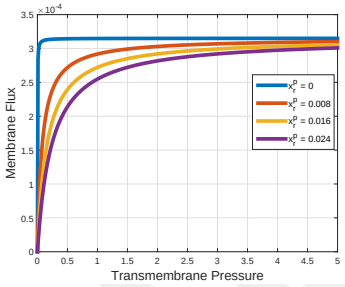
6.3 Offline Model-based Applications with Ultrafiltration Membrane Model

In this section we discuss three offline model-based applications cast on the UF membrane process. The first one is on the offline optimization studies conducted with the UF membrane model. Secondly, we discuss a novel operating strategy and design aspects of the UF membrane process. Lastly we discuss the sensor selection problem, which signals in UF membrane process to measure to deduce maximum information on the latent variables.

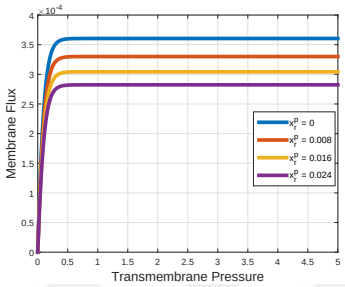
6.3.1 Optimal Operation and Scheduling of UF Membrane Stacks

We investigate the predictive capabilities of the dynamic model of the UF membrane unit and calculate optimal operation strategies using this model. To this end, we consider several fundamental operation scenarios listed as;

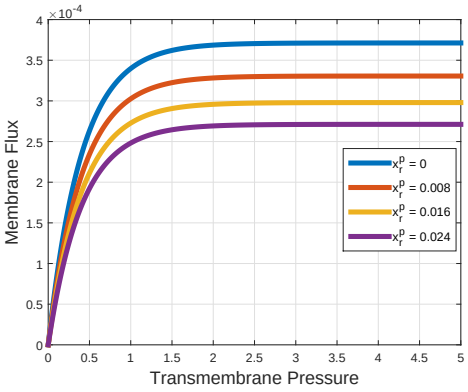
1. Addition of a clean membrane stack and the effect of concentration change in the feed stream;
2. Optimal scheduling and operation of membrane stacks;
3. Introduction of new membranes to the operation;
4. Constant feed mass flow operation;
5. Constant pressure operation;



(a) Flux vs. TMP graph for $R_{poly}^{j,emp}$.



(b) Flux vs. TMP graph for $R_{log}^{j,emp}$.



(c) Flux vs. TMP graph for $R_{exp}^{j,emp}$.

Figure 6.7: Flux versus transmembrane pressure and retentate protein mass fraction graphs.

6. Optimal feed and pressure profile design under accumulated performance specifications.

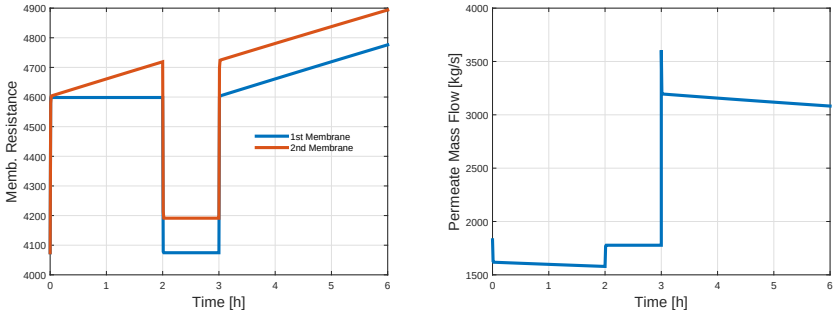
For all of the simulation studies, we use the membrane model with the exponential empirical resistance function, with the parameters given as in Equation (6.16). Furthermore, we assume that for all of the simulation studies, membranes are introduced to the operation being filled with water and cleaned initially, hence with zero initial condition for dynamic membrane resistance.

1) Addition of a clean membrane stack, the effect of concentration change in the feed stream and optimal pressure profiles: We study whether the UF model shows dynamic behaviour similar to the observations in practice for the case of concentration changes in the feed stream. Furthermore addition of a fresh membrane and its effects on the performance indicators are also simulated. To this end, we create a test case consisting of two membrane stacks with the following strategy, where t here stands for time;

- For $t < 2$h : The 2^{nd} membrane stack is operational, while the 1^{st} membrane stack is left non-operational.

- At $t = 2$h : A step change in dry matter mass fraction of feed stream is introduced, increasing the mass fraction of water from 0.904 to 0.99.
- At $t = 3$h : Feed stream is back to its original mass fraction of dry matter and the 1^{st} membrane is added to the process.

We design the simulation to observe the dependency of the process to the concentration of feed stream. After we introduce a new membrane to the process, we effectively increase the concentration of feed entering to the already operational (2^{nd}) membrane. This effect is tracked with the permeate mass flow values and fouling indicator, the membrane resistance. Throughout the experiment we set the feed mass flow and pressure to be constant. In Figure 6.8 we provide the membrane resistance and permeate mass flow (determines the VRF hence the performance) values over the simulation horizon. We observe that during the time window between $2 \leq t \leq 3$, where the protein concentration of feed is dropped, the membrane resistances are not increasing, since almost no fouling occurs for water dominant feed inlet. This observation can be explained by Equation (6.5) as the effect of lower protein mass fraction on the membrane resistances and the operating point. Furthermore, the permeate mass flow results show that the increased water permeation after the 1^{st} membrane is turned on, introducing an instantaneous increase in the permeate flow. Next, we present the



(a) The membrane resistance profiles for both of the membranes, effected by the mass fraction of the input.

(b) The total permeate flow, visualizing the dependence on the composition of the input and number of operational membranes.

Figure 6.8: The change in the behaviour of membrane with respect to changes in the feed concentration and including a new membrane to the process.

simulation results where the base-line pressure profiles are optimized, either with linearly parametrized or step-wise pressure profiles. This simulation study is done to show the effect of pressure profiles on the performance, the mass flow of permeate, and the fouling performance, the membrane resistance. The feed mass flow and concentration is kept constant over the simulation horizon. The pressure profiles and the results are visualized in Figure 6.9. As can be seen, the piecewise linear pressure case is able to track the desired permeate mass flow profile with a smoother profile, while the optimal step-wise pressure profile introduces undesired errors before and after the jumps in the pressure.

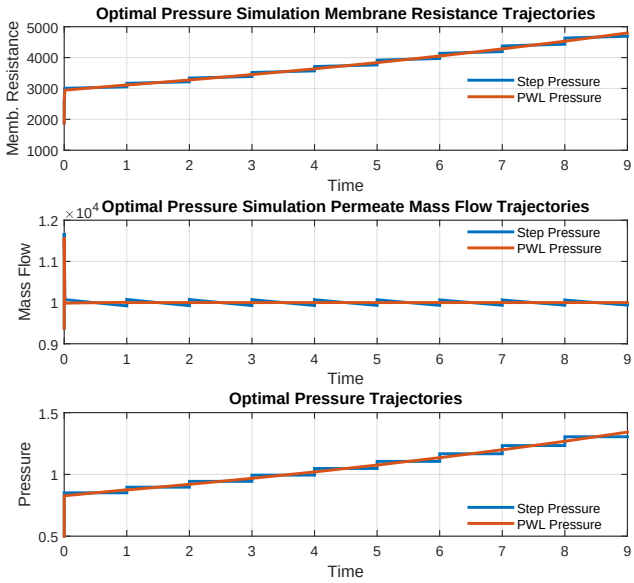


Figure 6.9: The open loop optimal pressure profiles and the simulation results or the trajectories obtained with optimal pressure profile.

2) Optimal scheduling and operation of membrane stacks:

We assume that the UF membrane process consists of $N_s = 6$ stacks of membranes, each denoted with the index i. We assume that the performance specifications are provided, thus the scheduler compares and optimizes over different scheduling strategies over similar performance levels. This allows us to investigate and slow down the membrane fouling induced deterioration due to different operation strategies, while guaranteeing satisfactory performance.

The optimization is conducted over both the on-off conditions of the membrane stacks, tracked with the scheduling variable $\sigma \in \mathbb{R}^{N_s}_{[0,1]}$ and the applied pressure profile under the scheduling strategy. The σ^i variable replaces the decision (integer) variable $\bar{\sigma}^i \in \{0, 1\}$ denoting the On-Off status of i^{th} membrane. We relax the decision variable in the optimization routine by replacing the non-smooth 0-1 decision with a smooth approximation, i.e.,

$$\bar{\sigma} = \frac{\sigma}{\sqrt{\sigma^2 + \epsilon_\sigma^2}}, \text{ for } \sigma \in \mathbb{R}^{N_s}_{[0,1]},$$

where ϵ_σ is a small relaxation constant. We assume that the applied pressure is a piecewise constant input sequence, in order to reduce the associated complexity of the optimization problem. Furthermore, the feed mass flow and concentration are considered as known disturbances.

We assume that the desired output values for the protein concentration at the last membrane and the total retentate yield is provided as performance specifications, given in following form of

$$\begin{aligned} |x_r^{N_s,prot}(\sigma,t) - x_r^{des}| \leq \epsilon_{x_r}, \quad t_0 \leq t \leq t_f, \\ Y_{lb} \leq Y_{ret} = \int_0^{t_f} F_r^{N_s}(\sigma,t)dt \leq Y_{ub}, \end{aligned} \tag{6.19}$$

where $x_r^{1,N_s}(\sigma,t)$ is the protein concentration at the retentate stream of N_s^{th} (final) membrane, x_r^{des} is the desired protein concentration level, ϵ_x is the allowed margin of protein concentration deviation from the desired concentration. Moreover the overall yield is required to be between the allowed lower and upper bounds, denoted with Y_{lb} and Y_{ub}, respectively.

Lastly we construct the cost function of the optimization problem. If the performance specifications are satisfied, then we claim that the operation can be improved by extending the operational time window. This can be done by slowing down the membrane fouling accumulation, which is adjusted by the scheduling strategy and the associated optimal pressure profile. Thus the cost function contains quadratic terms of the applied pressure, to increase the weight on the high operating pressures, and the membrane resistance, to slow down the effects of the membrane fouling, i.e.,

$$J = \int_{t_0}^{t_f} \Delta P(\sigma,t)^T Q_{\Delta P} \Delta P(\sigma,t) + \lambda R(\sigma,t)^T Q_R R(\sigma,t) dt$$

where $R(\sigma,t) \in \mathbb{R}_{>0}^{N_s}$ is the vector of membrane resistances and $\lambda \in \mathbb{R}_{\geq 0}$ is a weighting parameter.

Hence, the optimal membrane scheduling problem can be read as satisfying the performance specifications while decreasing the input energy and the accumulation of membrane resistance, i.e.,

$$\begin{aligned} (\sigma^*, \Delta P^*) = \quad & \arg \min_{\sigma \in \mathbb{R}^N, \Delta P} \int_{t_0}^{t_f} \Delta P(\sigma,t)^T Q_{\Delta P} \Delta P(\sigma,t) + \lambda R(\sigma,t)^T Q_R R(\sigma,t) dt, \\ \text{s.t.} \quad & \dot{x}(t) = f_d(x(t), z(t), u(t)), \;\; 0 = f_a(x(t), z(t), u(t)) \\ & y(t) = h(x(t), z(t), u(t)), \;\; t \in \mathbb{R}_{[t_0,t_f]}, \\ & |x_r^{N_s,prot}(\sigma,t) - x_r^{des}| \leq \epsilon_x, \;\; Y_{lb} \leq \int_0^{t_f} F_r^{N_s}(\sigma,t)dt \leq Y_{ub}. \end{aligned} \tag{6.20}$$

The feed trajectory $F_f^1(t)$ is given as in Figure (6.10) and feed protein mass fraction $x_f^{1,prot}$ is assumed to be 0.0045. The desired concentration and overall yield constraints' are taken as;

$$x_r^{des} = 4 * x_f^{1,prot}, \quad \epsilon_{x_r} = 0.1, \quad Y_{lb} = \frac{20}{100} Y_{feed}, \quad Y_{ub} = \frac{30}{100} Y_{feed}.$$

The simulation model and the optimization problem are implemented in gPROMS software environment. The optimization parameters are taken as

$$Q_{\Delta P} = 5, \quad Q_R = 10^{-3} I_6, \quad \lambda = 1, \quad \sigma_\epsilon = 0.001.$$

The results obtained from the the optimization procedure subject to the cost and the con-

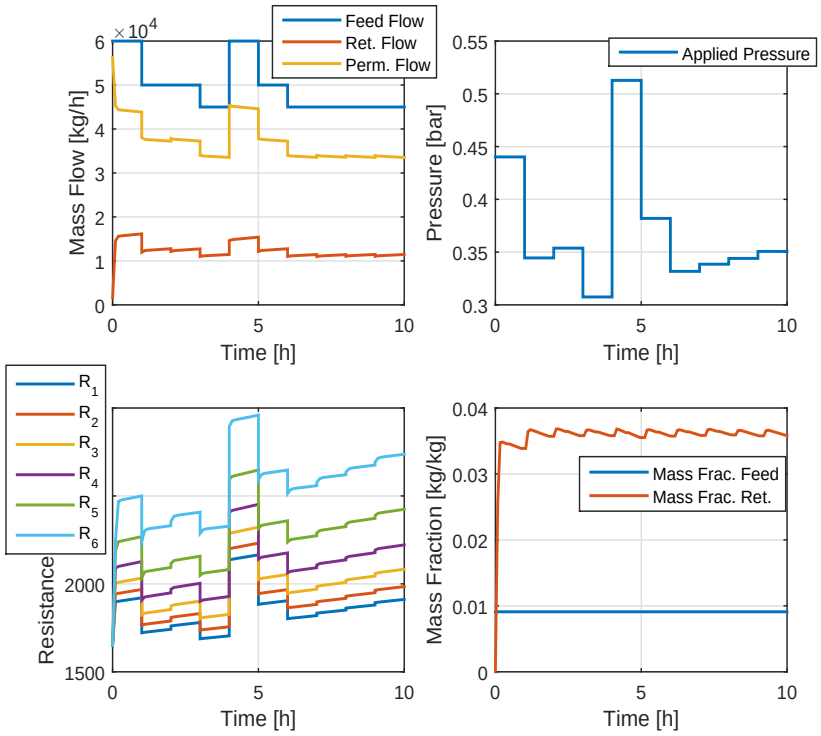


Figure 6.10: The simulation results for the optimal membrane scheduling strategy.

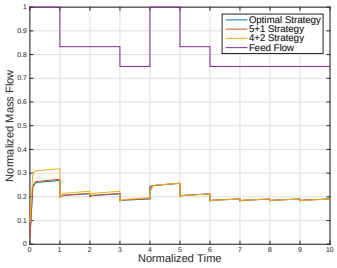
straints given in Equation (6.20) indicate that the best scheduling strategy is obtained by turning all of the membranes on at the beginning of the operation, i.e.,

$$\sigma_i(t) = 1, i = 1, 2, \ldots, 6,\ t_0 \leq t \leq t_f,$$

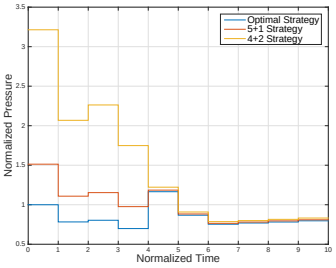
to distribute the permeate flow over all membrane stacks for the whole operation. Figure 6.10 shows that the optimal pressure profile can be lowered by operating more membranes to reach the performance specifications.

We also compare the optimal scheduling strategy and pressure profile with heuristic strategies. For this purpose we solve the optimal control problem for two different heuristic scheduling strategies:

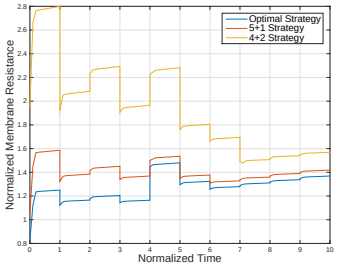
- 5+1 strategy, one membrane is added to five operating membranes at the 4^{th} time instant, when feed mass flow is increased, which mimics the common practice in the industrial applications.
- 4+2 strategy, two membranes are added to four membranes at the 4^{th} and 7^{th} time instants.



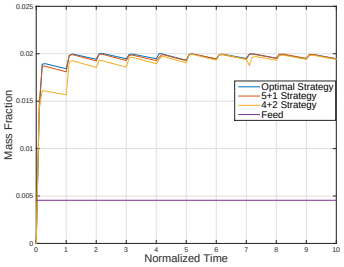
(a) Retentate mass flows.



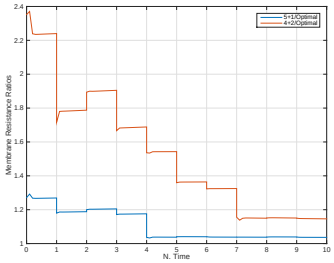
(b) Pressure profiles.



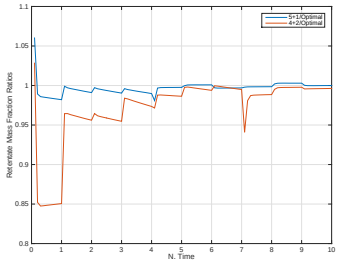
(c) Membrane resistance.



(d) Mass fraction.



(e) Ratio of membrane resistances.



(f) Ratio of mass fraction values.

Figure 6.11: The simulation results of optimal scheduling strategy compared with heuristic strategies.

For both of the cases, we obtain the optimal pressure profiles with given scheduling strategies. The main differences between these strategies can be seen in Figure (6.11). The optimal scheduling strategy (solution of Equation (6.20)) decreases the required pressure levels compared to other strategies which is also beneficial for the membrane fouling. The membrane fouling is increasing with the increased operating pressure, see Equation (6.5). Hence decreasing the pressure decreases the fouling accumulation, thus improving the operation. The differences between the membrane resistance trajectories of different strategies are presented in Figure 6.11c. The excess resistance of heuristic strategies are due to the

higher applied pressures at the beginning of the simulation. Lastly, the effect of adding a clean membrane into operation can be seen in Figure 6.11d at the start of the 7^{th} time instant, where the fresh membrane filled with water disturbs the mass fraction at the outlet.

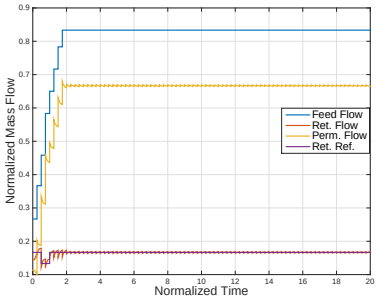
3) Introduction of new membranes to the operation: Similar to the previous studies, in this case we observe the effects of introducing membranes during run-time to the retentate mass flow and protein mass fraction, while the process is driven by the optimal operating pressure. Similarly we consider a setup consisting of eight membranes and in this case we introduce the membranes to the operation sequentially with 15 minute break in between two membranes during the beginning of the simulation. Since membranes are introduced to the operation filled with only water, i.e., the low protein mass fraction, the permeate mass flow has relatively large overshoots at each starting moment. It is desirable to suppress the overshoots at these time instants through control action. We assume that the feed mass flow trajectory is provided and we solve an optimization problem that minimizes the pressure values while achieving control specifications. The specifications are imposed as integral quadratic constraints in the optimization problem, i.e.,

$$\int_0^{t_f} (F_r^8(t) - F_r^{ref}(t))^2 dt \leq c_1, \qquad \int_0^{t_f} (x_r^{8,prot}(t) - x_{ref}^{prot}(t))^2 dt \leq c_2. \tag{6.21}$$

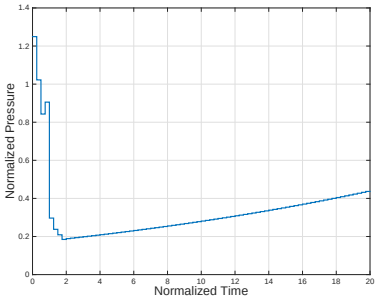
The resulting (inlet) feed, (outlet) permeate and (outlet) retentate mass flows are visualized in Figure 6.12a, while the optimal pressure profile is shown in Figure 6.12b. The protein mass fraction at the outlet, presented in Figure 6.12c, is highly effected during the initial phase where the fresh membranes are introduced one after another. However, once all the membranes are operational, the pressure values regulate the retentate flow and the protein content with high performance.

4) Constant feed mass flow operation: In this case we study the membrane unit operation in which the (inlet) feed mass flow is kept constant and the retentate mass flow and protein content are regulated. We consider eight membranes, all of which are operational during the whole simulation, and solve an optimization problem to optimize the pressure profile with performance constraints as given in Equation (6.21). The results are demonstrated in Figure 6.13a, which shows the given feed mass flow and resulting permeate and retentate mass flows; Figure 6.13b which demonstrates the resulting pressure trajectory; and lastly Figure 6.13c, which visualizes the protein mass fraction of feed and retentate streams. As expected for constant feed flow case, the regulation is obtained by increasing pressure values which adjusts the filtration to the permeate line to satisfy the specifications.

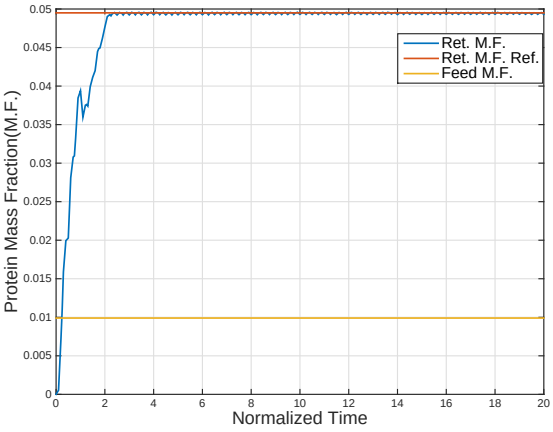
5) Constant pressure operation: In this case, we consider a case where the pressure is kept constant and the mass flows in the outlets are adjusted via the inlet feed mass flow. Again we consider eight membrane units, but we simulate this case for two different pressure values, a low operating pressure case, where the membranes are operating in the linear region of the flux-pressure graph (Figure 6.3), and a high operating pressure case, where membranes are operating in the saturated flux region. For both of the results we solve an optimization problem minimizing the feed mass flow entering to the process while satisfying the performance constraints in the form of Equation (6.21). We set the low pressure value to 0.341 normalized units and high operating pressure value to 4.54 normalized units. The results are demonstrated in Figure 6.14, in which Figure 6.14a and Figure 6.14b shows



(a) Mass flows.



(b) Optimal pressure profile.

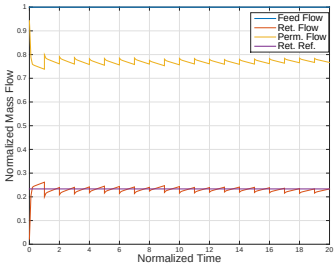


(c) Protein mass fraction trajectories.

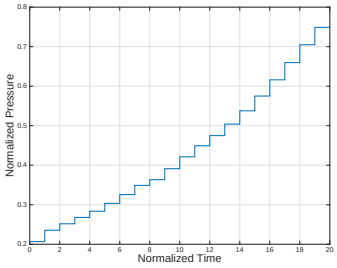
Figure 6.12: The normalized results of the sequential membrane scheduling strategy. (6.12a): Mass flows of the feed of first membrane, the retentate and the permeate of final membrane; (6.12b): The optimal pressure profile;(6.12c): The protein mass fractions in feed of first membrane, the retentate of final membrane and the reference value.

the optimal feed mass flow trajectory for low and high pressure case, respectively, while Figure 6.14c and Figure 6.14d visualizes the retentate protein mass fraction results for low and high pressure values. It is important to note that both of the optimization problems are ill-conditioned, resulting in many re-runs to satisfy the performance constraints which are relaxed (by increasing the c_1 and c_2 values in Equation (6.21)) after each failure. Since for constant pressure operation, the permeate flow is predetermined, satisfying both of the performance requirements is difficult, hence leading to poor performance in both cases, which can be observed in the resulting protein mass fraction profiles.

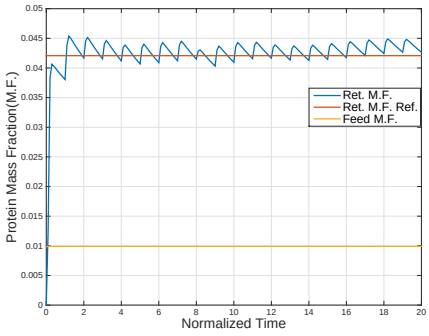
6) Optimal feed and pressure profile design under accumulated performance specifications: Lastly we consider a different specification in which the performance is not de-



(a) Mass flows.



(b) Optimal pressure profile.



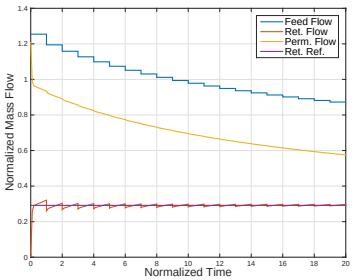
(c) Protein mass fraction trajectories.

Figure 6.13: The normalized results of the constant feed mass flow case. (6.13a): Mass flows of the feed of first membrane, the retentate (both real and reference) and the permeate of final membrane; (6.13b): The optimal pressure profile; (6.13c): The protein mass fractions in feed of first membrane, the retentate of final membrane and the reference value.

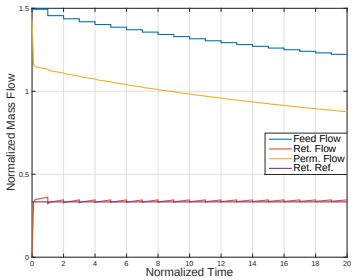
fined instantaneously but the overall product (yield) should satisfy the given specification. We let the feed mass flow and pressure be variable to regulate the retentate mass flow while the protein mass fraction of the accumulated retentate mass flow (overall yield) should be close to the desired levels at the end of operation. For this case we construct an optimization problem which minimizes both the feed mass flow and pressure values for the constraints

$$\int_0^{t_f} (F_r^8(t) - F_r^{ref}(t))^2 dt \leq c_1, \qquad \left| \frac{\int_0^{t_f} x_r^{8,prot} F_r^8 dt}{\int_0^{t_f} F_r^8 dt} - x_{ref}^{prot} \right| \leq c_2. \tag{6.22}$$

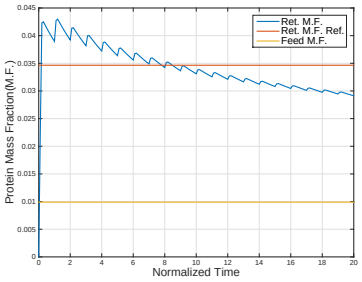
The second constraint is shaped in such a way that the total protein mass fraction of the yield at the end of the simulation (hence a point-wise constraint) is close to the reference protein mass fraction, while during the operation there is no constraint to the cast the protein mass fraction to track the specification instantaneously. The results are visualized in Figure 6.15a, Figure 6.15b and Figure 6.15c which show the mass flows, pressure and protein mass



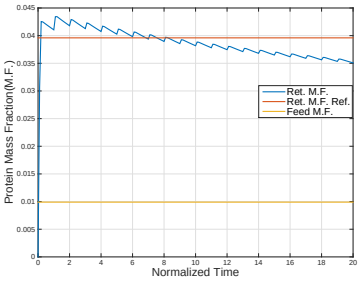
(a) Mass flows for low pressure operation.



(b) Mass flows for high pressure operation.



(c) Protein mass fractions for low pressure operation.



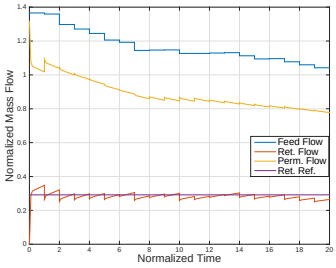
(d) Protein mass fractions for high pressure operation.

Figure 6.14: The normalized results of the constant (low and high) pressure case. (6.14a): Mass flows of the feed, the retentate (both real and reference) and the permeate of final membrane for low pressure case; (6.14b): Mass flows of the feed, the retentate (both real and reference) and the permeate of final membrane for high pressure case; (6.14c): The protein mass fractions in the feed, the retentate of final membrane and the reference value for low pressure case; (6.14d): The protein mass fractions in the feed, the retentate of final membrane and the reference value for high pressure case

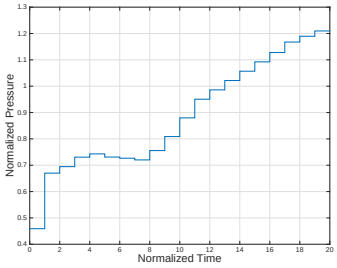
fractions, respectively. The resulting profiles for the performance variables (retentate mass flow and protein mass fraction) demonstrate that the specifications are met. However the optimization problem is highly ill-conditioned, which is due to the dependency of the controlled variables on both of the feed and the pressure values. In order to obtain meaningful results, we solved multiple optimization problems where we iteratively fixed one of the feed or pressure values while leaving the other one free to be optimized over.

6.3.2 Exploratory Study on Dynamic Operation of UF Membrane Stacks

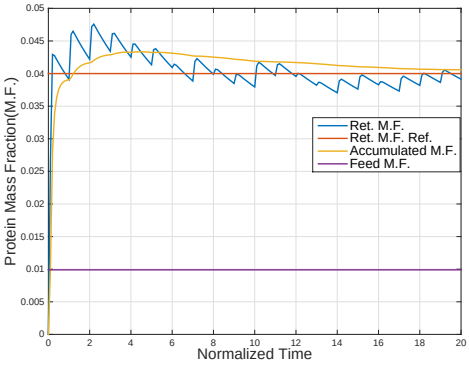
The fouling of the membrane is a difficult modelling task due to the multiple phenomena happening. The fouling is expressed by the membrane resistance concept in the model, which builds up over time, see Equation (6.5). In the resistance equations, the TMP has a



(a) Mass flows.



(b) Optimal pressure profile.



(c) Protein mass fraction trajectories.

Figure 6.15: The normalized results of the constant feed mass flow case. (6.15a): Mass flows of the feed, the retentate (both real and reference) and the permeate of final membrane; (6.15b): The optimal pressure profile; (6.15c): The protein mass fractions in the the feed, the retentate of final membrane and the reference value.

strong influence, which is determined by the pressure before and after the UF membrane stack. If the TMP values are varied between the loops, then it is possible to manipulate the development of fouling at each membrane stack separately. In this chapter the possible gains by variable TMP's per membrane stack are discussed.

Resistance build-up and Variable Transmembrane Pressure

In the UF process the first membrane stack removes a large quantity of permeate. The remaining retentate with a higher dry matter content flows to the next membrane stack. This phenomena repeats between every membrane stack. The feed flow into the second or further membrane stacks with a higher dry matter content increases the rate at which the membrane fouls, causing higher resistances compared to the previous membrane at the same time. This problem is shown in Figure 6.16. The magnitude of the resistance determines how thoroughly the cleaning process must be, but more importantly the reduction of production

capacity for each membrane, which eventually results in the maximum time for a single batch. From Figure 6.16 it is observed that the resistance of the last membrane stacks is

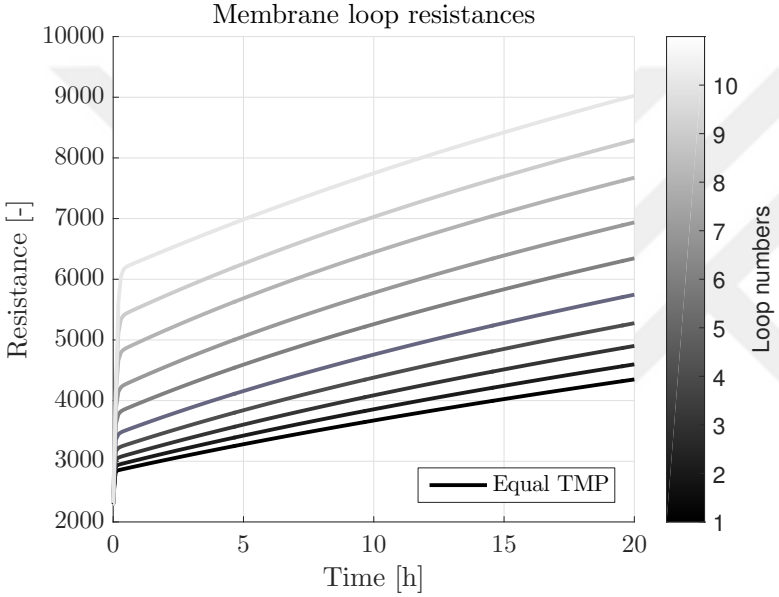
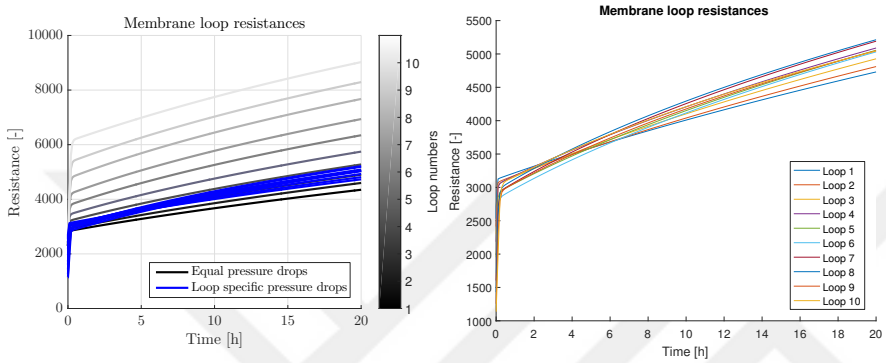


Figure 6.16: Membrane resistances of ten membrane stacks, where all membranes are switched on at $t = 0$. The plots show how the resistance of the last membranes is much higher than the resistance of the first one. Also the slope of the resistance built-up for the tenth membrane is steeper than that of the first membrane.

about twice as high as that of the first membrane at the end of the $20h$ batch.

If one is able to manipulate the TMP for each membrane stack separately, it becomes possible to influence the resistance accumulation (fouling) of individual membrane stacks. By decreasing the TMP of the later membrane stacks, the empirical membrane resistance can be reduced, which however causes a decrease in throughput. This can be compensated by increasing the TMP of the first membrane at the cost of increase in resistance. By tuning the TMP's for every loop in the right way, the UF capacity and output can be kept equal while reducing the maximal resistance build-ups over all membrane stacks. We construct a simulation case study to demonstrate the possible effects of a variable TMP strategy. In the simulation the TMP's of the individual membrane stacks are tuned, and kept constant during the batch. The resulting resistances are shown in Figure 6.17a. Furthermore the actual resistances for this case can be seen in Figure 6.17b.

Figure 6.17b shows that by adjusting the TMPs and mass flows, the membrane stacks keep a specific resistance gradient (accumulation), so that the resistances can be made exactly equal, which in turn distributes the total permeate flux to each membrane stack evenly. This could possibly be accomplished by implementing a supervisory controller which can control the independent TMP's during a batch.



(a) Membrane resistances where all membrane loops on.

(b) Membrane resistances for all loops on and pressure drops set per loop.

Figure 6.17: For the case depicted in (a) the TMP is set to the lowest for loop 10 and to the highest for loop 1, resulting in the blue trajectories.

In this section, we show that the membrane resistances and the accumulation can be reduced while rigorously adjusting the operating conditions of each membrane unit. Therefore the production capacity of the UF is not increased yet. However, with the decrease of the operating resistance values there is room for further improvements in the ultrafiltration process. One option could be to operate the process with a higher feed pressure leading to higher throughput. This increased pressure would induce an increase in fouling. Another option could be to extend the batch times. The increased operating time window would reduce the production to cleaning ratio, and therefore increases the capacity of the UF membrane unit.

6.4 Observability and Identifiability Analysis of the UF Membrane System

Some of the recent research on OMBA methods has been in the area of maintaining the accuracy of the rigorous models since the performance of any model-based application is affected by the quality of the model ([192]). A research topic which has been investigated to improve and maintain model accuracy is the area of parameter estimation ([192, 225, 346]). Parameter estimation problems are concerned with obtaining accurate estimates of unknown or varying parameters by using the measured input-output data. These works include parameter estimation algorithms, observer design techniques and also methods to decide which parameters to estimate ([212, 225]). Similar to these the quality of the estimation results depend on what is being measured. In this section, we address the optimal sensor configuration (OSC) problem for state and parameter estimation purposes in order to achieve high model accuracy and improved maintenance in the whey process.

The problem of state and parameter estimation requires the concepts of observability and identifiability, which indicate whether the initial state or parameters can be uniquely distinguished from the input and output measurements. However determining whether the

system is observable or not, is not enough to decide what to measure. The OSC problem might require comparison of different measurement channel configurations with respect to the resulting variance of the estimates or the numerical reliability of the estimates, while all of the output configurations might render the system to be observable ([342, 352]). To this end, we use the *degree of observability measures* to indicate how informative sensors are. Many of these measures are based on the spectral properties of the observability gramian ([251]). These different measures on the observability gramian may lead to different 'optimal' sensor configurations (with respect to (w.r.t.) different measures). Thus, a multi-objective optimization scheme could be utilized to solve the OSC problem ([335]).

In order to use the degree of observability measures for the OSC problem, we need to construct the observability gramian. This matrix reflects the influence of the states on the measurements and hence provides a description of the state-output behaviour of a system ([342]). In the literature, many of the studies consider the linear time-invariant (LTI) or linear time-varying approximations of the complex system for deciding which signals to measure. This is due to the ease of solving the OSC problem with the observability gramian ([352]), or the observability matrix ([345]), or the Hautus' test, ([362]). However, all of these methods are only valid for linear systems. One can make use of the differential-geometric techniques ([258]) or the tools from lie algebra ([64]) for addressing the observability property of nonlinear systems. For large-scale systems, these techniques are almost impossible to apply, since the computational demand required to solve the problem is tremendous ([322]). Many other algorithms are developed to address the OSC problem in relation to the estimation quality. For example, metrics on the error covariance matrix of the implemented Kalman filter are developed in [261] and [195]. Although the connection between the Fisher information matrix and the gramian is known for a long time, ([25]), the analytic relations for the bounds on the spectral properties of the state estimation covariance matrix derived from the observability gramian is missing. A discussion on these bounds can be found in [165, §4.3].

Recently data-based methods are introduced to approximate the controllability and observability gramians of nonlinear systems with empirical counterparts ([198]). These (covariance) matrices are used for the balanced truncation approach in the model reduction problems for nonlinear systems ([149]). These type of algorithms suffer from the number of experiments/simulations that are needed to be done. Generally one needs to conduct many simulations to include the effect of nonlinear behaviour into the empirical gramian matrix. However once the data is gathered, the effect of the states (and consequently the parameters) can be observed in the gramian matrix which is obtained only from the data. As indicated in [333], combining the computational ease of the calculating output covariance matrix with the established measures for OSC problem that are valid for linear systems is a big step for solving the OSC problem for nonlinear systems.

Here we present a procedure to decide upon the output measurement channels of UF membrane process, represented with unstable nonlinear differential algebraic equations (DAEs), to estimate the parameters and states effectively. The decision is made by using the measures cast on the observability gramians of the system. Due to the nonlinearities and instabilities, the conventional observability gramians are not available. Instead, we make use of the empirical observability gramians. The effectiveness of selected sensor channels are compared with respect to the degree of observability measures cast on empirical gramians.

6.4.1 Problem Formulation and Background Information

The OSC problem is connected to the effect of the states and parameters to the output measurements, which can be tracked with the spectral properties of the observability gramian. One important spectral property of the observability gramian, i.e., its positive definiteness, is also connected to the observability (and identifiability) concept, from which we start our discussion. For this purpose we consider a DAE system in the form of Equation (6.23)

$$\Sigma : \begin{cases} \dot{x}(t) = & f_d(x(t), z(t), p, u(t)), \\ 0 = & f_a(x(t), z(t), p, u(t)), \\ y_i(t) = & h_i(x(t), z(t), p, u(t)), \end{cases} \tag{6.23}$$

where $x(t) \in \mathbb{R}^{n_x}$ $(t \in \mathbb{R}_{[0,t_f]})$, is the differential state, $z(t) \in \mathbb{R}^{n_z}$ is the algebraic state, $p \in \mathcal{P} \subseteq \mathbb{R}^{n_p}$ is the (unknown) parameter vector associated with the model, $u : [0, t_f] \to \mathbb{R}^{n_u}$ is the input and $y_i(t) \in \mathbb{R}^{n_{y_i}}$ is the measured output for the i^{th} possible output configuration, in which $i \in \mathcal{I} := \{1, 2, \ldots, n_i\}$ and h_i is the associated sensor configuration, being an element of the set of possible measurement channels $\mathcal{C}$. Our task is to decide the 'best' sensor configuration h_i to improve the state and parameter estimation accuracy. The initial condition at $t = t_0$ of Σ is denoted with x_0. It is assumed that the functions f_d and f_a are analytic, the Jacobian $\partial f_a/\partial z$ is invertible at all points (x, z, u) and the differentiation index of DAE model is equal to one, hence there exists a linear approximation Σ_{lin} on the trajectory $(\bar{x}(t), \bar{z}(t), \bar{u}(t))$ with deviation variables $\delta x, \delta u$, i.e.,

$$\Sigma_{lin} : \begin{cases} \delta\dot{x}(t) = & A(t,p)\delta x(t) + B(t,p)\delta u(t) + g(t,p), \\ \delta y_i(t) = & C_i(t,p)\delta x(t) + D_i(t,p)\delta u(t) \end{cases}$$

The observability and identifiability properties play a crucial role in state and parameter estimation problems to result in correct estimates. We start with the observability concept.

Definition 6.4.1 *([160]) The model Σ, with known parameters p and input $u(t)$, is observable if the initial state x_0 can be uniquely determined from the inputs and outputs $u(t), y(t)$ for $t \in \mathbb{R}_{[t_0,t_f]}$;*

$$\exists u(t), \exists p, \forall x_0^1, x_0^2, t \in \mathbb{R}_{[t_0,t_f]} : y_i^1(t) = y_i^2(t) \implies x_0^1 = x_0^2,$$

where $y_i^1(t)$ and $y_i^2(t)$ are the output values obtained from input $u(t)$ and the initial conditions x_0^1 and x_0^2, respectively.

One of the main tools for analyzing observability is the observability gramian, $W_o : \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0} \to \mathbb{R}^{n_x \times n_x}$. For the linearized model Σ_{lin} and time instants $t_f > t_0 \geq 0$, $W_o^i(t_0, t_f)$ is defined as

$$W_o^i(t_0, t_f) := \int_{t_0}^{t_f} \psi_i^T(p, \tau, t_0)\psi_i(p, \tau, t_0)d\tau. \tag{6.24}$$

where $\psi_i(p, t_1, t_0) = C_i\phi(p, t_1, t_0)$ and $\phi(p, t_1, t_0)$ is the state transition matrix from t_0 to t_1. Observe that $\psi_i(p, t_1, t_0) = \partial(\delta y_i(t_1))/\partial(\delta x_0)$, hence $W_o^i(t_0, t_f)$ summarizes the effects on the deviated output δy by the perturbations on the initial state δx_0.

For linear deterministic systems, the observability gramian is full rank if and only if the system is observable hence the states are uniquely reconstructible. Furthermore, for nonlinear systems with differentiable function f_d, the nonsingular observability gramian implies local observability around x_0.

Now we define the identifiability concept which relates the effects of parameter variations to the output measurements.

Definition 6.4.2 *([142]) The parameterized model Σ with a set of parameters $p \in \mathcal{P}$ is locally identifiable if the parameters are distinguishable from each other given the initial state x_0, the inputs $u(t)$ and the outputs $y_i(t)$ for $t \in \mathbb{R}_{[0,t_f]}$;*

$$\begin{aligned} \exists x_0,\ \exists u(t), t \in \mathbb{R}_{[0,t_f]}, \quad & \forall p_1, p_2 \in \mathcal{P}: \\ & y_i^1(t) = y_i^2(t) \implies p_1 = p_2. \end{aligned} \tag{6.25}$$

where $y_i^1(t)$ and $y_i^2(t)$ are the output values obtained from input $u(t)$, the initial condition x_0 and the parameters p_1 and p_2.

The identifiability property guarantees that any deviation in the parameters eventually effect the outputs on the operating trajectory. One general way of checking identifiability, given input and initial condition, is by constructing the parameter-to-output sensitivity matrix, i.e., $\mathcal{W}_p^i(t) := \partial y_i(t)/\partial p$. This matrix is connected with the observability gramian of the (new) system constructed with the parameters, i.e., p in Equation (6.23), augmented as new states, i.e., $\dot{p} = 0$. Then the observability matrix of the augmented system contains the information deduced from the parameter-to-output sensitivity matrix, i.e.,

$$\bar{W}_o^i(0, t_f) = \begin{bmatrix} W_o^{i^{xx}} & W_o^{i^{xp}} \\ W_o^{i^{px}} & W_o^{i^{pp}} \end{bmatrix} = \int_{t_0}^{t_f} \begin{bmatrix} \left(\frac{\partial y_i}{\partial x_0}\right)^T \frac{\partial y_i}{\partial x_0} & \left(\frac{\partial y_i}{\partial p}\right)^T \frac{\partial y_i}{\partial x_0} \\ \left(\frac{\partial y_i}{\partial x_0}\right)^T \frac{\partial y_i}{\partial p} & \left(\frac{\partial y_i}{\partial p}\right)^T \frac{\partial y_i}{\partial p} \end{bmatrix} dt.$$

Under this construction the parameter identification problem boils down to the problem of the observability of states of the augmented system.

Since the observability gramian is utilized for both the state and parameter estimation problems, we consider methods that construct the gramian for large-scale and nonlinear models with relatively low computational complexity. For nonlinear systems the construction of the observability gramian does not follow from the linear case trivially. One can overcome this problem by the empirical observability gramian. This matrix is constructed via the output measurements which are obtained from (nominal and perturbed) initial states (and initial parameters) by calculating the covariance matrix of the output measurements. The empirical observability covariance matrix for system Σ in Equation (6.23) is calculated as ([334]);

$$\hat{W}_o^i(t_0, t_f) := \sum_{l=1}^{r} \sum_{m=1}^{s} \frac{1}{r} \frac{1}{s} \frac{1}{c_m^2} \int_{t_0}^{t_f} T_l \Psi^{lm}(t, i) T_l^T dt, \tag{6.26a}$$

where r and s are the cardinalities of the sets of perturbation directions (T^{r,n_x}) and magnitudes (M);

$$\begin{aligned} T^{r,n_x} = & \{T_1,\dots,T_r; T_l \in \mathbb{R}^{n_x \times n_x}, T_l^T T_l = I, l = 1,\dots,r\}, \\ M = & \{c_1,\dots,c_s; c_s \in \mathbb{R}, c_m > 0, m = 1,\dots,s\}, \end{aligned} \tag{6.26b}$$

and the $(jk)^{\text{th}}$ element of the matrix $\Psi^{lm}(t,i) \in \mathbb{R}^{n_x \times n_x}$ is defined as

$$\Psi_{jk}^{lm}(t,i)(t) = (y_i^{jlm}(t) - \bar{y}_i(t))^T (y_i^{klm}(t) - \bar{y}_i(t)),$$

with $y_i^{jlm}(t)$ being the output of the system corresponding to the initial condition $x^{jlm}(0) = c_m T_l e_j + \bar{x}_0$, e_j being the j^{th} basis direction in $\mathbb{R}^{n_x}$ and $\bar{y}_i(t)$ is the nominal-unperturbed output of the system initiated from the nominal state $\bar{x}_0$. All of the output trajectories, $y_i^{jlm}(t)$ and $\bar{y}_i(t)$, are steered with input $\bar{u}(t), t \in [0, t_f]$. For stable systems one can use the steady state values of states and outputs, see, e.g., [333]. However for an unstable system, since no steady state values exist, we propose that one should take the $\bar{y}_i$ values as the nominal (unperturbed) trajectory while calculating the empirical gramian. In this case observe that $y_i^{jlm} \approx \bar{y}_i + \delta y_i^{jlm}$, hence $\Psi_{jk}^{lm}(t,i) = (e_j^T T_l^T c_m)\psi_i^T(p,t,t_0)\psi_i(p,t,t_0)c_m T_l e_k$ (since $\delta u(t) = 0$ for $t \in [t_0, t_f]$), which yields to true observability gramian after evaluating the integral and the sums in Equation (6.26a). Hence if the model equations Equation (6.23) are sufficiently smooth and the linear approximation for perturbations defined by T_l and c_m is valid, then the empirical observability gramian converges to the true observability gramian. Thus, from a system theoretical point of view, both the theoretical and computational problems induced by nonlinearities or instabilities are discarded with the empirical observability gramian approximation.

The operating region is highly effective on the calculated gramian which should reflect the nonlinear behavior in the covariance matrix. The approximation quality of the empirical observability gramian to the real observability gramian of the system is determined by the selection of sets T^{r,n_x} and M. The set of perturbations T^{r,n_x} should be selected such that $c_m T_l e_j$ terms excite the nominal initial state $\bar{x}_0$ in all of the directions spanning the state space, both the positive and negative directions to approximate $\partial y_i(t)/\partial x_0$; while the perturbation magnitudes c_m should be selected wisely to capture the whole region of interest by adjusting the magnitudes of the elements of M, see e.g., [333].

In practice, the initial state might be difficult to reconstruct from data, if the observability gramian is ill-conditioned. To briefly elaborate, consider the gramian based state estimates $\delta\hat{x}(t)$ for model Σ_{lin} with $\delta u(t) = 0, t \in [t_0, t_f]$,

$$\delta\hat{x}(\tau) = W_o^{i^{-1}}(t_0,\tau) \int_{t_0}^{\tau} \phi^T(p,s,t_0) C_i^T \delta y_i(s) ds,\ \tau \in (t_0, t_f].$$

If the observability gramian is ill-conditioned, this will cause poor results for the estimates. In this section this issue is taken into account with the degree of observability measures. These measures are cast on the (true or empirical) observability gramian to compare different output channels with respect to the informativity of the measured signals. Several of these measures are listed below, for the set of possible output channels $\mathcal{C}$.

Eigenvalues: The minimum eigenvalue of the observability gramian indicates how close the system is to becoming unobservable, hence higher values of the smallest eigenvalue imply a greater degree of observability measure, i.e.,

$$\mu_1 := \max_{h_i \in \mathcal{C}} \arg\min_{\lambda} \{\lambda(W_o^i(t_0, t_f)\}\}.$$

However the eigenvalue measure does not reflect the effect of perturbations in other directions than the eigenvector associated with the minimal eigenvalue. Hence in general it is difficult to assess the overall performance through the eigenvalue measure.

Trace: The second measure compares the trace of the observability gramians w.r.t. different sensor selections, i.e.,

$$\mu_2 := \max_{h_i \in \mathcal{C}} \operatorname{trace}(W_o^i(t_0, t_f)) = \max_{h_i \in \mathcal{C}} \left\{ \sum_{j=1}^{n} \lambda_j(W_o^i(t_0, t_f)) \right\}.$$

The trace is the sum of eigenvalues, hence representing the maximum effect on the output w.r.t the perturbations in all directions. It is observed that ([333]) for OSC problem with estimation purposes, the measure μ_2 yields better results compared to the other measures defined here.

Determinant: The third measure is the determinant of the observability gramian. A larger determinant of $W_o^i(t_0, t_f)$ shows that at least one eigenvalue is larger compared to the other sensor configurations. Therefore the largest determinant case of the gramians constructed for different sensor selections corresponds to a better output channel configuration, i.e.,

$$\mu_3 := \max_{h_i \in \mathcal{C}} \det(W_o^i(t_0, t_f))^{\frac{1}{n}}.$$

The determinant measure is numerically problematic and varies with a high dependence on the elements of the gramian matrix, thus depends heavily on the selection of T^{r,n_x} and M matrices for empirical observability gramian calculation.

Condition number: For state estimation quality, one needs to consider the maximum and minimum singular values and associated directions of the observability gramian. Another method that can be utilized for the OSC problem is the condition number of the gramian, which is defined as

$$\kappa(W_o^i(t_0, t_f)) := \frac{\sigma_{max}(W_o^i(t_0, t_f))}{\sigma_{min}(W_o^i(t_0, t_f))}.$$

This method can be viewed as a tool for analyzing the sensitivity of the observability property w.r.t. different measurement channels ([114]). A small valued $\kappa(W_o^i(t_0, t_f))$ implies that the magnitude difference between $\lambda_{max}(W_o^i(t_0, t_f))$ and $\lambda_{min}(W_o^i(t_0, t_f))$ is small, hence a less ill-conditioned observability gramian. This improves the estimation quality, hence this measure minimizes the condition number, i.e.,

$$\mu_4 := \min_{h_i \in \mathcal{C}} \kappa(W_o^i(t_0, t_f)) = \min_{h_i \in \mathcal{C}} \left\{ \log \left(\frac{\sigma_{max}(W_o^i(t_0, t_f))}{\sigma_{min}(W_o^i(t_0, t_f))} \right) \right\}.$$

Apart from these, one can make use of the spectral radius or the near singularity measures as indicated in [333].

Remark 6.4.1 *Since the analytic relations between the spectral properties of observability gramian and state (or parameter) estimation covariance matrix are missing, there is no straightforward way to select the OSC minimizing estimation error variance through the above mentioned measures. These measures are only used as an indicator for the selection procedure.*

6.4.2 Simulation example for Ultrafiltration Membrane System

In this section we implement the sensor selection method via empirical observability gramians to a *simplified* ultrafiltration membrane process. The membrane set-up consists of two membrane stages which are connected in series to each other. There are two outlet flowing streams, hence possible output channels, of each membrane stack, the permeate and the retentate streams. We provide a simple schematic of a membrane system with two stages in Figure 6.18. In this simulation study we make use of a simplified UF membrane model

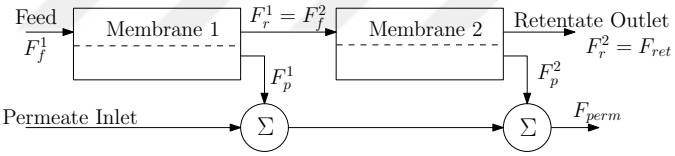


Figure 6.18: The two membrane stacks system.

which is taken from [146] as

$$\begin{aligned} F_f^j &= F_r^j + F_p^j, & F_f^j C_f^j &= F_r^j C_r^j, \\ F_p^j &= AJ^j, & J^j &= \tfrac{\Delta P^j}{R^j}, \\ \Delta P^j &= P_{in}^j - \tfrac{1}{2}\delta P^j - P_{atm}, & \dot{R}^j &= \alpha(\Delta P^j)^\beta J^j C_r^j, \\ P_{in}^2 &= P_r^1 = P_{in}^1 - \delta P^1, & j &= 1, 2, \end{aligned} \tag{6.27}$$

where $j = 1, 2$ is the index over two membrane stacks, the feed, retentate and permeate mass flows are denoted with F_f^j, F_r^j, F_p^j, respectively, C_f^j and C_r^j denote the concentrations of feed and retentate stream, J^j is the mass flux, and R^j is the membrane resistance. The state and parameter vectors are defined as

$$x = \begin{bmatrix} R^1 & R^2 \end{bmatrix}^T, \qquad p = \begin{bmatrix} \alpha^1 & \alpha^2 \end{bmatrix}^T.$$

The parameter values, initial state and operating conditions are tabulated in Table 6.1. The operating conditions are selected in such a way that the performance outputs, F_r^2, C_r^2, are kept in acceptable levels with the linear increase of the pressure, the input u. Furthermore, the initial conditions are determined through the geometry (pore size/density etc.) and operating conditions of the membranes. Lastly, we assume that the membranes foul at $t_f = 10h$, thus the process execution is interrupted.

Area of membrane stack (m^2)	A	74.0350
Pressure drop over membrane (Pa)	δP_i	100
Atmospheric pressure (Pa)	P_{atm}	101.3
Fouling parameter 1 (-)	α	0.0005
Fouling parameter 2 (-)	β	0.5
Inlet pressure (Pa)	$P^1_{in}(t)$	$10^3(1/2 + t/t_f)$
Inlet mass flow (kg/h)	F^1_1	5000
Inlet protein concentration (kg/L)	C^1_1	5
Initial membrane resistance (-)	$R_1(0) = R_2(0)$	15

Table 6.1: Nominal parameter values and operating conditions.

Estimation of States with Empirical Observability Gramian

The possible measurement channels are taken as the mass flows at the retentate and permeate ports. There are seven different output channels which are characterized as

$$y_i = C_i^T \begin{bmatrix} F_{perm} \\ F_{ret} \\ F_r^1 \end{bmatrix},$$

where the output channels C_i are the elements of the set $\mathcal{C}$;

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}. \tag{6.28}$$

The F_{perm}, F_{ret} and F_r^1 are visualized in Figure 6.18 and given as

$$\begin{aligned} F_{perm} = &\quad F_p^1 + F_p^2 + F_{perm}^{inlet},\ F_{perm}^{inlet}(t) = 0, \\ F_{ret} = &\quad F_r^2. \end{aligned} \tag{6.29}$$

For the given set of possible output configurations, we calculate the empirical observability gramian with Equation (6.26a) to calculate the degree of observability measures μ_k. We select the perturbation directions as the positive and negative directions in the state space and the perturbation magnitudes are less then 10% of the initial conditions (see Table 6.1), i.e.,

$$T^{2,2} = \{I_2,\ -I_2\},$$

$$M = \{1.5,\ 1,\ 0.75,\ 0.5,\ 0.25\}.$$

The degree of observability measure results for the membrane system and the empirical observability gramian are visualized in Figure 6.19. From these results we observe that output channel cases 3-5-6 and 7 are providing sufficient information regarding the states of the system, whereas the case 7, measuring the sum of three variables, is the optimal choice for the state estimation purposes. This is due to the output measurements that contain

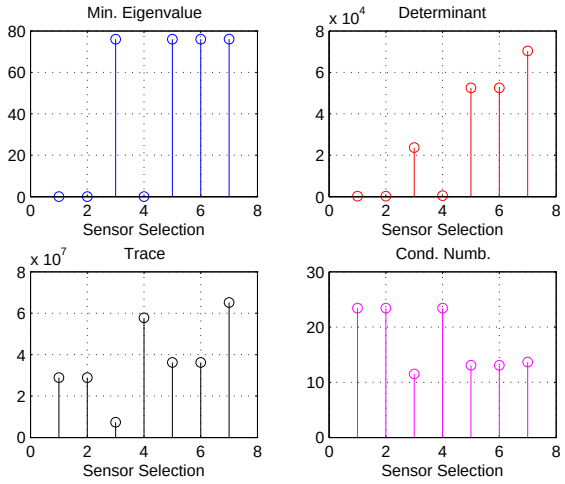


Figure 6.19: Degree of observability measure results obtained for two membrane stack system for state estimation problem.

the largest amount of information. Case 5 and 6 represent the instances measuring the sum of retentate mass flow of first membrane with the outlet permeate or retentate streams, respectively. Observe that all of the degree of observability measures are equal to each other in these two output configuration cases due to the model equation

$$F_f^i = F_r^i + F_p^i,$$

where F_f^1 is assumed to be known. Furthermore, measuring the two streams together causes the system to become unobservable, as seen from the case 4. The optimal sensor selection results correspond to the industrial implementations, where the mass flows of each membrane stack are tracked.

Estimation of Both the States and Parameters with Empirical Observability Gramian

In this part, we utilize the empirical observability gramians of different sensor configurations for parameter estimation purposes. We model the (empirically determined) fouling parameters α^1 and α^2 as constants up to an additive disturbance effecting their initial conditions, i.e.

$$\dot{\alpha}^j = 0,\ \alpha^j(0) = 0.0005 + w^j, j = 1, 2.$$

The Gaussian disturbance terms $w^j \sim \mathcal{N}(0, \Sigma_w)$ are independent and standard deviations are equal to $2.5 * 10^{-5}$, while the initial conditions for the membrane resistances are set similar to the previous case. The perturbation directions and magnitudes are taken as;

$$T^{2,4} = \{I_4,\ -I_4\},$$

$$M = \{\ 2.5,\ 2,\ 1.5,\ 1,\ 0.5\},$$

where we scaled the perturbations with 10^{-5} for the cases perturbing the fouling parameters α^i. The results are tabulated in Figure 6.20 for four different degree of observability measures. The resulting degree of observability measures are similar with the results of Section

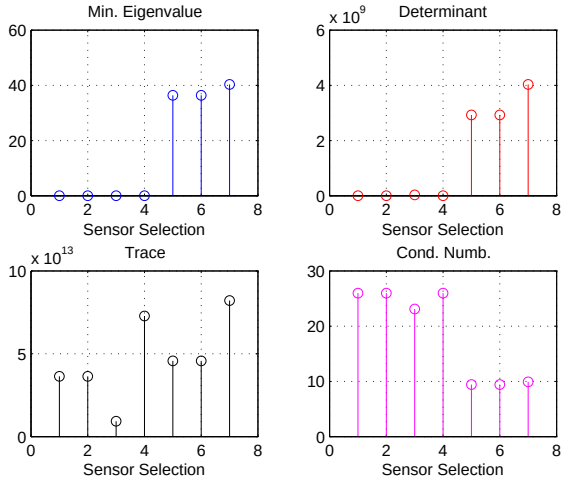


Figure 6.20: Degree of observability measure results obtained for two membrane stack system for state and parameter estimation problem.

3.2. The cases 5, 6 and 7 are providing the best sensor channels, again, while for the case 3, the information required for estimating the second membrane fouling coefficient is not present in the measurements. Furthermore the output configuration cases 1, 2 and 4 are prone to be numerically problematic for estimation purposes which can be observed by the ill-conditioning of the observability gramian, see the condition number subfigure of Figure 6.20.

Lastly, to check the validity of the implementation and the results, we designed Kalman filters for the membrane system with the output measurements varying over the measurement channel cases. In Figure 6.21, we visualize the results for the states and the parameters, including the true values of the states (R_1 and R_2) and the parameters (α^1 and α^2) and the estimation results for the output channel cases 5, 6 and 7. From these results we observe that the optimal output channels are providing an acceptable estimation performance for both of the states and the parameters. The estimation errors of parameters α^i are relatively high but convergent to the true value, however the state trajectories are estimated with high precision for the optimal sensor configurations. For the other sensor configuration cases, which are not reported here, the results are not satisfactory, either due to poor estimation quality of α^i or the true variables being not observable at all.

As an outlook, different degrees of observability measures are indicating different aspects of the model. The eigenvalue measure should be used in all cases, to check whether the system is observable or not, however the trace measure is the most reliable one about the informativity of the sensor configuration. The condition number measure indicates whether

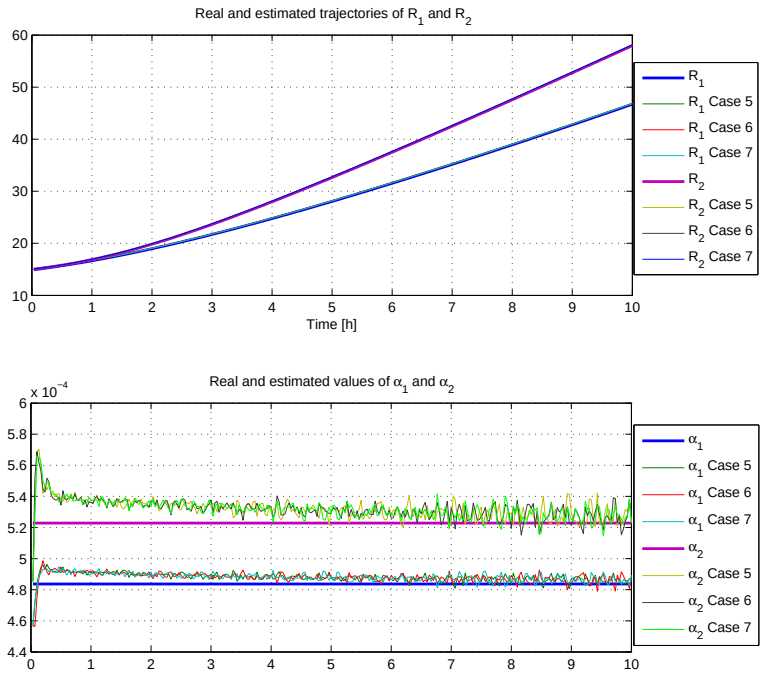


Figure 6.21: Simulation results for the simplified membrane system. The upper figure visualizes the true and estimated membrane resistance (state) trajectories while the bottom figure visualizes the true and estimated fouling coefficient (parameter) trajectories for output channel cases 5, 6 and 7.

the estimation algorithm is prone to be numerically erroneous with quickly varying estimates. Since the empirical gramian is calculated through noisy simulation data, it is not advised to make use of the determinant measure.

6.5 Conclusions on Modeling of Ultrafiltration Membrane Units and Model-based Applications

In this chapter, we discussed a complex industrial process, whey protein separation process, and the dominant unit operation in this process, the ultrafiltration membrane units. A mathematical ultrafiltration membrane model for whey separation process is developed. Lastly, the effect of operating conditions are mapped, through both static and differential relations, to the membrane flux and retention factor variables. For the constructed model, a parameter estimation problem is solved, regressing the gathered industrial data to the mathematical model. The statistical properties indicate that the model describes the process sufficiently and thus the model can be used for control and estimation purposes. With the validated

model, we have implemented several simulation studies to stress-test the model. We implement and challenge the model via optimization studies by comparing it with heuristic operating schedules which shows that the optimization based strategy is preferable since the fouling accumulation is slowed down with the lower operating pressure levels. Finally, a section is dedicated to the possible gains in resistance decrease when the TMP's are tuned for every membrane loop separately. These simulation studies present the potential use of the membrane model for optimal operation and scheduling strategies. We have made use of the developed membrane model in the observability and identifiability analysis. We consider the optimal sensor configuration problem for a simplified UF membrane system, to extract useful information for tracking the fouling. For this purpose we extend the definition of the empirical observability gramians for unstable nonlinear systems to decide on the optimal sensor configuration.

- The UF model incorporates the fouling effect as an additional state, the membrane resistance, on top of the mass balance relations.
- Comparison between three different static membrane resistance functions is conducted and both the simple (polynomial case) and the complex (logarithmic-exponential) approaches fail to reach statistical validation compared to the exponential static membrane resistance.
- All the membranes should be turned on at the beginning of the operation, since the performance improves by distributing the mass flux over all membrane stacks.
- The maximal membrane fouling could be decreased by tuning the TMP's per membrane loop separately. This, with either increased operating pressure or longer batch time, could lead to a higher production capacity while providing desired performance specifications.
- The sensor configuration can be decided based on the empirical observability gramians for estimation purposes.

Chapter 7

Online Model-based Monitoring and Control Applications for Ultrafiltration Membrane Process

Here I opened wide the door;
Darkness there and nothing more.
Deep into that darkness peering,
long I stood there wondering.

Edgar Allan Poe - The Raven

Model-based control and monitoring techniques have been widely accepted and used in oil and petrochemical industries. The benefits of these methods have been documented extensively ([288, 289]). A natural step for using the dynamic model of the UF unit is to show the benefits of model-based technology in the whey protein separation process. To this end, we have first used two observer design methods, namely EKF and MHE, and compared their performances for monitoring the protein content in the retentate and the evolution of fouling phenomena in the membrane stacks. Secondly, model-based control with a learning feature has been developed for controlling the protein content within the batch operation and improving it over batch-to-batch iterations. Finally, the performance of such a controller is compared with the conventional control designs, to demonstrate that with the multivariable model-based control methods, one can extend the whey protein separation operation while maintaining the performance goals on the protein concentration in the processed whey.

7.1 Monitoring of Fouling in Ultrafiltration Membrane Stacks

7.1.1 State Estimation Problem for Ultrafiltration Membrane Stacks

A crucial class of OMBAs is online state or/and parameter estimation. In this application the states or the parameters and more importantly the key performance indicators (KPIs) are continuously monitored. In addition to the monitoring of the process operation, the variables which are tracked by soft sensors can be used in model-based controllers which

require information about the current states or parameters of the process. This necessitates soft sensor implementations in cases where not all states of the system are measured or measurable. Currently, a noticeable amount of soft sensor implementations (both in simulation and industrial cases) are reported in the literature ([338]).

Similarly, soft-sensor implementations for the UF membrane unit are expected to improve the whey protein separation process operation. In the current practice, (physical) sensors to measure protein concentrations in thick whey are costly to install after each and every membrane unit, and even in the ideal case of large number of installed sensors, the measurements of the variables are generally sporadic rather than being continuous over time. The possible delays in the measurements have detrimental effects in the model-based control implementations. Furthermore cost-efficient sensors, such as the ones measuring the (mass) flows in the UF streams, are highly inefficient due to the low resolution and low signal to noise ratio factors. Thus a model-based soft-sensor synthesis has both monitoring and control benefits.

There are many different types of monitoring routines that are proposed in the literature, see [299] for a thorough discussion of different observer design techniques. The choice of the observer type for a certain application varies with several factors such as the dynamics of the process, existence of constraints, the available computational power and the statistical nature of the state or parameters. However, an industrial standard for tracking unmeasured or corrupted states is the Kalman filter [154, 174]. The Kalman filter is proven to be the ($\mathcal{L}_2$) optimal estimator provided that the system dynamics are linear and the covariance levels of disturbances are known to the practitioner. For practical implementations where incorporation of nonlinear dynamics is necessary, the extended Kalman filter (EKF) is used which accounts for nonlinear dynamics of the system to the expense of optimality. Nonlinear dynamics are linearized around the operating point to approximate the nonlinear behaviour of the process, [215], but such techniques result in sub-optimal estimates. EKF provides a simple and efficient approximate solution for the state and parameter estimation problem, while guaranteeing several system theoretic properties such as asymptotic convergence and stability ([215, 304]). EKF based monitoring methods have been successfully implemented in many industrial cases as reported in [113, 205, 338].

Once the estimation algorithm is required to incorporate constraints, such as the physical bounds on the states or disturbances, the Kalman filter loses its effectiveness. In its simplest formulation, the Kalman filter does not incorporate constraints. This might lead to estimates which are not feasible in physical systems, such as negative concentrations in UF membrane units. In addition, there are frequently reported problems with robustness associated with EKF ([300]). To overcome the linearization errors and constraint violations while achieving guaranteed convergence of estimates, moving horizon estimation (MHE) algorithms are introduced in the academic literature, see [252]. MHE techniques are a member of optimization based estimators that rely on model predictions in backwards (time) direction. The corresponding algorithms involve solving an optimization problem constrained to the process model and constraints as specifications or limitations for a finite number of output measurements. Depending on the structure of the optimization problem, the result can be the estimated state, the estimated process parameters or the estimation of exogenous disturbances ([193]).

In this section, both of the mentioned estimation techniques, EKF and MHE, are ap-

plied to UF membrane systems in the whey protein separation process. Our goal is to track the KPIs of the UF unit, but more importantly we want to monitor the fouling accumulation. The fouling phenomenon, being the main bottleneck in the UF operation, impedes the desired performance of membrane units. One way to effectively cope with the fouling phenomenon is to track the fouling effects by employing soft sensing algorithms and then controlling the process such that the fouling accumulation is slowed down and hence the process efficiency is increased. This line of reasoning has been used in several works reported in the literature e.g. [168, 330]. However, these efforts did not consider large-scale rigorous models of membranes but rely on simple black box models. Since we desire to track the fouling, these models are not detailed, or accurate, enough for soft-sensor implementation. Furthermore, soft-sensor methods for large-scale models require special attention to issues such as numerical ill conditioning and calculations over large sets of decision variables. We analyze rigorous soft-sensor design techniques for UF membrane units in this section due to these considerable problems.

We introduce the estimation problem and present the notation for the next sections. Here we represent the UF membrane model developed in Chapter 6 with the continuous time system Σ,

$$\Sigma : \begin{cases} \dot{x}(t) = & f_d(x(t), z(t), u(t), w(t)), \\ 0 = & f_a(x(t), z(t), u(t), w(t)), \\ y(t) = & h(x(t), z(t), u(t), v(t)), \end{cases} \tag{7.1}$$

where $x(t)$, $z(t)$, $u(t)$ and $y(t)$ denote the dynamic state, algebraic state, the input and the output vectors, respectively. Furthermore, $w(t)$ and $v(t)$ represent process and measurement noise with covariance matrices Q^w and R^v, respectively. Finally, f_d, f_a and h are nonlinear functions relating states, inputs and noise to states and measurements at $\{y(\tau), \tau \leq t\}$. The aim of state estimation then becomes reconstruction of the dynamical state vector $x(t)$ based on the measurements and the model. Let us denote the estimate of $x(t)$ by $\hat{x}(t)$. The estimation error then becomes $e(t) = x(t) - \hat{x}(t)$.

Extended Kalman Filter

The first estimator type that we discuss is the Kalman filter. The Kalman filter addresses the problem of finding $\hat{x}(t)$ by minimizing the error covariance $P(t) = E[e(t)e(t)^\top]$ on the basis of output measurements $\{y(\tau) : \tau \leq t\}$. This results in a predictor-corrector algorithm, which consists of two main steps:

- The predictor step calculates an estimate of the state and output based on the system dynamics Σ in Equation (7.1) without the disturbance terms corrupting the evolution, $w(t) = 0, v(t) = 0$;
- The corrector step updates the estimation by penalizing the resulting model-based output prediction $\hat{y} = h(\hat{x}(t), \hat{z}(t), u(t), 0)$ in comparison to the output measurement.

The predictor step is mainly concerned with iterating the state and covariance estimates forward in time by using the model. In the correction step the observer gain is calculated, the a-posteriori estimate of the states and the error covariance are updated.

In this thesis, we implement the estimators in digital devices, so we operate with the discrete time EKF algorithms, which are explained next. As a starting point, we linearize

the dynamics of the UF process model in Equation (7.1) around the operating conditions. Assuming a general case where direct feedthrough terms are also present, linearization of the DAE model leads to:

$$\Sigma_{lin} : \begin{cases} \delta\dot{x}(t) = & A(t)\delta x(t) + B(t)\delta u(t) + B^w(t)w(t), \\ \delta y(t) = & C(t)\delta x(t) + D(t)\delta u(t) + D^w(t)w(t) + v(t), \end{cases} \tag{7.2}$$

where δx contains the deviations from dynamical states of the UF model[1]. These equations are subsequently discretized, by sampling the system at time instants $t = kT_s$, $k \in \mathbb{Z}_{\geq 0}$ and $T_s \in \mathbb{R}_{>0}$ is the sampling time, in order to apply the soft sensing algorithms once new measurements are conducted. The linearized and discretized model is then represented as

$$\Sigma_{lin}^{DT} : \begin{cases} \delta x_{k+1} = & A_k\delta x_k + B_k\delta u_k + B_k^w w_k, \\ \delta y_k = & C_k\delta x_k + D_k\delta u_k + D_k^w w_k + v_k, \end{cases} \tag{7.3}$$

These state space system matrices and the initial guesses on the state estimate and the covariance matrix $(\hat{x}_0, P_0)$ are used in the EKF algorithm, which is visualized in Figure 7.1. The EKF algorithm initializes at the time update with the initial conditions $(\hat{x}_0, P_0)$ and $Q_k = Q^w$, $k \in \mathbb{R}_{\geq 0}$. With nominal dynamics, we integrate the prior estimated state $\hat{x}^-(t)$ and iterate the prior error covariance matrix P_{k+1}^-. The corrector, also called as measurement, step updates the prior estimates according to the measurement mismatch $\tilde{y}_{k+1}$ and the variance model of disturbances affecting the P_k and L_k. This update finalizes the Kalman filter algorithm, by sending the estimated states back to predictor step.

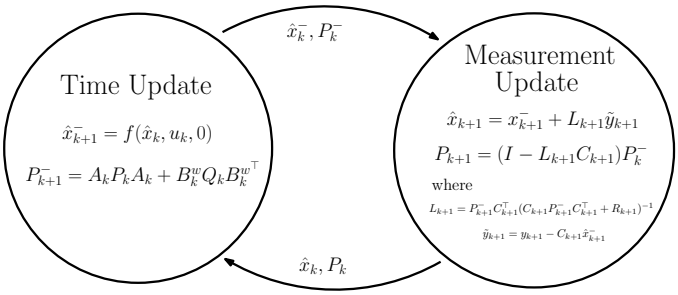


Figure 7.1: The EKF algorithm divided into prediction (Time update) and correction (Measurement update) updates.

Choice of Covariance Matrices and the Numerical Properties of EKF The weighting matrices Q_k and R_k used in the Kalman filter update equations are the tuning parameters of the EKF. These parameters indicate whether the measurements of outputs or the model-based predictions are trusted. With a low R_k (compared to the singular values of the Q_k matrix), measurements are trusted more, so that the resulting estimates are aligned more with the

[1]The algebraic states $z(t)$ vanishes, since the DAE system is an index-1 DAE system, which can be written as a causal LTI system, see also [77, 78, 79].

measured outputs, while for the opposite case the model is the trusted more than the measurements. In other words, the observer is tuned not by the individual covariance matrices but by the ratio between their singular values.

In practice it is rarely the case that the covariance matrices Q_k and R_k are known or accessible. Hence, these weighting matrices are tuned heuristically and independent of k. The tuning is accepted once the desired estimation results are obtained. In this dissertation, we compare two different ways of designing the weight matrices. First one is using static, heuristically tuned $Q_k = Q^{C1}$, where $C1$ stands for case 1, and we detail the second case next. In a scenario where process noise covariance levels Q^w and R^v are known, the Q_k and R_k matrices can be computed as follows:

$$\begin{aligned} Q_k &= B_k^w Q^w {B_k^w}^\top, \\ R_k &= D_k^w Q^w {D^w}^\top + R^v \end{aligned} \tag{7.4}$$

where B_k^w and D_k^w follow from the linearized and discretized dynamics of the membrane model and R^v is the measurement noise covariance matrix. With this method, the choice of the covariance matrices becomes adaptive.

A second issue with the EKF based monitoring applications is its numerical implementation for large-scale rigorous systems. In the implementation of EKF towards the UF membranes, numerical issues are observed at the measurement update phase. On top of it large-scale system dynamics leads to difficulties in the inversion operation of EKF corrector step, see Figure 7.1. To be specific, numerical computations in this step yield a P_k which is not nonnegative. However this is an impossible situation as we analyze the update equations. To cope with this issue, the square root of the covariance matrix P_k, found by Cholesky factorization, is proposed to be used, [49]. This method was further improved by using the UDL decomposition of the covariance matrix instead of Cholesky factorization. The algorithm presented in [366] addresses the numerical stability issues in the case of multiple measurements. Moreover, this algorithm exhibits higher numerical stability compared to other alternatives. In this thesis, we use the UDL decomposition based algorithm to overcome the numerical problems.

Moving Horizon Estimation

Both the Kalman filter and the EKF do not consider any constraints on the states and disturbances. Including inequality constraints to preserve the physical bounds is, in general, expected to yield better estimation results. Furthermore, applying the Kalman filter to nonlinear systems by means of the EKF comes at the cost of optimality as a result of linearization effects.

To improve the estimation results by taking the constraints on the states and the disturbances into account, the estimation problem can be formulated as a series of optimal control problems ([300]). However, solving optimization problems which take all of the available measurement data into account has an extremely high computational cost. In fact the size of the optimization problem increases with every new measurement that is made at every time instance. Moving horizon estimation (MHE) techniques are used to limit the computational size of the estimation problem by employing a time window of finite length while incorporating the constraints. Once new measurements are made, the old measurements

are discarded and new estimates are found by solving the finite horizon estimation problem again for which constraints and nonlinear system dynamics are taken into account.

Technically, we make use of linearization dynamics Σ_{lin}^{DT} in Equation (7.3) to estimate the state deviation that minimizes the measurement mismatch while satisfying constraints on the estimated trajectories. Consider a time instant k and given constraints on the dynamical states, i.e., $c_j(\hat{x}_{k-i}) \leq 0$, where $j = 1, \ldots, N_j$, N_j is the number of constraints, $i = 0, 1, \ldots, N_w - 1$ and N_w is the horizon length of measurement signals. Then the optimization problem reads as;

$$\mathcal{P}^{MHE} : \begin{cases} \delta\hat{x}_k^* = \arg \min\limits_{\delta\hat{x}_k} \sum\limits_{i=k-N_w+1}^{k} ||\tilde{y}_i - \delta y_i||^2_{Q_M} + ||\delta\hat{x}_k||^2_{R_M}, \\ \text{subject to} \quad \delta\hat{x}_{i|k-N_w+i} = A_{i-1}\delta\hat{x}_{i-1|k-N_w+1} + B_{i-1}\delta u_{i-1}, \\ \qquad\qquad \delta y_i = C_i\delta\hat{x}_i + D_i\delta u_i, \\ \qquad\qquad \tilde{y}_i = y_i^{meas} - \hat{y}_i^-, \\ \qquad\qquad c_j(\hat{x}_i) \leq 0, \\ \qquad\qquad i = k - N_w + 1, \ldots, k, \end{cases} \tag{7.5}$$

where $\hat{x}_k^* = \hat{x}_k^- + \delta\hat{x}_k^*$ is the optimal state estimate at time $k \in \mathbb{Z}_{\geq 0}$, $\hat{x}_k^-$ is the prior state estimated obtained from integrating the previous optimal state estimate $\hat{x}_{k-1}^*$ with the nonlinear dynamics in Equation (7.1); $\tilde{y}_k$ is the difference between the measured output and prior estimate of output, y_k^{meas} and $\hat{y}_k^-$, respectively; δy_k is the deviation at the output parametrized with δx_k; and lastly Q_M and R_M are the weighting matrices in the MHE problem. Once the MHE problem (Equation (7.5)), is solved, the optimal deviation in the state estimate $\delta\hat{x}^*$ is obtained and estimated states and outputs $\hat{x}_k^-, \hat{y}_k$ are updated with respect to this deviation variable, i.e.,

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + \delta\hat{x}_k^*, \\ \hat{y}_k &= \hat{y}_k^- + C_k\delta\hat{x}_k^* + D_k\delta u_k. \end{aligned}$$

7.1.2 Implementation of Monitoring Algorithms for Ultrafiltration Membrane Process

Operating Conditions of Ultrafiltration Membrane Stacks

Two different sets of measurements are used in the estimator design for the UF membrane process. In the first study, an optimal operating strategy is constructed as in [316]. We refer to the resulting input-output data set as the *simulation data set*. The *simulation data set* and the (virtual) measured outputs are generated from the dynamical model with the given optimal inputs, which provide the trajectories of the variables in the process. To mimic realistic operating conditions of the UF process, perturbations and noise are added to the outputs and inputs. To this end, input and output noise levels are selected as 5% of the corresponding nominal signals. The uncertainties are taken as; i) for state estimation purposes, we consider input and output noise levels of 5% of the applied value; ii) we further introduce varying perturbations (a sine wave) for the fouling parameter α for each membrane that is equal to the 5% (at maximum) of the nominal value, i.e.,

$$\alpha^i(t) = \alpha_0^i + \frac{\alpha_0^i}{20}\sin(\frac{2\pi t}{10}).$$

The input variables, feed mass flow and pressure trajectories, are visualized in Figure 7.2 and Figure 7.3.

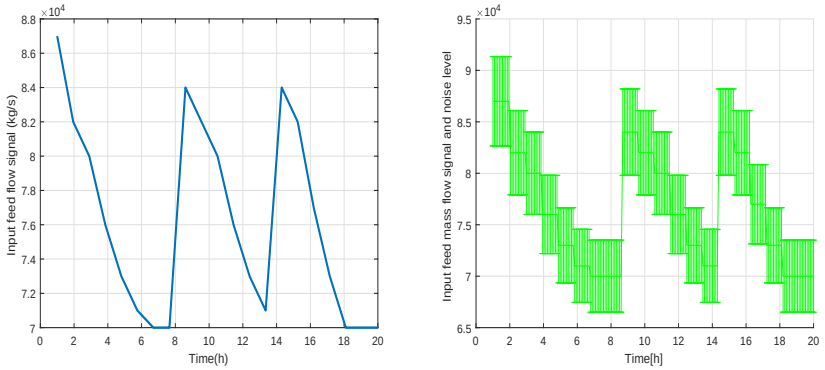


Figure 7.2: True and noisy input feed flow in the simulation data set.

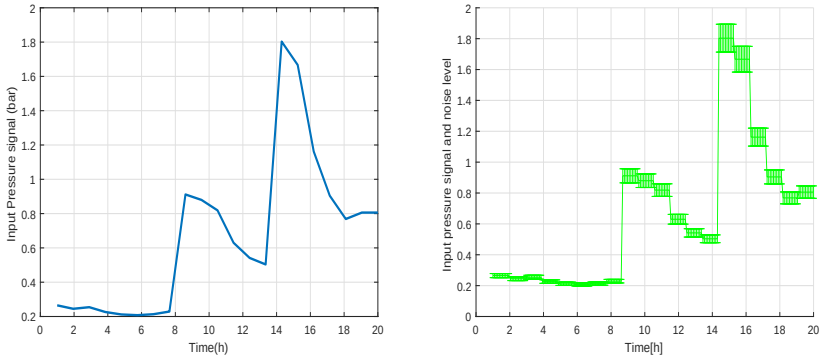


Figure 7.3: True and noisy input pressure in the simulation data set.

As the second simulation experiment, an industrial data set is gathered from a whey protein separation plant which is run by operators. The operating strategy used by the operators is then mimicked by the dynamical model to generate the variables which are not measured. We refer to this data set as the *Industrial data set*. This data set consists of mass flow measurements with the associated operating pressure profile. This operation depends on the operators' personal skill and experience level. In this experiment, not all of the membranes in the plant are operational at all time instants. This is different from the simulation data set, in which 8 membranes are kept operational during the whole batch time. After the start of the batch, the operator introduces two membranes to compensate the performance drop due to the fouling effects, approximately around the $t = 3h$ and $t = 15h$. This introduces hybrid behaviour to the monitoring model, which deteriorates the estimation performance. In the case of the industrial data set, additional disturbances are

not added to the input and output signals since the collected data from the plant operation is already noisy. The true and noisy input feed flow and pressure signals used to generate this data set are presented in Figure 7.4 and 7.5 respectively.

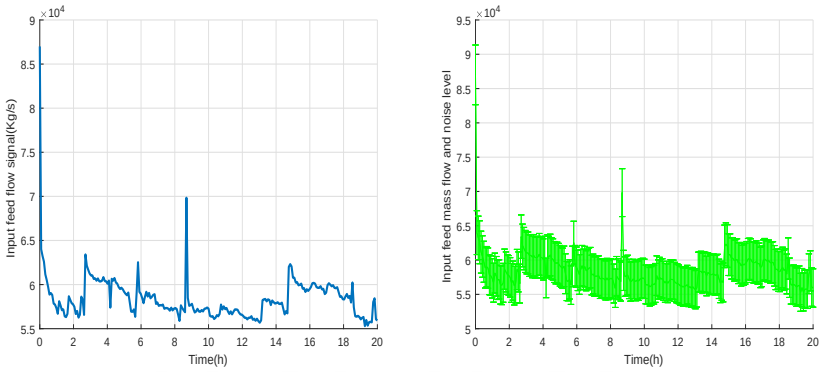


Figure 7.4: True and noisy input feed flow in the industrial data set.

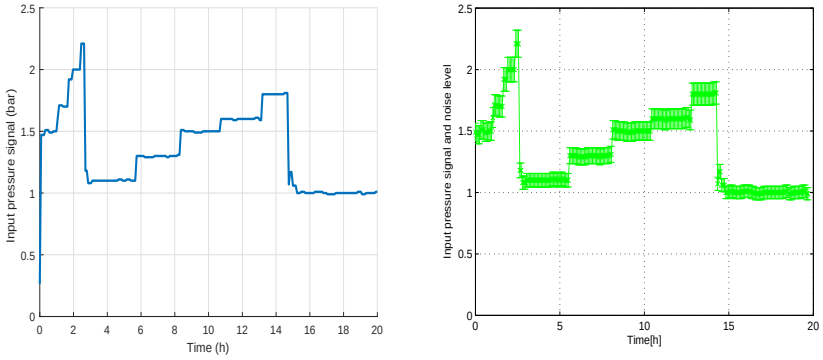


Figure 7.5: True and noisy input pressure in the industrial data set.

Parameter Configuration of Estimators

We implement three different monitoring algorithms and calculate the estimates, namely

- EKF with heuristic weight selection;
- EKF with time varying weights;
- MHE.

For the first case, the EKF weights are tuned with trial and error. In the second case, EKF with adaptive weight tuning is implemented; as described in Equation 7.4; which we refer to as the Case 2. The last implementation is the MHE based estimation routine.

- **EKF Case 1:** In order to overcome the numerical issues encountered during the measurement update phase, we implemented the UDL decomposition based EKF algorithm presented in [366]. For EKF Case 1, the covariance matrices Q_k and R_k are tuned by trial and error, until the desired estimation error results are obtained. The reported estimation results are obtained with the (constant) covariance matrices as,

$$Q_k^{C1} = 0.01^2 I_{104}, \quad R_k^{C1} = \begin{bmatrix} R_F & 0 \\ 0 & R_x \end{bmatrix},$$
$$R_F = 1500 I_9, \quad R_x = 0.012.$$

  The EKF Case 1 algorithm does not take constraints into account. As a result of this, we observe that some states are estimated in infeasible regions of the state space and cause the simulations to leave the operation window for which the model is descriptive. More specifically at the start of the batch operation, several of component concentrations are estimated to be negative, which is violating the model constraints. To circumvent this problem, we apply clipping to satisfy the constraints. With this technique, if the estimates violate physical constraints, they are simply discarded and the prior estimate is used to continue the simulation.

- **EKF Case 2:** In this case, the EKF estimator implementation is the same algorithm as the EKF Case 1, yet the choice of the covariance matrices are different. This leads to a crucial difference since the choice of covariance matrices significantly affects the resulting estimates. In EKF Case 2, the linearized dynamics and the time varying input covariance levels $Q_w(t)$ are used in place of the constant guess. However during the error covariance update step of the EKF, the covariance matrices are taken as

$$Q_k^{C2} = B_k^w Q_k^w {B_k^w}^\top,$$
$$R_k^{C2} = D_k^w Q_k^w {D_k^w}^\top + R,$$

  which leads to adjustable covariance levels. This improves the estimation quality of the EKF. In this case, clipping is used once again for dealing with infeasible estimates during the start of the batch operation.

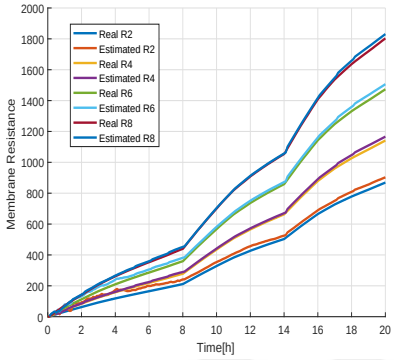
- **MHE:** In the case of MHE with a prediction horizon of $N_w = 5$, the algorithm presented in the this section is simulated in UF membrane model. With trial and error, the weighting matrices Q_M and R_M are set to

$$Q_M = 0.01^2 I_{10}, \quad R_M = 500^2 I_{104}.$$

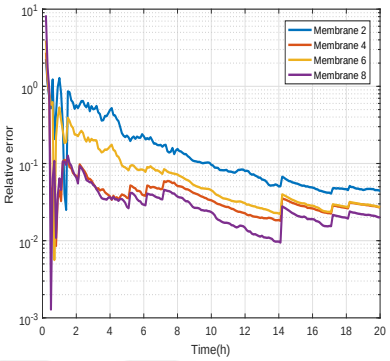
Simulation Results of Monitoring Algorithms

The resistance tracking results for the simulation data set for EKF case1, EKF case 2 and MHE are visualized in Figures 7.6a-7.7a-7.8a, respectively. Here we only present the tracking for the even numbered membranes for brevity. The results for the rest of the membrane

units show a similar trajectory. Moreover, the corresponding relative estimation error (magnitude of error divided by the true magnitude of the resistance) is presented in Figures 7.6b-7.7b-7.8b. As it can be seen in these results, the estimation results are highly accu-

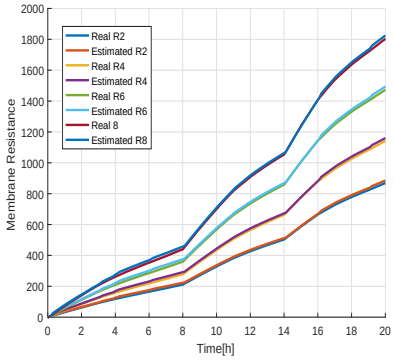


(a) Resistance tracking for even numbered membranes.

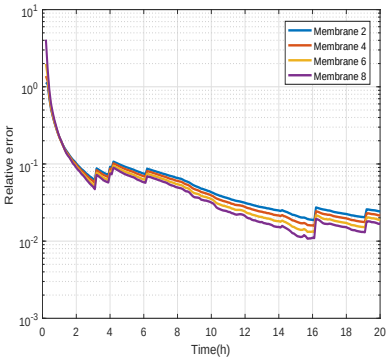


(b) Relative estimation error.

Figure 7.6: Resistance tracking and relative estimation error trajectories with the EKF algorithm (Case 1 - Q_K^{C1} and R_K^{C1}) for the simulation data set.



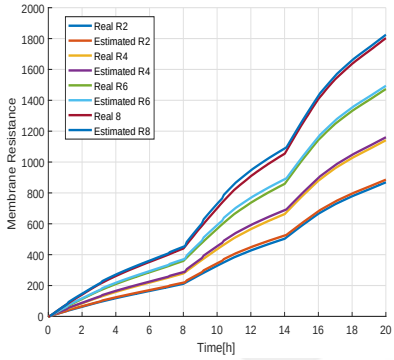
(a) Resistance tracking for even numbered membranes.



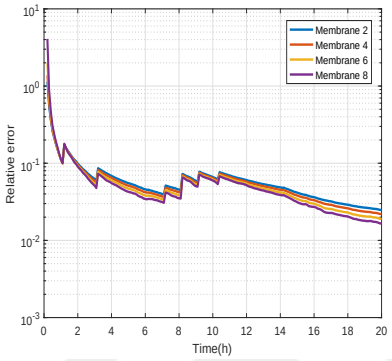
(b) Relative estimation error.

Figure 7.7: Resistance tracking and relative estimation error trajectories with the EKF algorithm (Case 2 - Q_K^{C2} and R_K^{C2}) for the simulation data set.

rate. More specifically, the estimators always keep up with the true fouling trajectories and the relative magnitude of error is decreasing over time if there is no operating point change. Furthermore, it is observed that at time instants when there is a transition in the transmembrane pressure (TMP), the estimation error have sudden jumps. This can be attributed to



(a) Resistance tracking for even numbered membranes.



(b) Relative estimation error.

Figure 7.8: Resistance tracking and relative estimation error trajectories with the MHE algorithm (Q_M and R_M) for the simulation data set.

the mismatch between the estimated and the true parameters and the algebraic mappings inherent to the process model. The key performance indicators (KPIs) which are the outlet protein mass fraction and the mass flow together with their estimates obtained from the three soft sensors are presented in Figure 7.9 and Figure 7.10, respectively. These results indicate that the estimates of the performance indicators are well aligned with the true trajectories. Hence the results support the use of these soft sensors for OMBAs purposes in UF membrane systems.

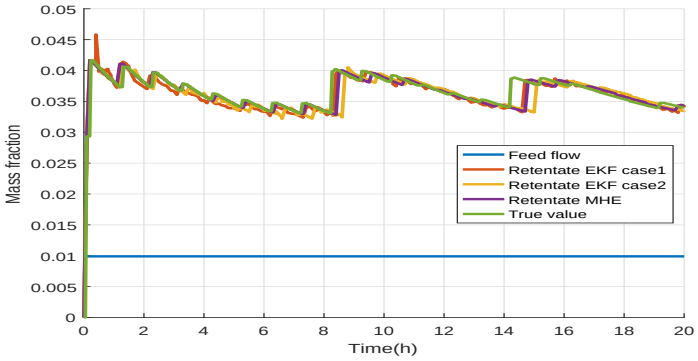


Figure 7.9: True and estimated mass fraction of proteins at the retentate and feed streams for the simulation data set.

.

For the industrial data set, the resistance tracking results are presented in Figure 7.11a,

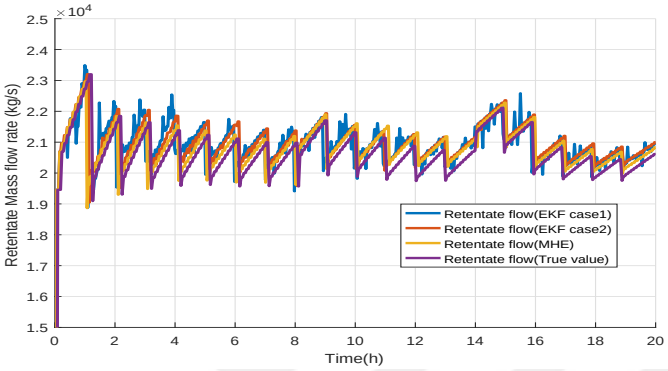
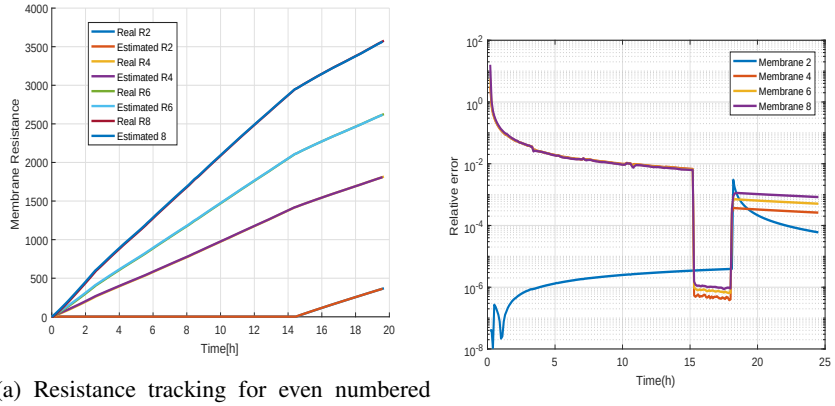


Figure 7.10: True and estimated mass flow rate at the output retentate stream for the simulation data set.

Figure 7.12a and Figure 7.13a for EKF case 1, EKF case 2 and MHE, respectively. Once again, only resistance tracking for evenly numbered membranes is presented for brevity. Moreover, the resulting relative estimation errors are visualized in Figure 7.11b, Figure 7.12b and Figure 7.13b, respectively. The observers run by the measurements gathered

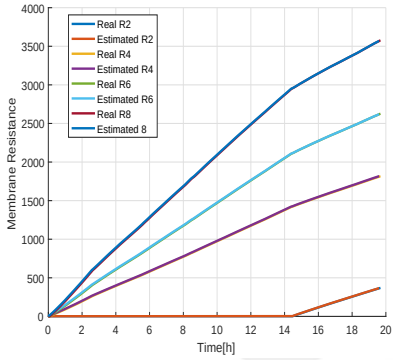


(a) Resistance tracking for even numbered membranes.

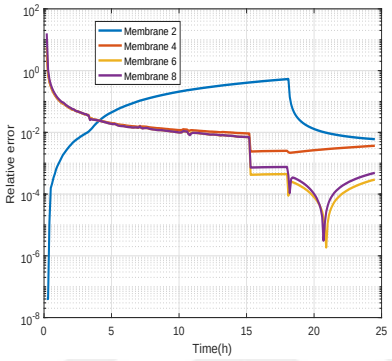
(b) Relative estimation error.

Figure 7.11: Resistance tracking and relative estimation error trajectories with EKF algorithm (Case 1 - Q_K^{C1} and R_K^{C1}) for the industrial data set.

from the industrial plant are all capable of tracking the resistances throughout the batch operation and the estimation error is bounded throughout the simulations. Furthermore, the magnitude of estimation error relative to the magnitude of membrane resistances is reduced over time. Once again, jumps in the estimation error are observed when there are transitions in the TMP profile and when new membranes are switched on. In addition, the estimation

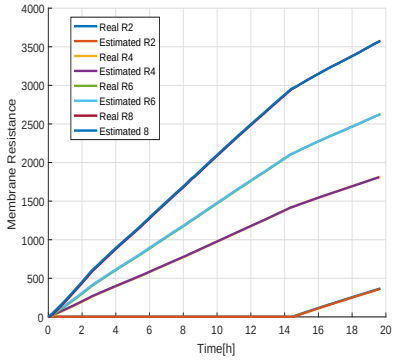


(a) Resistance tracking for even numbered membranes.

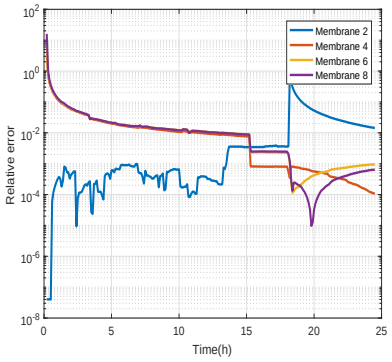


(b) Relative estimation error.

Figure 7.12: Resistance tracking and relative estimation error trajectories with EKF algorithm (Case 2 - Q_K^{C2} and R_K^{C2}) for the industrial data set.



(a) Resistance tracking for even numbered membranes.



(b) Relative estimation error.

Figure 7.13: Resistance tracking and relative estimation error trajectories with MHE algorithm for the industrial data set.

results of KPIs are presented in Figures 7.14-7.15.

The estimation results for KPIs deduced from the industrial data set are not as good as the case of the simulation data set. Specifically, although the estimation of the protein fraction and mass flow at the retentate output stream are accurate at the first few hours of operation, once a new membrane is switched on there is an offset in the estimation of the performance indicators. Nonetheless, the fouling effect is still estimated with high accuracy.

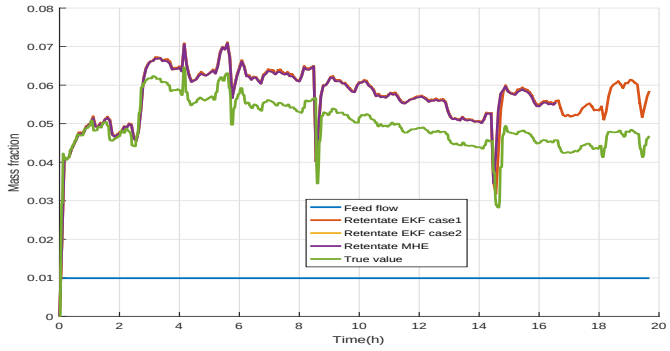


Figure 7.14: True and estimated mass fraction of proteins at retentate and feed streams for the industrial data set.

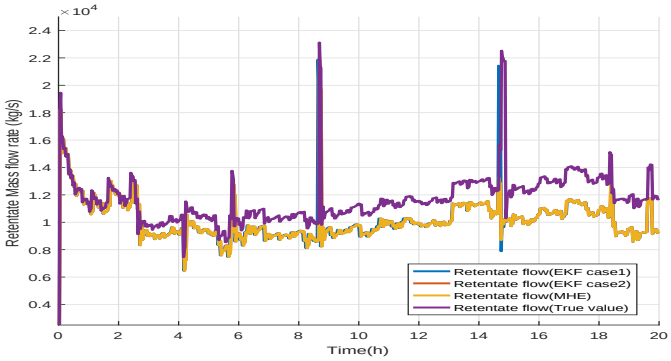


Figure 7.15: True and estimated mass flow rate at the output retentate stream for the industrial data set.

Comparison of Different Observers

In this section, the results presented in the previous section are analyzed and the performance of the estimators is compared.
Although all three estimators perform sufficiently well, by producing comparable results on the estimation accuracy for both of the data sets, there are differences present in the results. In case of the simulation data set, investigating the error dynamics portrayed in Figures 7.6b-7.7b-7.8b reveals that EKF Case 2 (slightly) outperforms the other two estimators. This was expected since adaptive tuning of covariance matrices is implemented, which dynamically adjusts the reliance to the model or the measurement. The other two estimators also perform very well as the estimation error has low magnitude compared to the true value of the resistance and accurate estimates of the performance indicators are obtained. It should however be mentioned that initially, EKF algorithms were implemented

without clipping, meaning that the concentration variables are allowed to take negative values for the EKF based estimation cases. In that case, MHE clearly outperformed both of the EKF algorithms significantly. This means that without (ad-hoc) clipping of the estimates, the performance of EKF Cases 1 and 2 deteriorates. Stated otherwise, with clipping the concentration estimates by imposing the prior physical (positivity) constraints, performance of the EKF algorithms are drastically improved. This clearly shows the strength of MHE and vulnerability of EKF when constraints are present and in fact frequently active.

Furthermore, the fact that EKF algorithms are performing well implies that the linearization errors are not very significant and the linearized system dynamics are faithful representations of the true system at every time instant. It should however be noted that using MHE algorithms with the nonlinear dynamics of the system is expected to yield better results than the EKF based methods. However, such an algorithm would have a higher computational cost since the optimization problem to be solved will have a drastically higher complexity.

In the case of the industrial data set, investigating Figures 7.11b-7.12b-7.13b reveals a similar behavior of the observers for a significant part of the simulations. Addition of new (clean) membranes significantly improves the estimation results. One observation which is actually surprising is that EKF Case 1 outperforms the other two estimators after addition of the last membrane. All three estimators yield stable and bounded error dynamics and low magnitude of relative error in the presence of hybrid behavior. Although the fouling effects are estimated with high accuracy, the estimation of the performance indicators is biased. In other words, the dynamics of the performance indicators seem to be estimated very well by the observers but there is an offset between the estimates and the performance indices. This implies that in this case there is a constant (static) parameter (possibly the empirical membrane resistance variables) or a slowly varying parameter/state (possibly the coefficient α for the dynamic membrane resistance) which is not estimated correctly as a result of the hybrid behavior introduced by switching on a clean membrane. It is expected that the estimation performance can be improved, either by resolving the offline parameter estimation problems discussed in previous chapter or by incorporating online parameter update schemes, as discussed in Section 6.3, to track the estimation error and map it back to the parameters occurring in empirical membrane resistance functions.
A summary of the results is provided in terms of the $\mathcal{L}_2$ norm of the estimation error for resistance, mass fraction of proteins and mass flow rate at the output retentate stream in Table 7.1.

7.2 Low-Level Control of Ultrafiltration Membrane Units

7.2.1 Control Problem in UF Membrane Process

The control goal of the UF membrane separation process is to regulate the composition properties of the final retentate flow, i.e., to achieve the desired protein level at the retentate, and to increase the throughput. We consider an UF membrane process scenario that contains ten parallel membrane loops, is shown in Figure 6.2 where the feed, permeate, and retentate flows are inscribed in. The feed (F_f), permeate (F_p) and retentate flow F_r in the UF unit are shown in the simplified instrumentation diagram of the UF installation in Figure 7.16 and in Figure 7.17 we visualize the input and output variables of a single membrane unit.

Data Set	Physical Variable	Metric	EKF Case 1	EKF Case 2	MHE
Simulation Data	Resistance	$\|e_R\|_2$	910.69	668.33	969.24
	Mass Flow	$\|e_F\|_2$	6.39×10^4	6.43×10^4	6.32×10^4
	Mass fraction	$\|e_x\|_2$	0.1863	0.1775	0.1679
Industrial Data	Resistance	$\|e_R\|_2$	353.78	373.94	417.10
	Mass Flow	$\|e_F\|_2$	6.55×10^4	6.55×10^4	6.54×10^4
	Mass Fraction	$\|e_x\|_2$	0.1739	0.1735	0.1773

Table 7.1: Error norm results for three algorithms applied to both data sets.

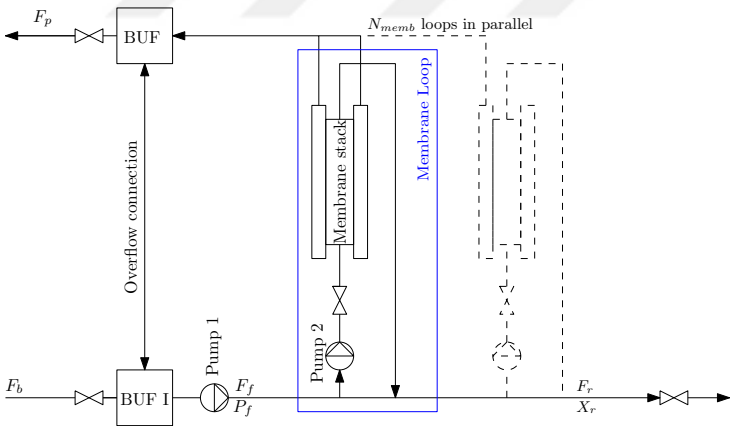


Figure 7.16: Simplified instrumentation diagram of the UF membrane unit, showing a single membrane loop (blue) and the product flows.

During the operation, whey is transported by a pump (Pump 1) which carries the main feed line up to pressure P_f. Then the whey goes through a membrane loop where the pressure is increased by another pump (Pump 2). In a membrane stack, the feed stream is divided into protein-rich retentate and sugar, salts and water rich permeate. The permeate leaves the UF and accumulates in a buffer, and the retentate enters the main feed line again in the direction of the next UF membrane loop. In this way, whey circulates in membrane stacks and eventually the protein rich "whey-protein-condensate" (WPC) is sent to evaporators and dryers for further processing.

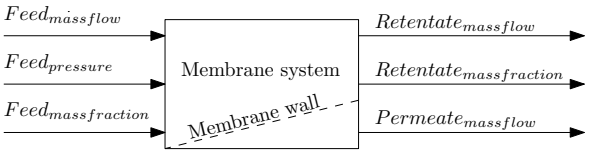


Figure 7.17: Input and output diagram of a single membrane

7.2.2 Improved Operation via Model-based Control Strategies

The measured variables and multiple actuators acting on the membrane system are indicated in Figure 7.18.

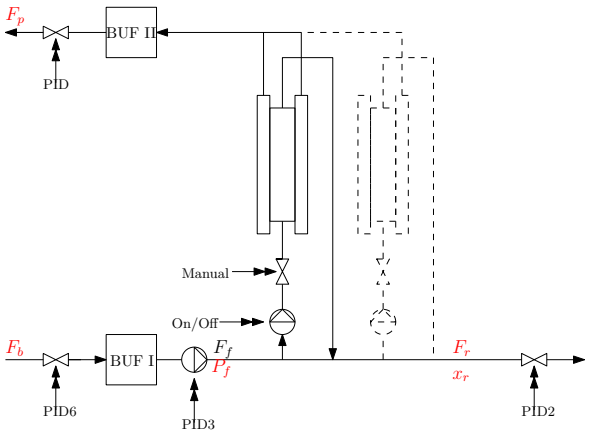


Figure 7.18: The simplified instrumentation diagram showing the measured variables in red color. The actuators and their controllers are indicated by the double arrows.

In the current configuration, the UF membrane units are controlled with three proportional-integral (PI) controllers to manipulate the process outputs. One PI controller (PID3) controls the feed pressure (P_f). The other two controllers (PID2 and PID6) together are used to control the retentate protein concentration (x_r), which can be seen in Figure 7.19. Since the buffer (BUF I) is small and PID6 is tuned to keep the whey level in the buffer at a certain level, the flows F_f and F_b can be assumed equal.

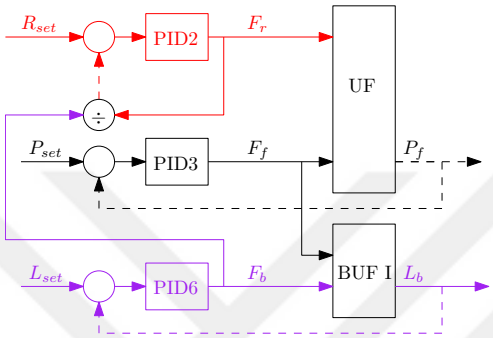


Figure 7.19: Control diagram for the UF process at the plant. The black loop shows the feed pressure control loop. The purple loop shows the buffer level control loop. And the red loop shows the feed/retentate ratio control loop, where the feed value is the F_b.

In this case the feed flow into the buffer (F_b) is used to determine the ratio between the feed and retentate flows as described in Equation (6.4). This ratio is an indicator of x_r, given the feed concentration, as the amount of permeate retracted from the feed flow contains dry matter salts and sugars but almost no protein.

One can improve the control structure by incorporating the instanteneous values of the protein mass fraction in the retentate flow. The protein mass fraction signal x_r is then used in the control loop instead of the VRF. This control loop is then called as a cascade control [214], where the primary (slow, master) loop is identified by the (blue) loop controlling the protein mass fraction of retentate stream. The secondary acting (fast, slave) loop is adjusting the retentate flow according to the ratio set points as indicated by the (red) connections in Figure 7.20, which simplifies the controller configurations by measuring the F_f directly. This excludes the buffer level from the control problem.

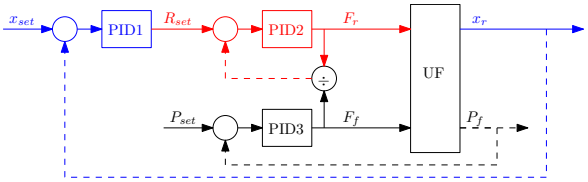


Figure 7.20: The simplification of the control diagram, considering that the F_f is measured and used for the ratio input.

In this section we direct our attention to the closed-loop responses of UF membrane systems with different controllers implemented similar to the reasoning given above. More specifically we design controllers that regulate the protein concentration while having stationary mass flows at the feed and retentate streams. We consider two different PID controller cases as well as two similar MPC controllers with learning feature across the batch iterations to achieve these goals, i.e.,

- PID Case 1 - One PID controller controlling the retentate protein by manipulating retentate flows;
- PID Case 2 - Two PIDs;
- IL-MPC Case - State measurements are available to the MPC, learning update occurs after every batch;
- EKF-IL-MPC - Outputs are filtered by an EKF and the estimated states are used in MPC with learning update that occurs after every batch.

For the PID controllers two separate feedback loops are used. The first feedback loop couples F_r as the manipulated variable and x_r as the controlled variable and the second feedback loop has P_f as the manipulated variable and F_f as the controlled variable. The control scheme used for the PID controllers is shown in Figure 7.21.

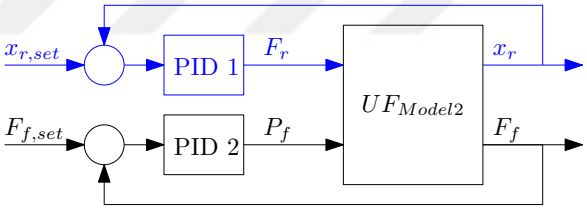


Figure 7.21: PID controllers block scheme, showing two separate control loops.

We first use only the PID 1 feedback loop in Figure 7.21 (Case 1), and then extend the control implementation to both of the loops (Case 2). For the single PID controller case the feed pressure P_f is kept constant at 1 bar. For the cases of IL-MPC with or without EKF, the prediction model inherently considers both controlled variables and decides on F_r and P_f simultaneously, as shown in Figure 7.22.

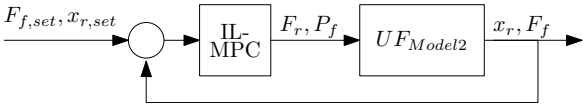


Figure 7.22: IL-MPC controller block scheme, showing one control loop.

During the operation, operators decide to introduce clean membranes based on the instantaneous feed mass flow F_f. The operator switches on another membrane loop when the feed flow drops below the limit threshold, which we mimic in the simulations with the threshold of $58000 kg/h$ for the PID controller cases. For the IL-MPC controller case a predefined switching sequence is used, as the learning aspect of the controller is based on the 'exact' repetition across the batches of operation. A short discussion can be found in [316] which also presents the optimal operating conditions.

7.2.3 Closed-Loop Response of UF Membrane System with Different Controllers

Simulation Results for PID Controllers

To test the performance of the PID controllers, we set a reference set point at 36.8% for the retentate protein fraction and while keeping the feed flow at a minimum flow of $58000kg/h$. Therefore the PID's performance is quantified by the maximal error on the x_r.

In the simulation studies for the UF membrane unit controlled with PID controllers, the system starts with 3 loops switched on at the beginning of the operation and a new loop is added when the F_f drops below a minimum level of $F_{f,min} = 58,000kg/h$. During the start up phase the control inputs are kept constant for a brief time period of $t \in [0, 0.3]h$ to increase the output protein mass fraction levels x_r^{prot}. Then the PID controllers become active for the rest of the operation. The PID controllers are implemented as a discrete time systems, with a zero order hold, for a sampling time of $t_s = 0.001h$. The discrete PID controller is described by the following set of equations,

$$\begin{aligned} e_p(k) &= x_{r,ref} - x_r(k), \\ e_i(k) &= e_i(k-1) + (e_p(k) * t_s), \\ e_d(k) &= \tfrac{e_p(k) - e_p(k-1)}{t_s}, \\ F_r &= k_B + K_p * e_p(k) + K_i * e_i(k) + K_d * e_d(k). \end{aligned} \tag{7.6}$$

In this set of equations e_p, e_i and e_d are the proportional error, the integral error and the derivative error from the reference value of retentate protein mass fraction, respectively. The retentate mass flow F_r is calculated by multiplying these errors with the controller gains K_p, K_i and K_d respectively and adding a bias term K_B to decide on the desired operating point. The tuning of the PID controller parameters is done manually, by evaluating the trajectories and adjusting the parameters to improve the closed loop performance. The tuning parameters of the PID controllers are given in Table 7.2.

The first two single-PID-loop controllers are designed to regulate the deviation from the reference ($x_r - x_r^{\text{set}}$) back to zero. Here we name these two single loop PID controllers according to their closed-loop responses, the mild PID for the milder response on the error and the aggressive PID for more aggressive response. The variable trajectories of the slow and aggressive controller can be seen in Figure 7.23 and Figure 7.24 respectively.

The mild controller has smoother input and output trajectories, due to the slowly accumulating error signal. The aggressive controller causes fluctuations in the input due to a stronger integration action and an added derivative action. However, the maximum error when a clean membrane loop is switched on, is smaller compared to the mild controller trajectories.

During the operation of the whey filtration process, we prefer to have a steady mass flow, whereas a constant feed pressure results in the saw-tooth profile as in Figure 7.23. Therefore the operators regularly adjust the feed pressure manually to flatten the feed mass flow profile. By manipulating the feed pressure with a second control loop the manual control action can be replaced. Then the second PI controller loop is implemented to control the feed flow F_f to a steady flow by manipulating the feed pressure. As the UF membrane unit has limits on the applied pressure levels, the controller's output is clipped to the range $P_f \in [0.8, 1.5]bar$, by using an anti-windup scheme. The tuning parameters for the double

	K_B	K_p	K_i	K_d
PID Case 1 Mild Config.	$12 \cdot 10^3$	$-100 \cdot 10^3$	$-2 \cdot 10^6$	0
PID Case 1 Aggresive Config.	$12 \cdot 10^3$	$-100 \cdot 10^3$	$-20 \cdot 10^6$	$-10 \cdot 10^3$

Table 7.2: Tuning parameter values for single loop PID controller case.

	K_B	K_p	K_i	K_d
PID 1	$12 \cdot 10^3$	$-100 \cdot 10^3$	$-2 \cdot 10^6$	0
PID 2	1	$5 \cdot 10^{-5}$	$5 \cdot 10^{-4}$	0

Table 7.3: Tuning parameter values for two PID controllers case.

PID case are provided in Table 7.3. The simulation results for the double PID controlled UF membrane system is visualized in Figure 7.25.

Simulation Results for IL-MPC Case

In this section we combine a model predictive controller (MPC) with an iterative learning controller (ILC) to control the UF system, for two different cases. In this configuration MPC is used for tracking the reference signals, while the ILC is used for learning a better operating trajectory across the batch iterations. We process tracking errors in the previous batches to adjust the operating trajectory, while MPC reduces to correcting the output tracking errors in the current batch operation, similar to the case in [96]. By combining MPC and ILC control structures, two IL-MPC is implemented which has enhanced prediction capabilities by evaluating the tracking errors from the past batches and improving the operating in the current batch. The controllers are differing from each other according to the accessibility of the state vector. In the first case, we assume that the states are directly accessible to the controller, whereas in the second case, we filter the outputs to deduce the state values.

To achieve the proposed control structure, the UF model is linearized along an optimal operating trajectory similar to the monitoring case study, where the resulting local state space description is used at the output predictions of the future time instants to predict the output tracking error. The predicted output deviation from the linearization point of i^{th} batch, y_k^i, is defined as,

$$y_k^i = C_k^{i-1}\left(x_k^i - x_k^{i-1}\right) + D_k^{i-1}\Delta u_k^i + y_k^{i-1}. \tag{7.7}$$

where x_k^i is the state vector at time instant k for the current batch i, Δu_k^i correction to the input of the previous batch u_k^{i-1} at time instant k. The measured output variable is y_k^{i-1}, i.e., $[x_r, F_f]^\top$, of the previous batch $i-1$. The C_k^{i-1} and D_k^{i-1} are the state space representations of the previous batch $i-1$ at time instant k. Hence the predicted output deviation from the linearization point is based on the linearization of the dynamical system, i.e., C_k^{i-1}, D_k^{i-1}, and the measured output (y_k^{i-1}) of the previous batch. We can describe

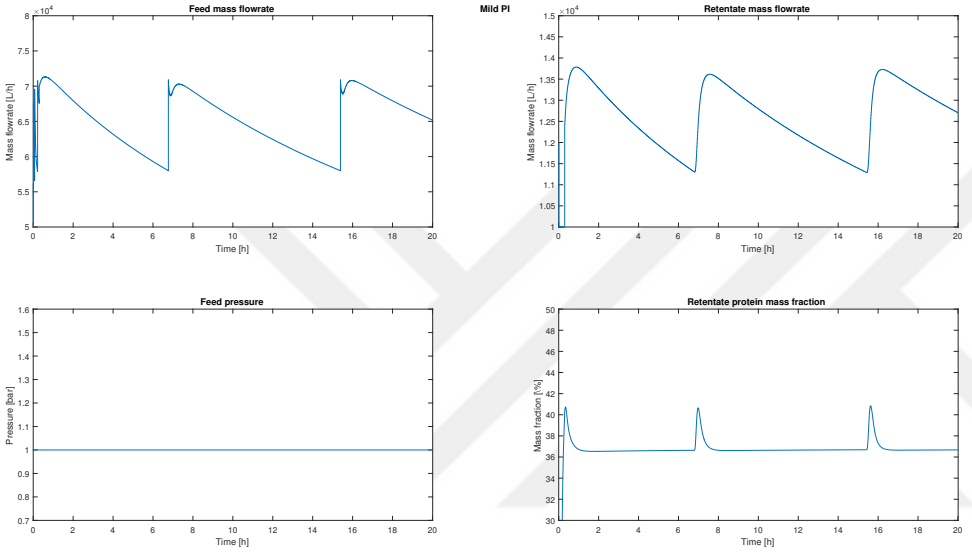


Figure 7.23: Manipulated and controlled variable trajectories for the case of the mild PI controller, with a smooth control behaviour

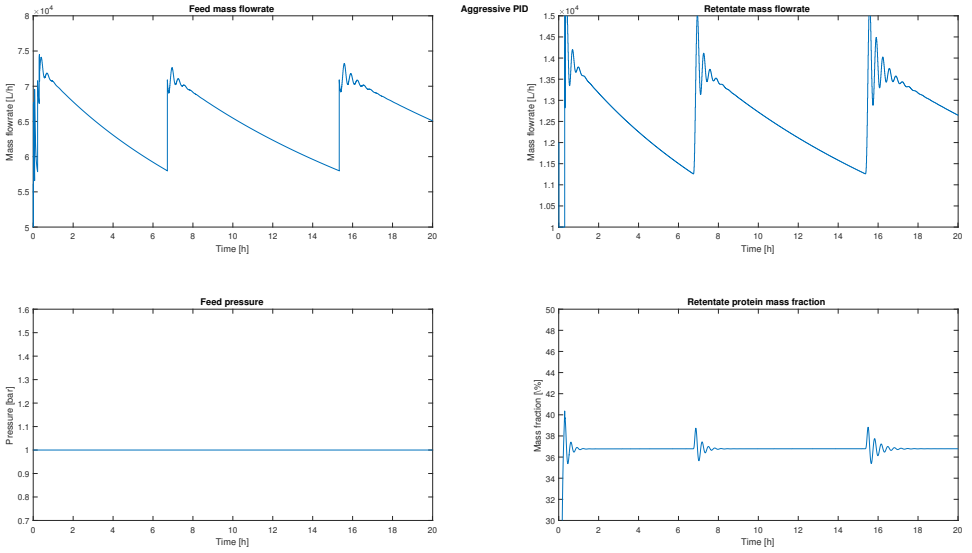


Figure 7.24: Manipulated and controlled variable trajectories for the case of the aggressive PID controller, which shows oscillatory behaviour

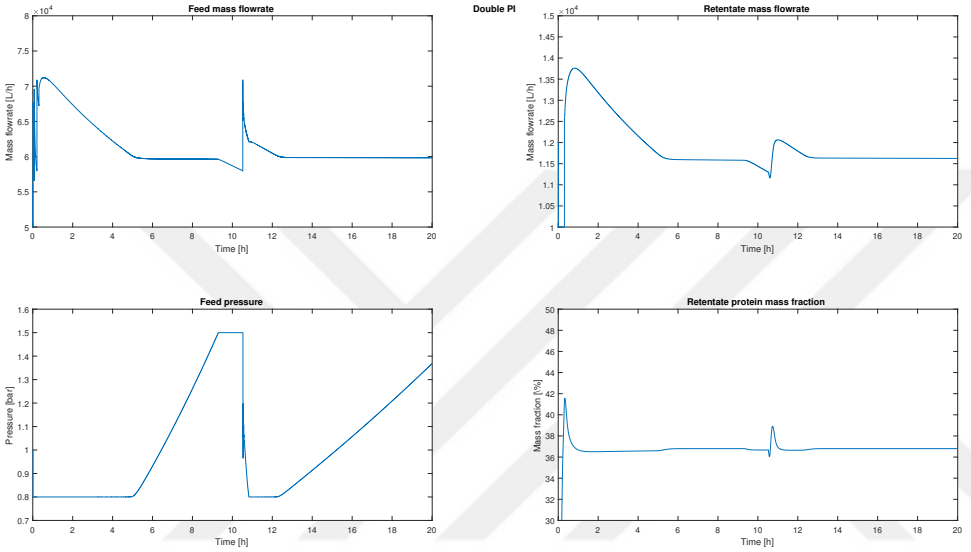


Figure 7.25: Two PID controllers response. The feed mass flow is regulated around a constant flow, while satisfying at the desired retentate concentration levels, as long as the pressure input is within the bounds

the evolution of deviation dynamics as

$$\Delta x^i_{j+1|k} = A^{i-1}_{j|k}\Delta x^i_{j|k} + B^{i-1}_{j|k}\Delta u^i_{j|k},$$

where

$$u^i_{j|k} = u^{i-1}_{j|k} + \Delta u^i_{j|k}, \quad x^i_{j|k} = x^{i-1}_{j|k} + \Delta x^i_{j|k}$$

and j is the index over the prediction horizon. To compute the optimal input actions which steer the system to the given reference trajectory, which is constant over batch iterations i r^{set}_k, at every time instant k the following nominal MPC problem is solved;

$$\mathcal{P}^{\text{IL-MPC}} : \begin{cases} \min\limits_{\Delta U_k} & \sum\limits_{j=0}^{N_p-1} \left((\bar{y}^i_{j|k} - r^{\text{set}}_{j|k}) Q (\bar{y}^i_{j|k} - r^{\text{set}}_{j|k}) + \Delta {u^i_{j|k}}^\top R \Delta u^i_{j|k} \right) \\ \text{subject to} & \Delta x^i_{j+1|k} = A^{i-1}_{j|k}\Delta x^i_{j|k} + B^{i-1}_{j|k}\Delta u^i_{j|k}, \\ & \bar{y}^i_{j|k} = C^{i-1}_{j|k}\Delta x^i_{j|k} + D^{i-1}_{j|k}\Delta u^i_{j|k} + y^{i-1}_{j|k}. \\ & \Delta u^i_{k,\min} \le \Delta u(k) \le \Delta u^i_{k,\max}, \\ & \Delta u^i_{k,\min} = u_{\min} - u^{i-1}_k, \quad \Delta u^i_{k,\max} = u_{\max} - u^{i-1}_k, \\ & u^i_{j|k} = u^{i-1}_{j|k} + \Delta u^i_{j|k}, \quad x^i_{j|k} = x^{i-1}_{j|k} + \Delta x^i_{j|k}, \\ & \Delta U^i_k = \left[\Delta {u^i_{0|k}}^\top, \Delta {u^i_{1|k}}^\top, ..., \Delta {u^i_{N_p-1|k}}^\top \right]^\top \\ & j = 0, \dots, N_p - 1. \end{cases} \tag{7.8}$$

where the $A^{i-1}_{j|k}$, $B^{i-1}_{j|k}$, $C^{i-1}_{j|k}$ and $D^{i-1}_{j|k}$ are linearization matrices stored from batch number $i-1$, and are used to predict the system output from current state $x^i_k = x^i_{0|k}$ till the end of the prediction horizon $k+N_p-1$. The prediction horizon is chosen as $N_p = 5$. We select a short prediction horizon since we approximate the current linearization with the linearization of UF membrane dynamics from previous batches, hence long prediction horizons are expected to yield wrong predicted trajectories as the control input, and hence the state trajectories, accumulates the learning action after every batch.

Next we present the simulation results for the UF membrane system controlled with IL-MPC. As the MPC controller requires data from a previous batch operation to base the predictions on, an initial run is performed in which the feed pressure and retentate mass flows (inputs to the system) are kept constant. To keep the simulation times practical for this simulation study, the sampling time for MPC controller is chosen as $t_s = 0.25h$ which results in approximately $20min$ of simulation for a single batch ($20h$) of operation for the 10 membrane stacks. This is due to the large state space dimension around the linearization point. The local behaviour of each membrane is described by 13 states, which then yields a 2 input, 2 output and 130 state UF membrane system. The second batch is started with the cost weighting matrices as in Equation (7.9).

$$Q = \begin{bmatrix} 10^{-3} & 0 \\ 0 & 10^{10} \end{bmatrix}, \quad R = \begin{bmatrix} 0 & 0 \\ 0 & 10^{-2} \end{bmatrix}, \tag{7.9}$$

where the parameter $Q_{1,1} = 10^{-3}$ is the cost on the F_f error and the $Q_{2,2} = 10^{10}$ is the cost on the x_r mismatch. The matrix R contains the cost on the input change rate from the last batch, in which a higher cost would decrease the abrupt changes in the inputs across batches.

The simulation results for the IL-MPC implementation with the mentioned parameters are shown in Figure 7.26. In this figure the initial batch with constant inputs is shown together with the six next batches, showing the learning aspect of the controller as the trajectory converges to the reference. The trajectories of controlled and manipulated variables are visualized in Figure 7.26. The error values, the mismatch between the reference and the instantaneous value of the controlled variable, can be minimized by the predictive controller. By using MPC one can obtain lower maximum errors compared to the PID controlled UF membrane process case.

As an extension of the IL-MPC control strategy, we also consider the case where the states are not accessible to the controller. In many cases, state measurements are not fully available to the controller directly, but deduced from a set of measured variables. In this case we reconstruct the states by incorporating an estimator into the loop. For this purpose an EKF is added into the control loop to reconstruct and send the states to the MPC controller. To compute the state estimates during run time, we run a second UF membrane model as soft-sensor in parallel to the true process. Within the simulation environment to mimic the actual plant conditions, we perturbed the optimal inputs calculated from MPC with disturbance signals. However the observer model does not have access to the perturbed values steering the true process. Under this setting the simulation results for the IL-MPC controller with EKF state estimator is visualized in Figure 7.27. The simulation results indicate that

the process trajectories for the output feedback IL-MPC case are similar to the trajectories induced from state feedback IL-MPC case. The estimation routine introduces small perturbations in the closed-loop system trajectories, however these undesired effects are suppressed with MPC controller. One possible drawback of the fluctuations can be observed if the learning controller has not smoothing (low-pass) effects and hence carries the noisy behaviour to the next batch iterations. However, we overcome this issue by filtering the errors with a unit gain (at DC) low pass filter which removes the fluctuations. Furthermore the closed-loop performance of both of the IL-MPC controllers can be increased, either by reducing the sampling time as the computation time only needs approximately $0.014h$ and the current sample time is $0.25h$, or by extending the prediction horizon in the expense of the computation time, or, finally, by incorporating high order learning techniques instead of incorporating the learning action as keeping the input signal from the previous batch.

Comparison of the PID, IL-MPC and Kalman IL-MPC Controllers

All controllers show similar closed-loop behaviours and performance results. To highlight the differences in closed-loop performances, the manipulated and controlled variables of the controllers are plotted in Figure 7.28. Due to the learning action, the IL-MPC and EKF-IL-MPC controller schemes improves the error rejection after every batch iteration, hence as the batch iteration number increases, the input-output trajectories improves. For comparison purposes the fourth iteration of both IL-MPC controller input-outputs and the PID controller input-outputs are shown in the Figure 7.28. In summary, we note that

- The mild PID shows the largest retentate protein mass fraction x_r error when switching on a clean membrane;
- The aggressive PID decreases the maximal error on the desired protein fraction x_r, but introduces larger and longer fluctuations during the operation;
- The double PID ends up with a similar level of maximal error on the desired protein fraction x_r as the case with aggressive PID, however in the double PID case, the trajectories converge to the references with less fluctuations, since the necessary action to control the retentate protein mass fraction is distributed among mass flow and pressure levels. This controller also yields a constant feed and retentate mass flow by increasing the pressure which overcomes the performance deterioration due to fouling;
- The IL-MPC shows similar behaviour as the double PID, since the MPC acts as a controller that is manipulating both of the input actions according to the outputs. Moreover IL-MPC controller is able to yield a smaller maximum x_r error and this error is decreasing monotonically over the batch operations;
- The realistic case for IL-MPC, the IL-MPC with EKF, leads to process trajectories that are similar to that of the IL-MPC while the input actions, and hence the output trajectories, are corrupted. However, the corrupted inputs are not leading to undesired responses in UF membrane operation, possibly due to the low learning rate. This keeps the operation around similar operating trajectories across batches.

7.3 Conclusions on Online Model-based Applications for Ultrafiltration Membrane Processes

In this chapter, we have discussed the online monitoring and (low-level and batch-to-batch) control of UF membrane processes. For the monitoring problem, we compare three soft-sensor implementations for the membrane process. Observer design for membrane operation is required to track the accumulating fouling effects. We implement extended Kalman filter (with two different choices of Kalman weights) and moving horizon estimation based soft-sensors for two data sets for the whey process, one set of data generated within the simulation environment and one set of measurements are gathered from an industrial plant. All three soft sensors are providing comparable estimates, while the implementation time and tuning (of weights) experience becomes increasingly important. The least effort consuming case (EKF-Case 2) yields the best results for the simulation data set case, while for the industrial data set case, this EKF case shows the worst estimation results out of three soft sensors.

The second part of this chapter presents the results on the control aspects of the UF membrane unit. We, first, present a brief discussion on the (manual) control strategies of the operators and controllers in the UF process. We have designed two types of controllers and tested them in simulation environment using the dynamical model. The first type of controllers are consisting of classical (PID) controllers, where the double PID configuration demonstrated the best results. Next, a model predictive controller with learning action across batches of operation is presented, which is both capable of estimating the future outputs and also learning from the errors based on the previous batch data.

- In all of the estimator cases, converging and reliable estimates for the states are achieved. The soft sensors are, also, able track the key performance indicators remarkably while hybrid changes are happening during the operation (such as switching on of membranes).
- To increase the monitoring performance and decrease the control structure complexity, the feed mass flow and its concentration should be measured.
- The feed pressure should be controlled(via as simple as a PID) to operate with a steady feed mass flow.
- An IL-MPC could be used to gain a lower maximal protein concentration in the retentate errors when a membrane is switched on.
- The IL-MPC performance could be increased by implementing an integral action reducing the steady state errors.

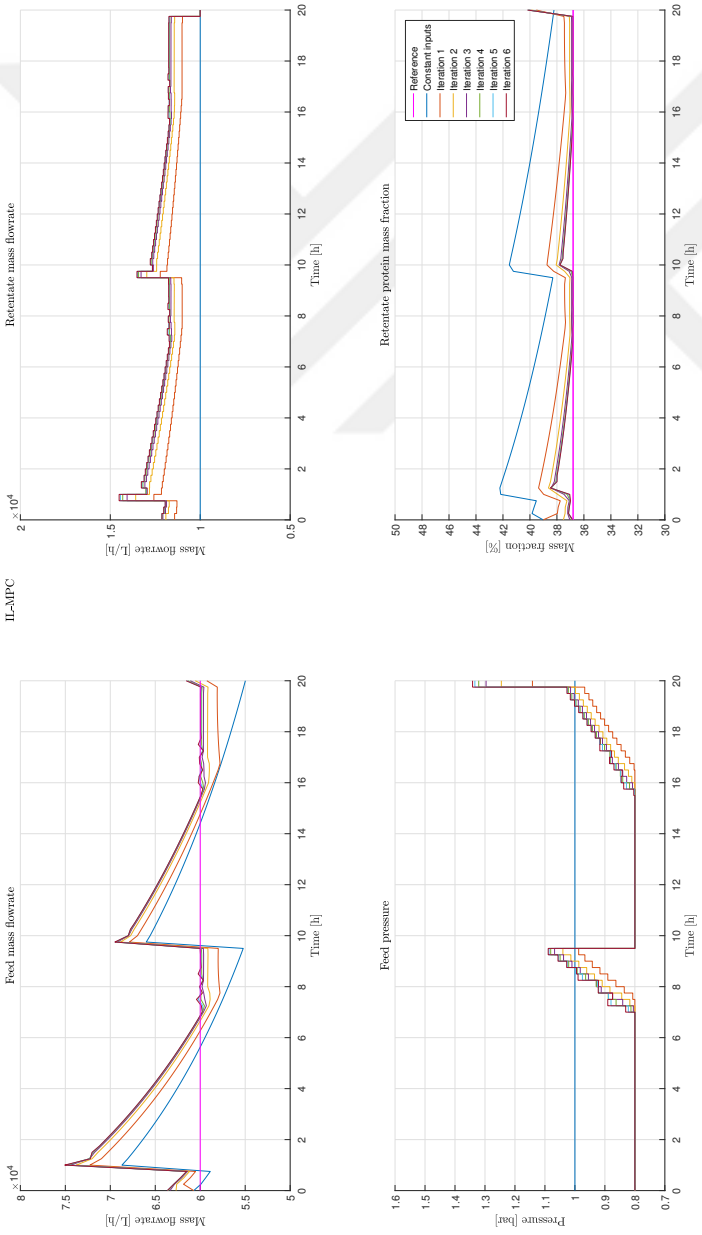


Figure 7.26: Simulation of IL-MPC, which shows the converging of trajectories towards the reference.

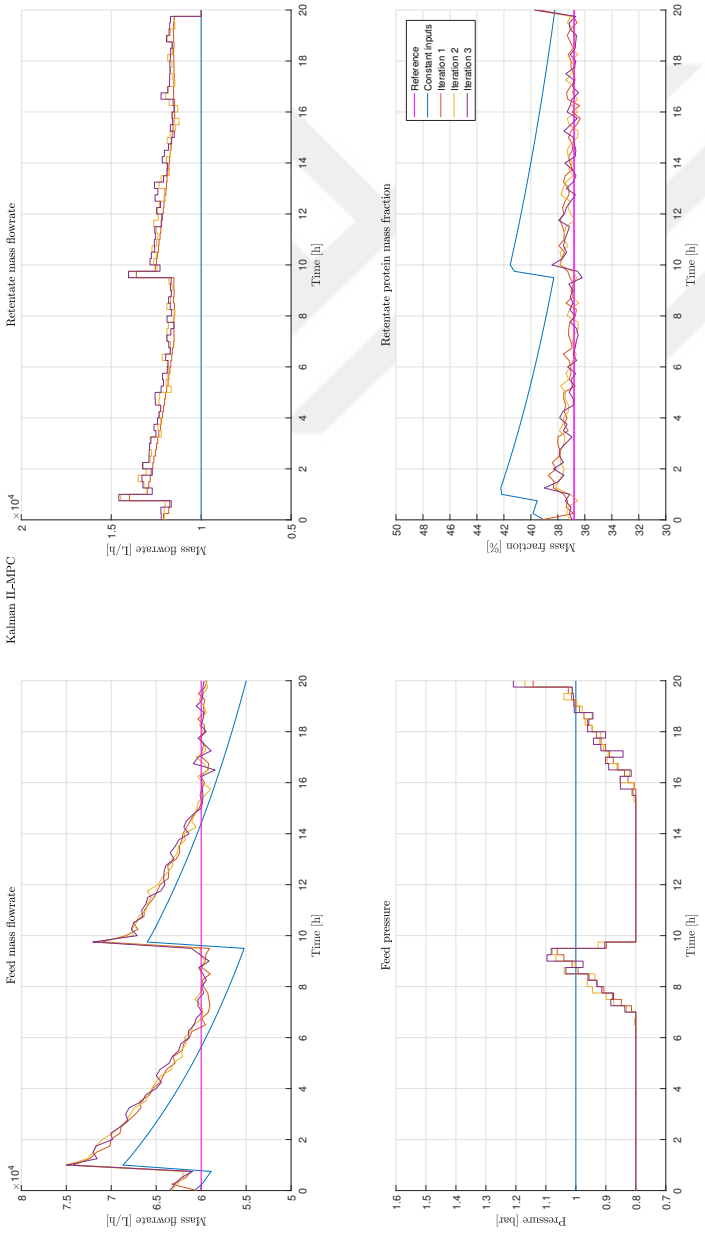


Figure 7.27: Manipulated and controlled variables of the UF model where a IL-MPC controller with Kalman filter, running a different UF model in parallel, is used.

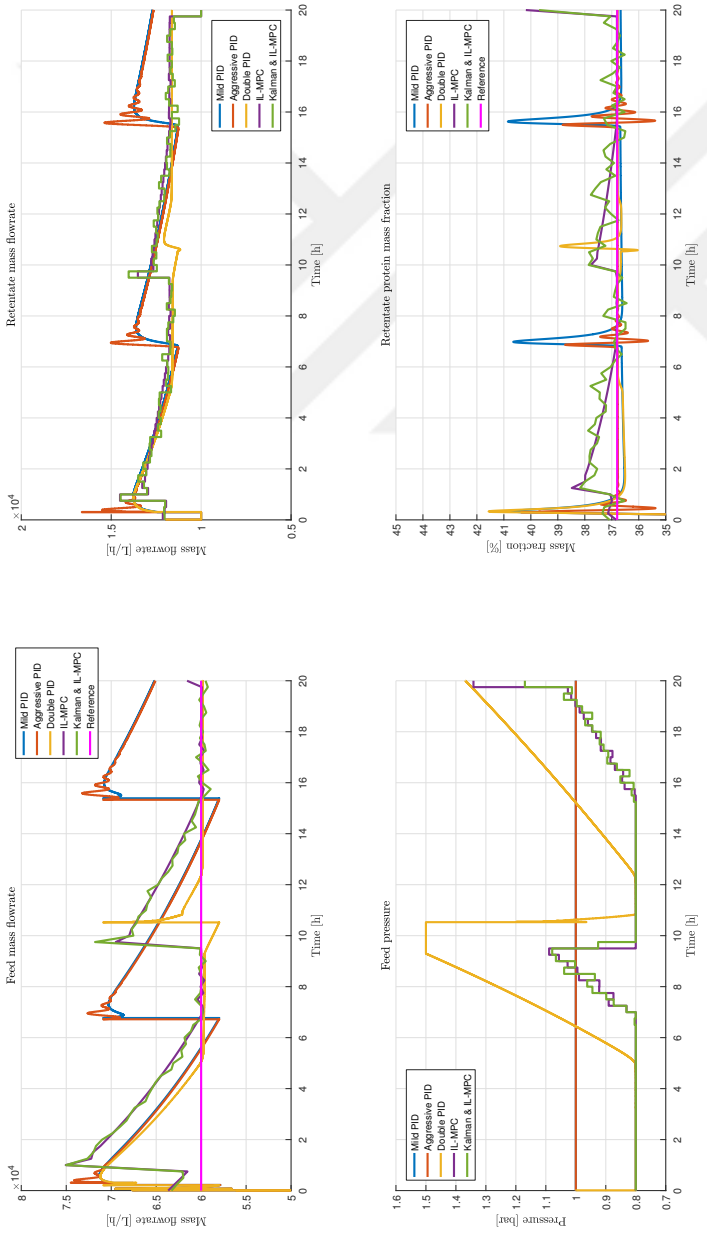


Figure 7.28: Manipulated and controlled variables of the different controllers, showing how the IL-MPC is capable of decreasing the maximum error to the reference. For the IL-MPC and Kalman IL-MPC only the fourth iteration is shown.

Chapter 8

Scheduling of Unit Operations in a Whey Separation Process

The Way never acts, yet nothing is left undone.

Dao De Jing 37

As a high-level model-based application, the scheduling of unit operations within the whey protein separation process is discussed in this chapter. We consider the whey plant with its subprocesses to discuss the safe scheduling problem in operating large and multi-unit plants. Safe operation means that the operational constraints, based on either the timing or physical limitations of the subprocesses are explicitly taken into account while constructing the operating schedule. A graph theory based approach is taken to guarantee temporal constraints on each of the unit operations, while reachability analysis is used to keep the plant operation continuing without any limitation violations. We make use of simulation examples to show the computational benefits of the graph based scheduling approach.

8.1 Introduction to Plant-wide Scheduling

In many of the manufacturing industries, it is common that the process, or enterprise, consists of several consecutive unit operations (UOs) that are subprocesses with complex dynamics. Furthermore, these UOs are turned on or off due to various planned activities, such as cleaning or maintenance. To achieve an efficient operation the schedules should be optimized in such a way that while the profits are maximized, the safety and performance constraints are guaranteed to be satisfied. Within these scheduling problems there exist (i)

[0]Substantial content of this chapter is also published or presented in:

- M.B. Saltik, N. Athanasopoulos, L. Özkan and S. Weiland. Safety Analysis for a Class of Graph Constrained Scheduling Problems.
- M.B. Saltik, N. Athanasopoulos, L. Özkan. Enterprise-wide Optimization.
- S. van Gameren, M.B. Saltik, L. Özkan, S. Weiland. Recovery scheduling for industrial processes using graph constraints.

logical rules for the decision variables (constraints on the schedules) over time and, (ii) a highly structured dynamics of the UOs. This is also the case in the whey protein separation plant. The whey process plant consists of a series of unit operations (UOs) as visualized in Figure 6.1. These UOs are turned on or off to cycle between operational and nonoperational phases during the process, see Figure 8.1 for an indication of operation times. The on-off status of an UO results in operational time window constraints and introduces integer valued actions, since an UO can be either operational or not for some amount of time but no other option exists. Furthermore, not all UOs operate (or, equivalently, be nonoperational) on the same time windows. The evaporator UO operates with time constants that differ from the time constants that one observes with membrane stacks. Similar differences between operating windows necessitates buffer tanks to accumulate outlets of UOs, before the operation starts in subsequent UOs. Furthermore, it is desired that the thin whey enters and leaves the process as quickly as possible. The necessity of a high processing rate is due to possible contamination of the eventual product, the whey powder. One of the important aspects of whey filtration process is to schedule the UOs without overflowing (or equivalently depleting) the buffer tanks. In the next section we present a scheduling technique that makes use of graphs as constraints and establishes links between buffer tank levels over time instances and the constraints on the on-off status of UOs. An optimization problem that takes the schedules as the decision variables is, in general, a mixed integer linear programming (MILP) problem, due to the integer valued actions in the schedules, which we optimize over. Formulating the scheduling problem as an MILP problem introduces computational challenges, see [325]. Indeed, many production planning problems have been shown to be NP-complete ([125]). Standard approaches concern either relaxing the binary optimization variables of the MILP problem into real ones [14], or constructing a decision tree which is expanding at each time instant, hence, affected by the curse of dimensionality ([26]). In general, the number of variables severely affects the solvability of scheduling problems. In industry, heuristic methods are applied that can end up far from optimality [280]. Moreover, these methods may not explicitly take the state and scheduling

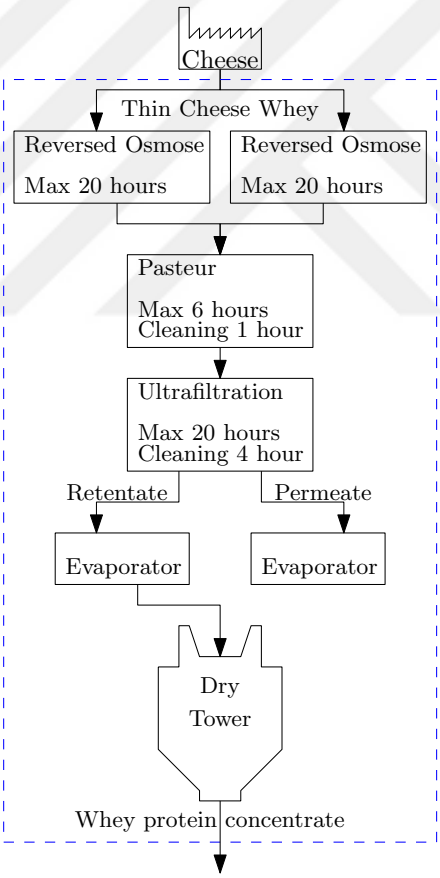


Figure 8.1: Schematic diagram of whey protein separation line.

constraints into account. Hence, one of the core directions in the scheduling area is the low complexity modeling of the problem. Such examples can be given as [204, 328][1].

In this chapter we present a safe scheduling algorithm which is based on a recently introduced switching based modeling approach by using directed graphs. Motivated by the work of [12], which considered the stability problem of switched linear systems under constrained switching, we express the operational constraints on the scheduling variables via weighted directed graphs. This approach allows to express scheduling constraints in a time-invariant setting, which simplifies the subsequent analysis. In these graphs, each vertex corresponds to an “on" or “off" status condition of subprocess while carrying the memory information, that is, we explicitly use the temporal constraints to construct a digraph which represents only the allowed transitions of the considered UO. The edges stand for logical rules inherent to the system, which models the switching between on-off conditions or staying in either condition. For example one can consider the logical rules such as maximal operational time, minimum offline time span, which shapes the graph of the considered UO. For safe operation guarantees, in particular, we utilize forward and inverse mappings of the dynamics applied on sets in order to find a sequence of safe sets. These sets are associated with scheduling paths generated from directed graphs, accounting for admissible values for operation and on-off configurations of UOs, respectively. Additionally, the structure of the dynamic process and of the state constraints allow us to use set-theoretic operations efficiently to simultaneously compute feasible scheduling profiles and the related safe initial condition sets on the state space. Furthermore, these sets are cast to be periodic over the scheduling path, by selecting the schedules as cycles of the constructed graphs. Similar operations can also be used to re-schedule the UOs in the face of undesired variations in the operation

Here, we solve a series of simple MILP problems in order to identify cycles in the directed graphs describing the process for which the sum of the labels over the cyclic paths is equal to zero. Next, we compute the sequence of admissible sets in the state space, for which the periodic scheduling profile can be implemented without violating the state constraints. Finally, using backward reachability operations with respect to the dynamics and the directed graph, we develop an iterative algorithmic procedure for enlargement of the safe set. Under this setting, the conventional decision making algorithms are reduced both in size and complexity. This allows the user to extend the predictions of the schedules up to infinitely long horizons for periodic schedules.

In the next subsection, 8.2, we describe the modeling of the overall process and the scheduling constraints, and then formulate the safe scheduling and steering problems. In Section 8.3 we present the proposed algorithms that construct safe schedules and show the theoretical properties of these safe sets. We apply the proposed approach to a simplified separation process in Section 8.4.

8.2 Modelling of Graph Constrained Scheduling Problems

Here we consider processes consisting of several connected UOs. Since the UOs in the production line are turned on and off in an asynchronous way, buffers are placed between

[1] A comprehensive study of MILP problems in process system engineering area can be found in [4], while works such as [112, 213] consider specifically scheduling problems. Scheduling problems are also common in operations research area, such as in supply-chain management [106] and in enterprise-wide optimization [143].

UOs. These buffers accumulate the outlet of previous UO until the next UO becomes operational again, as shown in Figure 6.1. Although more general connection schemes (parallel or feedback connections) of UOs and buffers can be modeled in the formalism introduced here, we consider only serial connections, because this connection scheme is the closest ti the case in whey process.

8.2.1 System Dynamics in Scheduling Problem

For the process visualized in Figure 6.1 we assume that the steady-state operation conditions are dominant in the behavior of the process. This means that deviations from the known optimal operating conditions are undesired and suppressed. This is the case in many industrial processes, e.g., see [382]. Hence, we model the process as a linear time-invariant system, in which the input vector takes its values from the discrete set $\{0, 1\}$. Next, we consider discrete time dynamics by applying mass balances for the buffer tanks. The inlet and outlet of each UO is denoted by $\phi_k^i, \bar{\phi}_k^i$, respectively, where $k \in \mathbb{Z}_{\geq 0}$ is the time index and $i \in \mathbb{Z}_{[1,n]}$ is the index of the i^{th} UO in the process loop. The external input flow to the process is denoted as ϕ_k^{in}, $k \in \mathbb{Z}_{\geq 0}$. In this work, we assume that the inlets and the outlets $\phi_k^i, \bar{\phi}_k^i$ are constant for all $k \in \mathbb{Z}_{\geq 0}$. Then, the following linear difference equations describe the (mass) levels in the buffer tanks,

$$x_{k+1}^i = x_k^i + \alpha_k^{i-1}\bar{\phi}_k^{i-1} - \alpha_k^i \phi_k^i, \tag{8.1}$$

where $i \in \mathbb{Z}_{[1,n]}$ is the index of each buffer tank, $x_k^i \in \mathbb{R}^n$ is the quantity accumulated in buffer tank i at time sample k, e.g., the mass hold up in a buffer tank and n is the number of buffer tanks. The scheduling variable α_k^i denotes the mode of the i^{th} unit operation at k^{th} time instant. If the i^{th} UO is operational (On) at time sample k then $\alpha_k^i := 1$, otherwise it is not operational (Off) abd we define $\alpha_k^i := 0$. We denote the decision vector as

$$\alpha_k^T = \begin{bmatrix}\alpha_k^0 & \alpha_k^1 & \dots & \alpha_k^n\end{bmatrix}^T,\ \alpha_k \in \mathbb{Z}_{[1,n]}^{n+1},\ k \in \mathbb{Z}_{\geq 0}.$$

We introduce an integer state, denoted by σ_k, that is standing for all different On-Off configurations of UOs including the memory effects, i.e., all possible α_k vectors that satisfy the known operating windows. Hence σ is utilized to realize satisfaction of the (timed) constraints on α variables (On-Off). The modeling formalism results in the following state space model.

$$x_{k+1} = Ax_k + b_{\sigma_k}, \tag{8.2a}$$

where

$$\begin{aligned} A &= I_n \\ b_{\sigma_k} &:= B_k\alpha_k, \\ &= \begin{bmatrix} \phi_k^0 & -\phi_k^1 & 0 & \dots & 0 \\ 0 & \bar{\phi}_k^1 & -\phi_k^2 & \dots & 0 \\ \vdots & \ddots & \ddots & \dots & \vdots \\ 0 & \dots & \dots & \bar{\phi}_k^{n-1} & -\phi_k^n \end{bmatrix} \alpha_k, \end{aligned} \tag{8.2b}$$

Here, the $(i-1)^{\text{th}}$ row of B_k matrix is constructed from the values of $\bar{\phi}_k^{i-1}$ and ϕ_k^i with i standing for the corresponding UO[2]. Furthermore, the vector b_{σ_k} indicates both the On-Off configurations and the inlet-outlet parameters of the system such as the ϕ_k^j and $\bar{\phi}_k^j$ values, which can be time-varying.

8.2.2 Time and Buffer Constraints in Scheduling Problem

We consider two different types of constraints, namely, the state and the scheduling constraints. The constraints on states represent the physical limitations of the buffers between the UOs. In general, these constraints are expressed by linear inequalities in the state space, forming polyhedral sets. The state variables are decoupled, since buffers are not effecting each other, we consider constraints expressed by hyper-rectangles, i.e.,

$$x_k \in \mathbb{X} := \{x \in \mathbb{R}^n : x_{\min} \leq x \leq x_{\max}\},\ \forall k \in \mathbb{Z}_{\geq 0}. \tag{8.3}$$

where inequalities are understood entry-wise.

The second type of constraints that we consider are the temporal constraints on the scheduling variable α_k. The a-priori knowledge of the operational structure casts the scheduling variable to evolve with regard to these scheduling constraints. To give an example, for each UO we consider an admissible scheduling pattern as described in the left part of Figure 8.2. These valid transitions model the mode of UO in the next time instant. Hence the On-Off configurations of each time instant $k \in \mathbb{Z}_{\geq 0}$, i.e., α_k, can be represented as a sequence of introduced graph state $\sigma_k \in \mathcal{V},\ k \in \mathbb{Z}_{\geq 0}$ which takes values of nodes and evolves over a deterministic graph. Hence, we introduce more nodes, see right hand side of Figure 8.2, which allows us to have temporal constraints explicitly taken into account. Furthermore, we assign labels $w(i,j)$ on each edge $(i,j) \in \mathcal{E}$ of the graph, corresponding to the admissible values the scheduling variable $b_k \in \mathbb{R}^n$ might attain. Consequently, the scheduling constraints can be expressed as $b_{\sigma_k} \in \mathcal{W}_{\sigma_{k-1}}$ where

$$\mathcal{W}_{\sigma_{k-1}} := \{b : \exists(\sigma_{k-1}, j) \in \mathcal{E} : b = w(\sigma_{k-1}, j)\}. \tag{8.4}$$

We proceed with modeling all the allowed mode transitions in a graph composed from all UO graphs, which is called the *enterprise graph*. To this end, we define the Kronecker product between two directed graphs.

Definition 8.2.1 *[245] The Kronecker product of two graphs $\mathcal{G}_1(\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2(\mathcal{V}_2, \mathcal{E}_2)$, denoted by $\mathcal{G}(\mathcal{V}, \mathcal{E}) = \mathcal{G}_1 \otimes \mathcal{G}_2$, is defined as*

$$\mathcal{V} := \{(p,q) | p \in \mathcal{V}_1 \text{ and } q \in \mathcal{V}_2\},$$

$$\mathcal{E} := \{(p_1,q_1) \to (p_2,q_2) | (p_1 \to p_2) \in \mathcal{E}_1 \text{ and } (q_1 \to q_2) \in \mathcal{E}_2\}.$$

Now we construct the so-called enterprise graph. Given the graphs $\mathcal{G}_i(\mathcal{V}_i, \mathcal{E}_i),\ i = 0, \ldots, n$ with vertex sets $\mathcal{V}_i$, $i \in \mathbb{Z}_{0:n}$, and directed edge sets $\mathcal{E}_i$, $i \in \mathbb{Z}_{0:n}$, then the enterprise graph

[2]We stress the fact that the serial connection of the unit operations is not a necessity. In fact, the balance equations lead to a specific structure in matrix A, which is equal to identity matrix I_n, and matrix B summarizes the inlets and outlets of serial, parallel or feedback connections of UOs that are not state dependent.

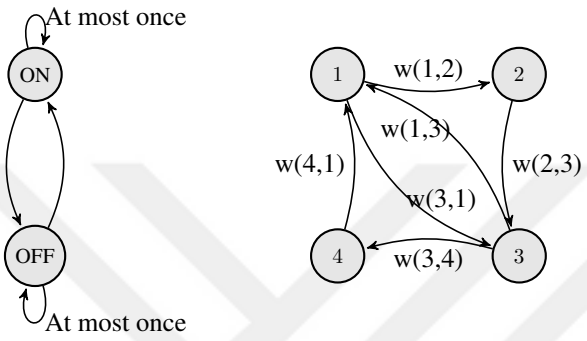


Figure 8.2: Temporal constraints to graph constraints. On the left the possible transitions and on the right the allowed transitions extracted from the a-priori information.

$\mathcal{G}(\mathcal{V}, \mathcal{E})$ is constructed by the Kronecker product of the given graphs, i.e., $\mathcal{G} = \mathcal{G}_0 \otimes \mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_n$. The vertex set of enterprise graph $\mathcal{V}$ is constructed by Cartesian product of all vertex sets, i.e.,

$$\mathcal{V} := \mathcal{V}_0 \times \mathcal{V}_1 \times \cdots \times \mathcal{V}_n,$$

and the edge set is constructed by the Kronecker product of all the adjacency matrices, i.e.,

$$A(\mathcal{G}) := A(\mathcal{G}_0) \otimes A(\mathcal{G}_1) \otimes \cdots \otimes A(\mathcal{G}_n).$$

8.3 Problem Formulation and Theoretical Results for Graph Constrained Scheduling

8.3.1 Research Problems for Graph Constrained Scheduling

In this section we formulate the safe scheduling problem for indefinite operation and steering. As a first step, we pose a verification problem. That is, from some initial points in the state space that lie inside the state constraint set, the verification problem amounts to deciding whether a scheduling sequence exists such that the constraint satisfaction is assured or not. By using the enterprise graph $\mathcal{G}_{Ent}$, which represents all possible On/Off combinations of the process as its nodes, the safe schedule and set can be defined as follows.

Definition 8.3.1 *A **schedule** is a sequence of admissible nodes $\sigma_{[i,j]} = \{\sigma_i, \sigma_{i+1}, \ldots, \sigma_j\}$, $i < j$. A **cycle** is a schedule (a sequence of admissible nodes σ) in which the starting and finishing nodes are same. A **zero-sum cycle** is a cycle in which the sum of the weights of transitions is equal to zero, $\sum_{i=0}^{N-1} b_{\sigma_i,i+1} = 0$, where N is the length of the cycle.*

Next, we define the forward path reachable set, the backward path reachable set and the safe set.

Definition 8.3.2 *For the system dynamics as in Equation* (8.2) *and the scheduling constraints as in Equation* (8.4)*, the one-step forward path reachable set from $\mathcal{S}$ with one step path $\sigma_{1:2}$ is*

$$\mathcal{F}(\mathcal{S}, \sigma_{1:2}) := \{x \in \mathbb{R}^n | \exists y \in \mathcal{S} : x = y + w(\sigma_1, \sigma_2)\}.$$

Likewise, the m-step, $m \in \mathbb{Z}_{\geq 0}$, forward path reachable set from set $\mathcal{S}$ with path (or schedule) $\sigma_{0:m}$ is denoted by $\mathcal{F}^m(\mathcal{S}, \sigma_{0:m})$ and is defined as

$$\mathcal{F}^m(\mathcal{S}, \sigma_{0:m}) = \begin{cases} \mathcal{S}, & m = 0, \\ \mathcal{F}(\mathcal{S}, \sigma_{0:1}), & m = 1, \\ \mathcal{F}^{m-1}(\mathcal{F}(\mathcal{S}, \sigma_{0:1}), \sigma_{1:m}), & m \geq 2. \end{cases}$$

The one-step and m-step backward path reachable sets from $\mathcal{S}$ with path $\sigma_{-m:0}$ are defined as follows.

Definition 8.3.3 *For the system dynamics as in Equation* (8.2)*, the constraints on the state variable $\mathbb{X}$ as in Equation* (8.3) *and the scheduling constraints as in Equation* (8.4)*, one-step and m-step backwards path reachable set are defined, respectively, as,*

$$\begin{aligned} \mathcal{C}(\mathcal{S}, \sigma_{-1:0}) &:= \{x \in \mathbb{X} | \exists y \in \mathcal{S} : y = x + w(\sigma_{-1}, \sigma_0)\}, \\ \mathcal{C}^m(\mathcal{S}, \sigma_{-m:0}) &:= \begin{cases} \mathcal{S}, & m = 0, \\ \mathcal{C}(\mathcal{S}, \sigma_{-1:0}) & m = 1, \\ \mathcal{C}^{m-1}(\mathcal{C}(\mathcal{S}, \sigma_{-1:0}), \sigma_{-m:-1}) & m \geq 2. \end{cases} \end{aligned}$$

Lemma 8.3.1 *The $m-$step forward and backward path reachability sets with paths $\sigma_{0:m}$ and $\sigma_{-m:0}$, respectively, are equal to*

$$\begin{aligned} \mathcal{F}^m(\mathcal{S}, \sigma_{0:m}) &= \mathcal{S} \oplus \{\sum_{i=1}^{m} w(\sigma_{i-1}, \sigma_i)\}, \\ \mathcal{C}^m(\mathcal{S}, \sigma_{-m:0}) &= \mathcal{S} \oplus \{-\sum_{i=1}^{m} w(\sigma_{-m-1+i}, \sigma_{-m+i})\}. \end{aligned} \tag{8.5}$$

Proof: Expand $\mathcal{F}^m(\mathcal{S}, \sigma_{0:m})$ as a vector summation. Re-express the iterated values of $w(\sigma_{i-1}, \sigma_i)$ as a sum which shows the claim. Same deduction is valid for backwards reachability case. □

Definition 8.3.4 *A set $\mathcal{S}$ is called a **safe set** with respect to the process model as in Equation* (8.2)*, the state and scheduling constraints as in Equation* (8.3) *and as in Equation* (8.4)*, respectively, if for any initial condition $x_0 \in \mathcal{S}$ there exists an initial node σ_{-1} and a scheduling sequence $\{b_{\sigma_k}\}_{k \in \mathbb{Z}_{\geq 0}}$ such that*

$$b_{\sigma_k} \in \mathcal{W}_{\sigma_{k-1}}, \; x(x_0, \sigma_{0:k}) \in \mathbb{X}, \quad \forall k \in \mathbb{Z}_{\geq 0}, \tag{8.6}$$

where $x(x_0, \sigma_{0:k}) := x_0 + \sum_{i=0}^{k-1} w(\sigma_i, \sigma_{i+1})$.

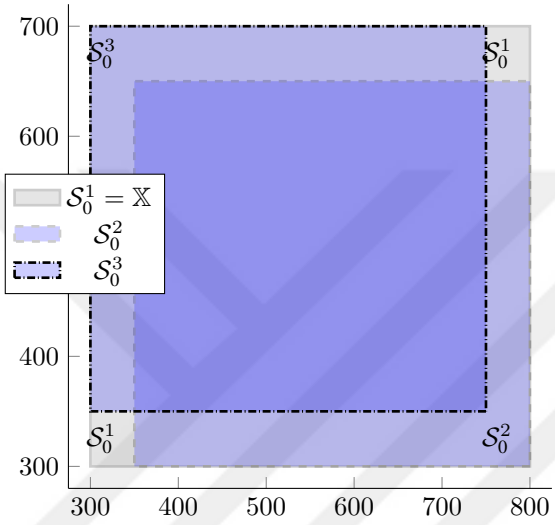


Figure 8.3: Partitioned safe set, from the example in Section 8.4.

We explain the concept of safe set through a visual example in Figure 8.3. We consider three different admissible schedules which are denoted as $\sigma^i_{0:k_i},\ i \in \{1,2,3\}$ which cast the corresponding sets $\mathcal{S}_0^i$ to be safe, respectively. This means that while applying the schedules $\sigma^i_{0:k_i}$, the corresponding set, i.e., $\mathcal{S}_0^i,\ i \in \{1,2,3\}$ remains inside of the constraint set, i.e., $\mathbb{X}$. With the definition of safe set, the research problem of this scheduling problem can be stated as:

Problem 8.3.1 *Consider the system* (8.2)*, the state and the scheduling constraints* (8.3) *and* (8.4)*, respectively. Compute a safe set $\mathcal{S}$ and corresponding admissible scheduling strategy $\sigma_{0:k}$ for each $x_0 \in \mathcal{S}$.*

Verifying the existence of a safe set and computing it assures that the scheduling problem remains to be feasible for the safe set over time iterations of the attached schedule to the safe set.

Once the system model and the safe set are constructed, we can pose the steering scheduling problem as follows.

Problem 8.3.2 *Given a process described as in the form of Equation* (8.2) *with $n \in \mathbb{Z}_{\geq 1}$ UOs that operate on a safe cycle $\hat{\sigma}_{[0,N_{cyc}-1]}$ with state constraints $\mathbb{X}$ and scheduling constraints $\mathcal{G}_i$. Compute an admissible recovery schedule $\sigma_{[0,N_r]}$ for the current state (x_0, σ_0) such that the recovery schedule executes safely, i.e., $x_k(x_0, \sigma_{[0,N_r]}) \in \mathbb{X}$ for $k \in \mathbb{Z}_{[0,N_r]}$, and the operation is connected back to the desired safe cycle ($\sigma_{N_r} \in \hat{\sigma}_{[0,N_{cyc}-1]}$)), while steering the buffer tank levels (states) as desired.*

We visualize the Problem 8.3.2 and an admissible solution in Figure 8.4.

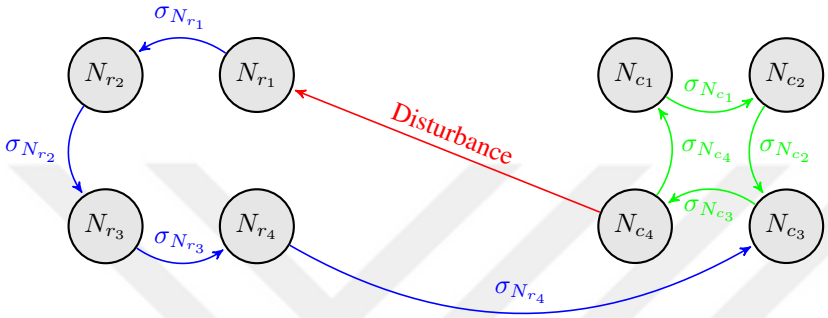


Figure 8.4: The research problem and a solution strategy visualized on $\mathcal{G}_{Ent}$. The green edges denote a safe cycle, the red schedule show the perturbation, and the blue schedule visualizes the recovery schedule.

8.3.2 Theoretical Results on Graph Constrained Schedulling

To solve Problem 8.3.1, we propose a three step approach. First, we construct the cyclic paths over the enterprise graph for which the sum of weights of the edges is equal to zero over the cycle. Second, we find the subsets of the state constraint set that generate the admissible trajectories of the system for the zero-sum cyclic schedules. In the last step we make use of inverse reachability arguments in order to recursively enlarge the safe set. In specific, at each iteration, we identify the sets of nodes and admissible states that reach the safe set in one step. In this way, the cyclic schedules yield periodic trajectories (both in state and graph spaces) for the partitions of constraint set that do not violate over forward trajectories and we expand the safe sets through backward paths. We provide Algorithm 8.3.1 to find and construct the safe set and the corresponding set of schedules that respect the state constraints.

Algorithm 8.3.1 *Input: System as in Equation* (8.2)*, the state constraints* $\mathbb{X}$ *as in Equation* (8.3)*, the scheduling constraints as in Equation* (8.4)*, maximum cycle circumference* N_p*, maximum backwards iterations* N_b*.*

1) Initialization: *Construct the enterprise graph, and generate all the possible paths of length* N_p *from all nodes. We denote the total number of cycles that start from node* σ_0 *and are of maximum length* N_p *with* M^{σ_0} *and the sum of the corresponding weights over each such cycle with* $d_m^{\sigma_0}$*,* $m \in \mathbb{Z}_{1:M^{\sigma_0}}$*. Finally, generate all the admissible backwards paths with length up to and including* N_b*. Denote each such path of length* p *with* $\sigma^q_{-p:0}$ *where* q *is the counter over paths with same length* p *starting from* σ_0*.*

2) Zero-sum Cycles: *For each and every node* $\sigma_0 \in \mathcal{V}$*, solve the integer linear programming problem* ***P.1****, which retrieves the combinations of cycles whose corresponding weights*

sum up to zero:

$$\textbf{P.1:} \quad \min_{\eta^{\sigma_0}} \ ||\eta^{\sigma_0}||_\infty, \\ \text{s.t.} \ \sum_{m=1}^{M^{\sigma_0}} \eta_m^{\sigma_0} d_m^{\sigma_0} = 0, \ \eta^{\sigma_0} \in \mathbb{Z}_{\geq 0}, \sum_{m=1}^{M^{\sigma_0}} \eta_m^{\sigma_0} \geq 1,$$

*where $\eta_m^{\sigma_0}$ is the number of times the cycle $d_m^{\sigma_0}$ is taken in the constructed zero-sum cycle. Denote such cycles as $\sigma_{0:n_j}^j$, where j is the counter over the schedules that are found as the solution of **P.1** with length n_j starting from σ_0. We denote the total number of such zero-sum cycles with N_c.*

3) Safety: *For each of the found schedules $\sigma_{0:n_j}^j$ and for each node $\sigma_0 \in \mathcal{V}$, solve the optimization problem **P.2**,*

$$\textbf{P.2:} \quad \begin{aligned} &max \quad \mathcal{V}(\mathcal{S}_0), \\ &s.t. \quad x_i(x_0, \sigma_{0:i}^j) \in \mathbb{X}, \forall i \in \mathbb{Z}_{0:n_j}, \forall x_0 \in \mathcal{S}_0. \end{aligned}$$

Furthermore, calculate and store the (forward path reachable) sets $\mathcal{S}_0^{j,i} = \mathcal{F}^i(\mathcal{S}_0, \sigma_{0:i}^j)$, for $i = 0, \ldots, n_j$. See Remark 8.3.2 for further discussion on the maximization of $\mathcal{V}(\mathcal{S}_0)$.

4) Enlargement: *Compute the backward reachability sets*

$$\mathcal{S}_0^{j,i,l,q} = \mathcal{C}^l(\mathcal{F}^i(\mathcal{S}_0, \sigma_{0:i}^j), \sigma_{i-l:i}^q),$$

*where $l \in \mathbb{Z}_{0:N_b}$, $i \in \mathbb{Z}_{0:n_j}$, $\sigma_{0:i}^j$ is the i^{th} forward iteration of j^{th} solution of Problem **P.1** and $\sigma_{i-l,i}^q$ is the q^{th} admissible path initiated from σ_i with l backwards reachability iterations.*

Remark 8.3.1 *The order of the cyclic paths that generate zero-sum cycles is important for the construction of different safe sets. Thus, one should check all the permutations of zero-sum cycles and solve Problem **P.2** accordingly.*

Remark 8.3.2 *The maximization of volume of the set $\mathcal{S}_0$ is, in general, a nonconvex problem. However, due to the structure of the system dynamics, i.e., A being equal to identity, not introducing any rotation or dilation to the trajectories and the constraint set $\mathbb{X}$ being hyperrectangle, maximization can be done by finding two vertices of set $\mathcal{S}_0$ that are on the space diagonal. This leads to a significant simplification of Problem **P.2** which is equivalent to solving two linear programming problems for $t = \underline{t}, \overline{t}$, expressed as follows,*

$$\textbf{P.2:} \quad \begin{aligned} \min_{x_t \in \mathbb{R}^n} \quad & ||x_t - \tilde{x}_t||_\infty, \\ s.t. \quad & x_{min} \leq x_t + \sum_{i=1}^{n_j} w(\sigma_i, \sigma_{i+1}) \leq x_{max}, \\ & x(x_t, \sigma_{0:i}^j) \in \mathbb{X}, \forall i \in \{0, 1, \ldots, n_j\}, \end{aligned} \tag{8.7}$$

where $\tilde{x}_{\underline{t}}$ and $\tilde{x}_{\bar{t}}$ are extreme elements of any space diagonal of the constraint set $\mathbb{X}$. The values of x_{min} and x_{max} are obtained from $\mathbb{X}$, which are constructing the faces of the hyperrectangle[3].

Theorem 8.3.1 *Consider the dynamical system as in Equation* (8.1)*, which is subject to constraints as in Equation* (8.3) *and as in Equation* (8.4)*. Suppose that* ***P.1*** *has $N_c \in \mathbb{Z}_{\geq 1}$ different solutions, resulting in $\sigma^j_{0:n_j}$, $j \in \mathbb{Z}_{1:N_c}$ distinct schedules. Then, for any $l \in \mathbb{Z}_{\geq 0}$, the set*

$$\mathcal{S}(N_p, N_b, \mathcal{G}) = \bigcup_{j=1}^{N_c} \bigcup_{i=0}^{N_j-1} \bigcup_{p=0}^{l} \bigcup_{\sigma^q \in \mathcal{I}^p_{\sigma^j_i}} \mathcal{C}^p(\mathcal{F}^i(\mathcal{S}^j_0, \sigma^j_{0:i}), \sigma^q_{i-p:i}),$$

is safe. Here $\mathcal{I}^p_{\sigma^j_i}$ is the set of schedules that are generated for backwards paths initiating from node σ^j_i of length p and indexed with q, i.e.,

$$\begin{aligned} \mathcal{I}^p_{\sigma^j_i} = \quad & \{\{\sigma_r\}_{r\in\mathbb{Z}_{1:p}} \in \mathcal{V}^p : (\sigma_r, \sigma_{r+1}) \in \mathcal{E}, \\ & \forall r \in \mathbb{Z}_{1:p-1}, (\sigma_p, \sigma^j_i) \in \mathcal{E}\}, \end{aligned} \tag{8.9}$$

Proof: For any triplet

$$(x_0, i, j) \in (\mathcal{F}^i(\mathcal{S}^j_0, \sigma^j_{0:i}) \times \mathbb{Z}_{0:n_j} \times \mathbb{Z}_{1:N_c}),$$

by solution of **P.1**, there exists a scheduling sequence $\sigma^j_{i:n_j}$ such that $\bar{x} := x(x_0, \sigma^j_{i:n_j}) \in \mathcal{S}^j_0$. Consider the periodic scheduling sequence $\hat{\sigma}$ generated by repeating $\sigma^j_{0:n_j}$, i.e., $\hat{\sigma} = \{\sigma^j_{0:n_j},\ \sigma^j_{0:n_j}, \ldots\}$. Then, from **P.2**, it follows that

$$x(\bar{x}, \hat{\sigma}) \in \bigcup_{i=0}^{n_j} \mathcal{F}^i(\mathcal{S}^j_0, \sigma^j_{0:i}) \subseteq \mathbb{X}.$$

Finally, for all triplets

$$(x_0, p, q) \in (\mathcal{C}^p(\mathcal{F}^i(\mathcal{S}^j_0, \sigma^j_{0:i}), \sigma^q_{i-p:i}) \times \mathbb{Z}_{0:l} \times \mathcal{I}_{\sigma^p_{i-p:i}}),$$

by definition of the backward path reachability mapping and its iterations, there exists a schedule $\sigma^\star$ of length p such that

$$x(x_0, \sigma^\star) \in \mathcal{F}^i(\mathcal{S}^j_0, \sigma^j_{0:i}).$$

This concludes the proof. □

[3]The validity of this argument can be seen through the volume of the constructed safe set, i.e.

$$\begin{aligned} \max \mathcal{V}(\mathcal{S}_0) = \quad & \max_{x_{\underline{t}}, x_{\bar{t}}} \prod_{i=1}^{n} |[x_{\bar{t}}]^i - [x_{\underline{t}}]^i|, \\ = \quad & \min_{v_1, v_2} \prod_{i=1}^{n} |[\tilde{x}_{\bar{t}} + v_1]^i - [\tilde{x}_{\underline{t}} + v_2]^i|, \end{aligned} \tag{8.8}$$

where $[x]^i$ is the i^{th} component of vector x, v_t denote the vectors from $\tilde{x}_t$ to x_t. Lastly, the set $\mathcal{S}_0$ has the points $x_{\underline{t}}$ and $x_{\bar{t}}$ as the extreme elements of its space diagonal.

Remark 8.3.3 *For all $l_1 \leq l_2$, the constructed safe sets $\mathcal{S}^{j,i,l_1} \subseteq \mathcal{S}^{j,i,l_2}$. Thus the sequence of safe sets is ordered and monotonically non-shrinking.*

Next, we propose an solution technique to Problem 8.3.2 in Algorithm 8.3.2, which uses the reachability operations for establishing safe steering while using admissible schedules.

Algorithm 8.3.2

Admissible Recovery Schedules: *Construct all admissible schedules that can reach to the target cycle within $\bar{N}_r$ steps, i.e.,*

$$\sigma^q_{[0,p-1]} \in \Delta_p := \left\{\sigma_{[0,p-1]} | \sigma_{[i,i+1]} \in \mathcal{E}_{Ent}, i \in \mathbb{Z}_{[0,p-2]}, \sigma_{p-1} = \hat{\sigma}_j, j \in \mathbb{Z}_{[0,N_{cyc}-1]}\right\},$$

where Δ_p is the set of admissible schedules of size p that reach to safe cycle $\hat{\sigma}_{[0,N_{cyc}-1]}$, q is the index over the schedules of length p and $\mathcal{E}_{Ent}$ is the edge set of enterprise graph.

Safe Steering: *Solve the feasibility problem **P1** for the admissible schedules $\sigma_{[0,p]}$, where $p = 1, \ldots, \bar{N}_r$ and $\mathbb{X}^*$ is a desired subset of state space inside constraints,*

$$\textbf{P1:} \left\{ \min_{\sigma^q_{[0,p-1]}} \quad p, \quad \textit{s.t.} \quad x_j(x_0, \sigma^q_{[0,j-1]}) \in \mathbb{X}, \;\; x_p(x_0, \sigma^q_{[0,p-1]}) = \mathbb{X}^*, \;\; j = 1, \ldots, p, \right.$$

Remark 8.3.4 *Due to the problem structure, one can speed up the calculations by;*

- *searching paths over the (powers of) adjacency matrix of the enterprise graph to find a connection of the starting node and a node on the desired safe cycle. If there is an n_r step path (admissible schedule) from σ_0 to $\hat{\sigma}_j \in \mathcal{V}_{Ent}$ then $\mathcal{A}^{n_r}_{Ent}(\hat{\sigma}_j, \sigma_0) = 1$;*
- *using the Equation* (8.5) *to check the set membership constraint, i.e.,*

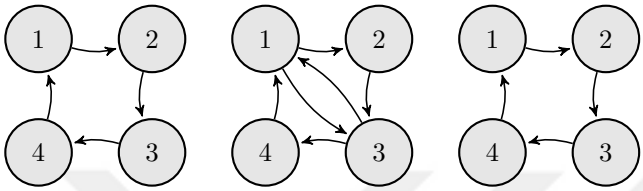
$$\begin{aligned} x(x_0, \sigma_{[0,j]}) \in \mathbb{X}, \implies \quad & x_{min} \leq x_0 + \sum_{i=0}^{j-1} b_{\sigma_{i,i+1}} \leq x_{max} \\ \implies \; & x_{min} - x_0 \leq \sum_{i=0}^{j-1} b_{\sigma_{i,i+1}} \leq x_{max} - x_0. \end{aligned}$$

Before providing simulation results, we discuss the reduction in the computational complexity of the scheduling problem. By employing the graph based modeling framework, one can drastically decrease the number of decision variables. Indeed, as an example, consider a case with 3 decision variables. In the conventional modeling case, one needs to optimize over a decision tree of $(2^3)^k$ different schedules, where k is the prediction horizon. However using the graphs, the number of possible schedules are reduced to almost 11000 for a prediction horizon of $k = 10$ instead of $2^{30} \approx 10^9$.

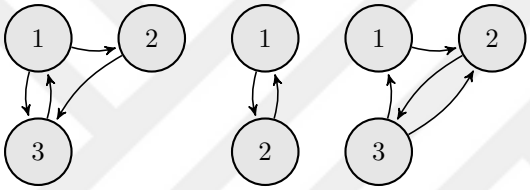
8.4 Simulation Study on Scheduling of Whey Protein Seperation Unit Operations

8.4.1 Safe Cyclic Operation

We consider the scheduling problem of cleaning times of a simplified separation plant, see [146] for details. The separation process has 2 UOs, which induces 3 binary decision variables α^j and 2 continuous states denoted as x_i which stands for the mass hold ups in the



(a) The graphs representing temporal constraints.



(b) The graphs of UOs that have different on-off structure.

Figure 8.5: Different On-Off structures of UOs for the case study.

tanks between the UOs. The model of the process is constructed as

$$\begin{aligned} x^1_{k+1} =& x^1_k + \phi^0 \alpha^0_k - \phi^1 \alpha^1_k, \\ x^2_{k+1} =& x^2_k + \bar{\phi}^1 \alpha^1_k - \phi^2 \alpha^2_k, \end{aligned} \tag{8.10}$$

where $x_k = \begin{bmatrix} x^1_k & x^2_k \end{bmatrix}^\top$. The parameter values are equal to each other which is set as $\phi = 50$. The constraint set is defined as $x_k \in \mathbb{X} = [300,\ 800] \times [300,\ 700]$, for all time instants $k \in \mathbb{Z}_{\geq 0}$. We consider two different cases for the graphs that stand for the temporal constraints. For the first case, the decision variables $\alpha^j,\ j \in \{0, 1, 2\}$, are constrained to 3 directed graphs, all of them with 4 nodes, see the Figure 8.5a.

For this example, we construct the enterprise graph which consists of 64 vertices. We name these nodes from the numbers on graphs in base 4, i.e.,

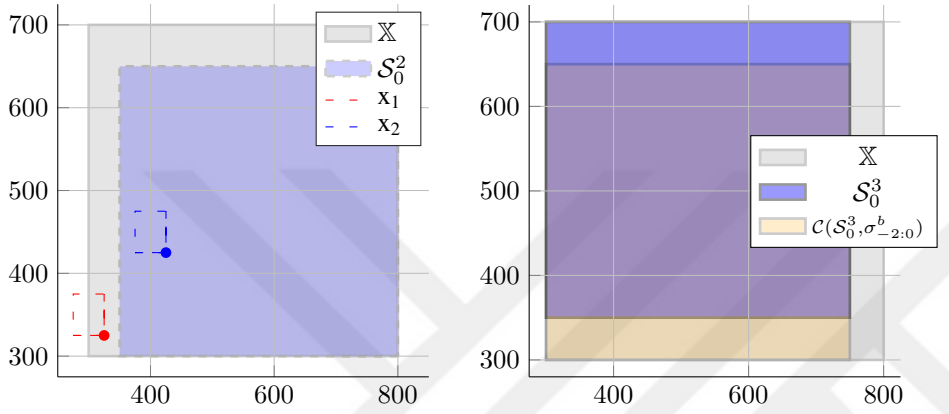
$$\sigma := (a - 1, b - 1, c - 1)_{4 \to 10} + 1,$$

where a, b, c corresponding to the numbers on the individual UO graphs and $(\cdot)_{4 \to 10}$ denotes the conversion of the considered number from base 4 to base 10. Then we apply Algorithm 8.3.1 to construct the safe set $\mathcal{S}$. A horizon of 8 prediction steps yields 121 different cycles in the enterprise graph. In Figure 8.3, we visualize the partitions of safe set with respect to three different schedules that can be generated from the graphs in Figure 8.5a. These (cyclic) schedules are;

$$\sigma^1_{0:4} = \begin{bmatrix} 1 & 22 & 43 & 64 \end{bmatrix},$$

$$\sigma^2_{0:7} = \begin{bmatrix} 1 & 26 & 35 & 60 & 26 & 35 & 60 \end{bmatrix},$$

$$\sigma^3_{0:8} = \begin{bmatrix} 20 & 41 & 50 & 11 & 20 & 37 & 58 & 15 \end{bmatrix},$$



(a) Safe set and trajectories.

(b) The constructed safe set.

Figure 8.6: Safe set from forward and backward iterations.

where all these schedules satisfy the zero-sum condition, which introduce different safe sets visualized in Figure 8.3. To demonstrate the trajectories in the state space, we select two different initial points for schedule $\sigma^2_{0:7}$. The first initial point is selected outside of the safe set $\mathcal{S}_0^2$, thus violating the constraints, whereas the second trajectory stays inside $\mathbb{X}$ as visualized in Figure 8.6a. Lastly, we apply backwards path reachability mapping for two steps for safe set $\mathcal{S}_0^3$ with backwards path selected as $\sigma^b_{-2:0} = \begin{bmatrix}50 & 11 & 20\end{bmatrix}$. The obtained set $\mathcal{S}^{20,0,2}$ is visualized in Figure 8.6b. In the second example we consider a different switching rule as presented in Figure 8.5b. Similar to previous example we take the maximal cycle length as 8. We start by constructing the enterprise graph, by using Kronecker product over graphs which is visualized in Figure 8.7. Then we have identified 198 cycles, including the permutations of cycles. However, due to the graphs corresponding to the inlet and UO2, there does not exist cycles longer than 2 steps which sum up to the zero vector. Thus the only cycles that yield infinitely long feasible schedules are:

$$\sigma^1_{0:1} = \begin{bmatrix}2 & 18\end{bmatrix}, \quad \sigma^3_{0:1} = \begin{bmatrix}3 & 17\end{bmatrix},$$
$$\sigma^5_{0:1} = \begin{bmatrix}5 & 15\end{bmatrix}, \quad \sigma^7_{0:1} = \begin{bmatrix}6 & 14\end{bmatrix},$$

and their permuted counterparts. All of these schedules have the zero-sum property over the cycle. Then, we proceed with the calculation of the safe set for each of the schedules in Equation (8.11). The safe sets generated for these schedules are equal to the constraint set $\mathbb{X}$. Thus, the safe set is equal to the constraint set and no further enlargement is possible via the backward reachability procedure.

8.4.2 Safe Reactive Scheduling after Disruptions

For steering purposes, we consider the same number of UOs for describing the whey process. For each UO same graph constraints are assumed as in Figure 8.2, leading to an enterprise graph consisting of 125 vertices. The constraints on the two buffer tanks are

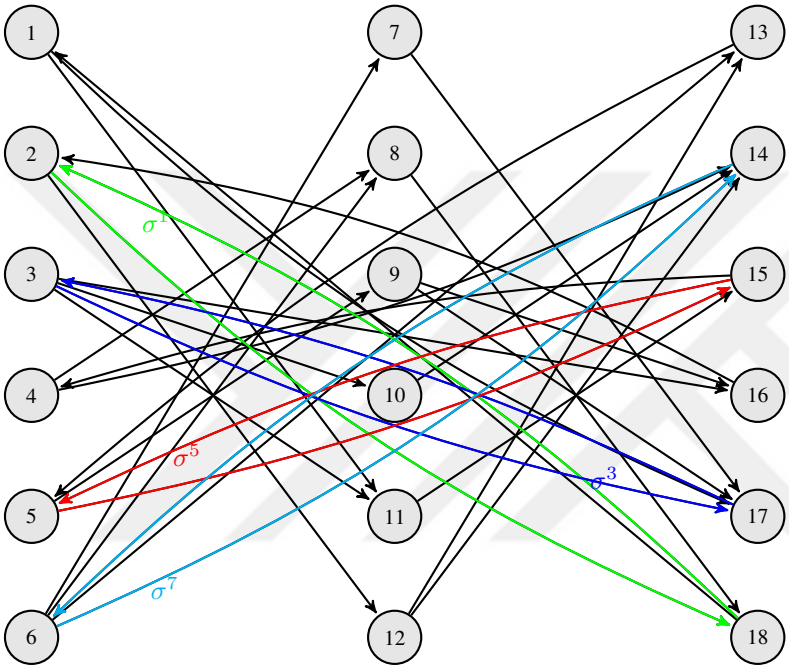


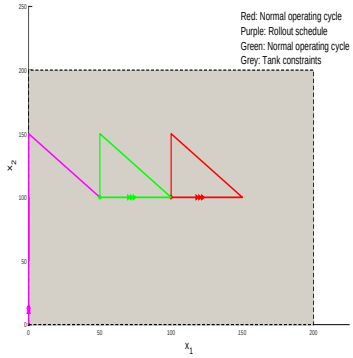
Figure 8.7: The weighted digraph for the graphs in Figure 8.5b.

given as $0 \leq x_k^i \leq 200$, $i = 1, 2$, $k \in \mathbb{Z}_{\geq 0}$. The values of the flow parameters are equal to each other and set to $\phi = 50$.

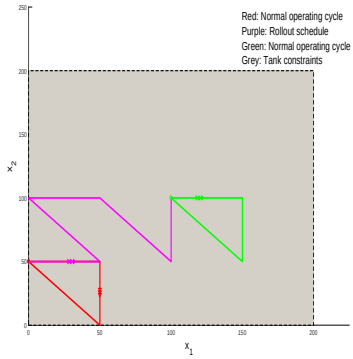
We consider two different simulation studies; i) the recovery case where both the state x_k and the mode σ_k are perturbed, ii) the steering case, no disturbances occur but the mode leaves the safe cycle to steer the process as desired. The operating zero sum cycle is selected as $\hat{\sigma}_{\{56,87,18,29\}}$.

Then a recovery schedule is generated towards the zero-sum cycle $\hat{\sigma}_{\{56,87,18,29\}}$ after applying different perturbations on the state and the mode. The results are shown graphically in Figures 8.8a-8.8b. On Figure 8.8a the first case study is visualized, the operation on the safe cycle, i.e., the red trajectory, is perturbed which transforms the states to origin and sets $\sigma_0 = 1$. We apply the safe recovery algorithm to steer the system back to the safe cycle and acceptable states through the recovery schedule $\sigma_{[0,N_r]} = \{1, 82, 113, 19, 76, 107, 18\}$, i.e., the pink trajectory. In the second case study the safe steering aspect is demonstrated. We initialize the process in a safe cycle (red trajectory), where the buffer tank values are initialized close to the constraints. Then by applying Algorithm 8.3.2, we find a safe schedule, that steers the buffer tank states to desired locations, i.e., pink trajectory, while returning back to the zero-sum cycle at the end of the recovery, i.e., the green periodic trajectory.

Results indicate that it is much harder for the buffer tanks to be filled in comparison to the case of draining, since there are more On nodes than Off nodes, thus more safe schedules to fill the tanks. As discussed above, the computational complexity is decreased since the recovery schedules are found by matrix multiplications and interval arithmetic.



(a) The state trajectories for case study 1.



(b) The state trajectories for case study 2.

Figure 8.8: Buffer hold up trajectories in two different steering cases.

8.5 Conclusions on Graph Constrained Scheduling for Whey Separation Processes

In this chapter, we have considered a type of scheduling problems which are difficult to solve directly due to the integer nature of the decision variables. In many scheduling problems there exist both (i) logical rules for the decision variables (constraints on schedules) over time; (ii) a highly structured system dynamics. We include both properties by introducing directed graphs that represent the temporal constraints on scheduling variables and by considering the process as a discrete-time LTI system. While the exposition is performed only for subprocesses connected in series, the results are applicable to any other interconnection scheme. We implemented a simulation example from an industrial process that is in line with the theoretical outcomes.

- One can extract all the admissible and safe schedules, with a significant reduction in

the number of schedules by using graphs as constraints on the evolution of enterprise.

- By computing a safe set, we guarantee existence of at least one admissible schedule for all initial conditions in the safe set.
- By using structure of the mass balance equations, the optimization problems boils down to interval arithmetic operations.

Chapter 9

Epilogue

> Being a realist does not mean seeing the truth for what it is. It is a question of determining our relationship with the truth in the way that is most beneficial for us.
>
> Ahmet Hamdi Tanpınar - The Time Regulation Institute

This chapter presents the concluding remarks on the research carried out in the topics of risk-aware model predictive control (MPC) and online model-based application for ultrafiltration (UF) membrane units and discusses future research directions on these topics.

Online model-based applications and their performance depend on the quality of the model. It is inevitable that uncertainties and unmodelled dynamical phenomena will be present in any model in comparison to the true trajectories. Therefore, the first part of this thesis has been directed towards incorporating these effects into the model predictive control technology. We base our discussion on an extensive classification of the robust MPC methods used in the literature according to the typical ways of treating the uncertain effects. In the course of achieving this objective, we have noticed that there is a fundamental conflict between the robust versus high performance operation. In short, the practitioner trades off the performance versus the pessimistic operation necessary for robustness. We show that weighting, through adjustable parameters, the predicted uncertainty into control actions with statistical (moment-based) reasoning allows us to design robustly operating closed-loop systems with desirable computational properties. The resulting robust MPC problem is constrained to the nominal model which is a beneficial aspect for using large-scale rigorous models in MPC problems. Furthermore, control theoretic properties, such as disturbance sensitivity, are improved in a systematic and intuitive way. Specifically, the Chapters 2-3-4-5 discuss the effect of various type of uncertainties in (un)constrained MPC problems and present the moment-based MPC techniques.

In the second part of this dissertation, we have directed our attention to (online) model-based control strategies applied to a whey protein separation process. Within this process, the UF membrane units lead to separation of whey proteins from the other components such

as sugars or salts. Substantial effort is directed towards the steps in rigorous model development. Furthermore this model is used to demonstrate the extended capabilities that can be achieved by offline model-based activities, such as optimal operation and sensor selection strategies. In the online stage, we use this model in monitoring (estimation) or control applications to improve the efficiency of the separation. Another research direction that is covered in this thesis regarding the process is on the scheduling of the subprocesses within the whole plant. We have used a novel modeling approach that is based on directed graphs to mimic the plant operation and through these graphs we deduce safe operating schedules. Specifically, Chapters 6-7-8 discuss offline or online use of model-based strategies towards a whey protein separation process.

In the next two sections, we detail our observations to address the research goal and questions identified in Chapter 1 and provide future research suggestions for the research themes respectively.

9.1 Conclusions and Future Directions for Risk-aware MPC

9.1.1 Conclusions on the Risk-aware MPC

In Chapters 2, 3, 4 and 5, we have considered a novel risk-aware MPC problem formulation to overcome high computational load and pessimistic closed-loop operation. These disadvantages are commonly observed in many of the robust MPC problem strategies as identified by the 1st Research Question in Chapter 1. In Chapter 2, we reflect on the robust MPC based closed-loop operation and identify that the practitioner decides on several important design questions that are directly affecting the resulting closed-loop operation:

- **"Does the nominal model generate acceptable predictions?":** The nominal prediction model radically affects the MPC actions through the future predictions. By using rigorous models, one expects to use (complex) nominal models with relatively insignificant uncertainty descriptions. This is expected to lead to an increase in the performance since the resulting control actions reduces the pessimism due to uncertain effects. There is no clear guideline to trade off the increase in the nominal model complexity against the uncertainty model that the practitioner is incorporating into the predictions in terms of the controlled system's performance.

- **"Do the modeled uncertainties mimic realistic scenarios?":** The performance loss due to robustness depends on how the uncertainty is evolving in the dynamics and whether the practitioner is evaluating the uncertain trajectories as risky or not. Current deterministic robust MPC methods lead to highly pessimistic operation, since all uncertain trajectories are deemed to be equally important. The stochastic robust MPC methods are based on the distribution of the predictions, hence some trajectories with low realization probability are not incorporated into the MPC problem. This leads to an improvement in the operation. However the optimization methods that operate on distributions are numerically tedious and always prone to wrong pdf models, especially evaluating the tails of distributions. Predictions based on scenarios is a key, and recent, approach, since one can effectively incorporate only the realistic uncertainty scenarios. Thus, the closed-loop operation is expected to bypass the unnecessary conservatism. Yet, the number of scenarios in MPC problems for operation with guaranteed properties is still too large for rigorous large-scale

models to handle.

The observations in Chapter 2 motivate the control theoretic analysis of moment-based MPC formulations. The moment-based MPC algorithms are based on the statistics of the cost and constraint functions for uncertain dynamical systems. To improve the response, we statistically re-evaluate the allocated uncertainty budget for each constraint and cost function to adjust the robustness levels[1]. We have evaluated the mean, mean-variance and higher order statistics of MPC cost and constraint functions for various control (regulation/tracking or state/output) and uncertainty configurations. By doing so, we have shown that:

- The moment-based MPC problems can be expressed in terms of nominal dynamics with stage-wise varying cost functions over the prediction horizon. This decreases the computational complexity drastically in comparison to the other robust MPC techniques since moment-based MPC does not use the pdfs or reachable intervals explicitly. Hence for processes with medium to large-scale representation of their dynamics, the moment-based MPC formulations remain computationally tractable.
- The ability to actively adjust the effect of overall spread of predictions also leads to adjustable performance versus robustness specifications. This allows the practitioners to implement nominal MPC schemes for process control systems with robustness guarantees.
- Explicit formulae exist for the reformulation of the weighting terms in MPC problems. These weights are depending on how the uncertainty is incorporated into the prediction model, their evolution and their statistical properties.

These observations provide a solid reasoning for research activities in moment-based MPC topic in the first research theme. Next, we have reformulated the robust counterpart MPC problems for systems under the effect of;

1. the additive perturbations, in which the uncertainty is accumulating in the future predictions. Thus, the optimization weights are also affected in an accumulative manner as the prediction stage increases.
2. the initial condition mismatches, in which the uncertain effect is only due to the lack of information at the current time instant. There is no accumulation of uncertain effects since there is no other source of uncertainty.
3. the plant-prediction model mismatch (or the multiplicative uncertainty), in which the predictions become more and more affected by the control decisions and states in the earlier prediction steps, which results in an increase in the weights of initial prediction stages in comparison to later ones.

These effects span a large section of uncertainty descriptions that are used in control literature. The effect of having a detailed uncertainty model on the closed-loop operation and performance, the second subquestion on the moment-based MPC topic has also important consequences:

[1] We remind that the stochastic MPC approaches are only valid for the cases where the constraint violations are practically allowable but undesired.

- For the additive perturbations and initial condition mismatch cases, we explicitly provide the effect of high-order statistics and hence the added value of detailed modeling of uncertain effects. In these cases, the MPC routine is suppressing the uncertainties by increasing the state/output weights in the optimization problems, as a function of the high order moments of the uncertainties.

- The high order moment-based MPC problems for the plant-model mismatch case are not completely formulated in this thesis, due to the associated increase in the MPC problems' complexity. However, due to the dependence of state trajectories to each other over time, cross-correlation (or higher order moment) terms are occurring in the robust counterpart formulation, which incorporates the uncertainty.

- We generalize the moment-based MPC towards non-Gaussian disturbances and report the mean, mean-variance and MVS MPC problems for these cases. We show the changes in the closed-loop operation by actively modeling the pdf, and hence demonstrate its effect on the control decisions.

The third research question of the first theme, addressing the cost functions and the selection of optimization parameters' values for incorporating the uncertain effects, have been addressed in the technical results given for the uncertainty types listed above. In each of these uncertainty families, we have established equivalence results on the optimization parameters of an MPC problem constrained to the nominal dynamics to let the closed-loop operation become aware of the uncertain effects predicted in the dynamics. Furthermore the effect of the robust counterpart formulations on the bandwidth or other closed-loop performance aspects is similar to the traditional robust control methods which are based on frequency domain performance filters designed in top of the nominal model.

As the fourth research question on this research area, we have formulated the constraint handling problem for the moment-based MPC formulations. We use different tools to achieve constraint satisfaction, such as distance concepts using distributions, overbounding the spread through variance operator or approximations of the predicted spread. Using the Mahalanobis distance and the nominal system dynamics, we provide the robust counterparts of constraints in moment-based MPC problems with additive perturbations. We observe that for the most frequently used constraint functions, i.e., affine functions of states and inputs, the variance based constraint handling technique increases the complexity class of the problem. Some conservative reformulations of constraints are reported which reduce the computational complexity of the robust counterparts of the constraints obtained after the applications of centralized moments.

9.1.2 Future Directions on the Risk-aware MPC

Our suggestions on future research towards a reliable operation with MPC based technology is presented next.

- First, we point out the process control hierarchy, Figure 1.2a, and a research direction along the line of decentralized control problems. The multi-layer strategies in process control hierarchy are resulting in separate optimization problems for each layer, generally due to different time scales, dynamics, uncertain effects and cost functions. The communication

of data and decisions in between these layers have a substantial effect on the resulting performance[2]. For this purpose;

1. Recourse based algorithms which incorporate lower or upper layers in their own formulation is expected to improve the operating performance.
2. The event (or self) triggering based update schemes can be used to achieve equivalent performance with less computational resources.

Thus, a rigorous approach on how to construct the hierarchy and the flow of information is still missing in the multi-step, multi-agent dynamical decision making research area.

- Secondly, it is evident that the MPC cost and constraint functions should be tailored to incorporate deviation and integral terms[3] of states or inputs based on the risk-aware predictions. Using previously calculated optimal actions and predictions in the current MPC problem to weight out the unpredicted mismatches is expected to provide the desired robustness properties without explicitly modeling or handling the uncertainties. In this way the MPC problem only depends on the nominal model of the process and the previous optimal solutions are expected to result in computational benefits, such as warm-starts.

- One positive aspect of the moment-based MPC formulations is the computational properties of the resulting problems. Using higher order moments is desirable since the true distribution of states can be approximated better with higher order moments, in the case of infinitely many moments, one can deduce deterministic (worst-case or others) properties for closed-loop operation. However the high order moment-based MPC problem becomes intractable. Yet, formulating approximation algorithms for the higher order moments of costs and constraints by polynomial optimization methods is expected to end up with functions that can be optimized over in a tractable way. This approach is also useful for the cases where the dynamics are described by nonlinear equations.

9.2 Conclusions and Future Directions for Model-based Operation of Whey Protein Separation and Ultrafiltration Membrane Processes

9.2.1 Conclusions on Model-based Operation of Whey Protein Separation and UF Membrane Processes

In Chapters 6, 7 and 8 we have considered a whey protein separation process and discussed the benefits of model-based applications for this process. Specifically, we have presented a rigorous large-scale model and facilitate it in offline and online optimization studies.

The first research question on this theme is on the modeling of ultrafiltration membrane units. A mathematical model consisting of differential and algebraic relations among the physics based variables describing the process through balance laws for ultrafiltration membranes is developed. We have estimated the parameters of three proposed models by measurements gathered from an industrial plant and we validate these models via statistical techniques. The statistical analysis allows us to select one of these models.

[2]The effect of solving one but larger problem by ignoring the layered structure is demonstrated in this thesis for scheduling and offline optimization of UF membrane operation.

[3]In the classical robust control theory, low and high pass filtering of the performance or input variables.

For the second research question, modelling and tracking the fouling, which is the major physical phenomena leading to performance loss in membrane based processes, we have incorporated via the membrane resistance concept. Hence monitoring and if possible controlling it during the operation can be achieved via the developed model. The dynamic membrane resistance is always accumulating and the empirical membrane resistance reflects the dependency on the operating conditions.

With the constructed model we implement several simulation studies in order to deduce useful operating strategies for the UF membrane processes, for the Research Question 2.2. The simulation studies on the optimal operation and scheduling strategies highlight the potential use and benefits of the developed model. We stress the conclusions in various optimization based studies such as;

- For extended (batch) operation of UF membranes satisfying desired goals, the total permeate flow needs to be distributed to all of the membranes to reduce the fouling accumulation. This can be achieved by operating all membranes in the subprocess from the start of the batch.

- The model can be used for monitoring and for selecting what to measure for the UF membranes. The spectral analysis of the observability gramian is used to deduce comparative results between different sensor configurations, yet it is often not enough to consider only one of these indices. We observe that the conclusions drawn from the data based (empirical) observability gramian correlate with the conclusions drawn by the experienced operators of the process.

In Chapter 7, we present the results of two crucial online model-based applications for UF membrane processes, the soft-sensor and the low-level control implementations addressing the Research Question 2.3;

- The first simulation study presents two soft sensing methods, Kalman filter and moving horizon estimation scheme, and their implementations on the UF membrane units. The online monitoring of the fouling in UF membranes helps with other online model-based activities such as improving the operator decisions, preventing disruptions and synthesis of model-based controllers. All of the constructed estimators produce acceptable and comparable results in terms of tracking the fouling effects even after some hybrid changes (switching on a new UF membrane, large changes in the operating conditions) observed. However, monitoring the key performance variables yield wrong estimates, possibly due to the mismatch in the operating conditions based on algebraic regression techniques.

- The controller implementation studies towards UF membrane process demonstrate that for desirable membrane operation both of the inlet variables, the mass flow and the operating transmembrane pressure, should be manipulated together. We have designed and tested four different controller structures using the UF membrane model: Two PID based control structures; a combined Iterative Learning Model Predictive Controller; and an extension of the latter with an Extended Kalman Filter. Both learning controllers and the double PID controller manipulate the applied pressure, which leads to smoother mass flow trajectories. Here, the IL-MPC controllers are capable of predicting the future outputs based on the past batch trajectories and learning from the errors provoked in the previous batches.

The final research topic discussed in this thesis addresses the safe and reliable scheduling of various subprocesses that construct the whey processing plant. In this separation process, many subprocesses operate in batches asynchronous to each other, thus scheduling problems are inherent to the operation. For this reason;

- We construct a type of scheduling problems that make use of directed graphs to model the subprocesses and temporal constraints (the operational or offline time windows) on them. This allows us to model the plant operation as a discrete-time LTI system with integer control actions.
- We use the graphs related to unit operations to extract all admissible schedules which is resulting in a significant reduction in the number of schedules to be analyzed.
- We compute safe sets and associated cyclic schedules, which guarantee the existence of at least one admissible schedule to continue the operation indefinitely. Furthermore, under this setting the optimization problems boil down to interval arithmetic operations.

9.2.2 Future Directions on the Model-based Operation for Whey Protein Separation and UF Membrane Processes

We list several recommendations on future research activities relevant to UF membranes in below:

- Regarding the proposed UF membrane model, the parameters in empirical membrane resistance terms are highly effective in the predicted flows. During start up procedures, the membrane process starts with high permeate flux conditions (due to thin whey) but the process quickly moves towards high fouling (due to thick whey) conditions. In the data used for the current parameter estimation results, we did not have access to sufficiently detailed/sampled data for the start up period, hence these procedures should be repeated once these data are available.
- The empirical observability and controlability gramians offer incredible computational benefits for approximating the true ones only from simulation data for nonlinear or unstable systems. These gramians can be used instead of the conventional ones in open loop analysis. We stress that maximal membrane fouling can be decreased by tuning the TMP's per membrane loop separately. This ends up with either increased operating pressure or longer batch time, which could lead to a higher production capacity while providing desired performance specifications. To achieve this, extracting more permeate mass flux at the initial membranes, in which thin whey (small protein concentration) circulates, is beneficial since thicker whey causes the later membranes to foul faster due to high protein concentration.
- We suggest expanding the monitoring routines by incorporating parameter and/or disturbance estimation aspects. The fouling parameters or the composition of inlet whey are affecting the performance of the operation, thus monitoring these parameters is an important step towards better operation. Similarly, we propose that incorporating learning action (across batches) into Advanced Process Control layer is crucial for high efficiency operation. Rigorous comparison between automatic model update mechanisms and incorporation of learning action into the controllers is a topic for further study.

Bibliography

[1] V. Adetola and M. Guay. Robust adaptive mpc for constrained uncertain nonlinear systems. *International Journal of Adaptive Control and Signal Processing*, 25(2):155–167, 2011.

[2] A. Alessio and A. Bemporad. A survey on explicit model predictive control. In *Nonlinear model predictive control*, pages 345–369. Springer, 2009.

[3] A. Alessio, M. Lazar, A. Bemporad, and W.P.M.H. Heemels. Squaring the circle: An algorithm for generating polyhedral invariant sets from ellipsoidal ones. *Automatica*, 43(12):2096–2103, 2007.

[4] R.J. Allgor and P.I. Barton. Mixed-integer dynamic optimization i: problem formulation. *Computers & Chemical Engineering*, 23(4):567–584, 1999.

[5] F. Allgower and A. Zheng. *Nonlinear model predictive control*. Birkhauser, 2000.

[6] J.C. Allwright and G.C. Papavasiliou. On linear programming and robust model-predictive control using impulse-responses. *Systems & Control Letters*, 18(2):159–164, 1992.

[7] M. Annergren, D. Kauven, C.A. Larsson, M. Potters, Q. Tran, and L. Özkan. On the way to autonomous model predictive control: A distillation column simulation study. In *Proceedings of the 10th International Symposium on Dynamics and Control of Process Systems*, pages 713–720, 2013.

[8] H. Arellano-Garcia. *Chance Constrained Optimization of Process Systems under Uncertainty*. PhD thesis, Technische Universität Berlin, 2006.

[9] H. Arellano-Garcia and G. Wozny. Chance constrained optimization of process systems under uncertainty: I. strict monotonicity. *Computers & Chemical Engineering*, 33(10):1568–1583, 2009.

[10] P. Artzner, F. Delbaen, J-M. Eber, and D. Heath. Coherent measures of risk. *Risk management: value at risk and beyond*, page 145, 2002.

[11] K.J. Aström and P.R. Kumar. Control: A perspective. *Automatica*, 50(1):3 – 43, 2014.

[12] N. Athanasopoulos and M. Lazar. Stability analysis of switched linear systems defined by graphs. In *Proceedings of the IEEE 53rd Conference on Decision and Control (CDC)*, pages 5451–5456, Dec 2014.

[13] Richard W Baker. *Membrane technology and applications*. John Wiley & Sons, Ltd, 2004.

[14] V. Bansal, V. Sakizlis, R. Ross, J.D. Perkins, and E.N. Pistikopoulos. New algorithms for mixed-integer dynamic optimization. *Computers & Chemical Engineering*, 27(5):647–668, 2003.

[15] M. Barger and R. P. Carnahan. Fouling prediction in reverse osmosis processes. *Desalination*, 83(1):3–33, 1991.

[16] Andrew R Barron. Entropy and the central limit theorem. *The Annals of probability*, pages 336–342, 1986.

[17] T. Barz, G. Wozny, and H. Arellano-Garcia. Robust implementation of optimal decisions using a two-layer chance-constrained approach. *Industrial & Engineering Chemistry Research*, 50(9):5050–5063, 2010.

[18] I. Batina. *Model predictive control for stochastic systems by randomized algorithms*. PhD thesis, Technische Universiteit Eindhoven, 2004.

[19] I. Batina, A. Stoorvogel, and S. Weiland. Stochastic disturbance rejection in model predictive control by randomized algorithms. In *Proceedings of the 2001 American Control Conference (ACC)*, volume 2, pages 732–737. IEEE, 2001.

[20] I. Batina, A. Stoorvogel, and S. Weiland. Optimal control of linear, stochastic systems with state and input constraints. In *Proceedings of the IEEE 41st Conference on Decision and Control (CDC)*, volume 2, pages 1564–1569. IEEE, 2002.

[21] M. Bauer and I.K. Craig. Economic assessment of advanced process control–a survey and framework. *Journal of process control*, 18(1):2–18, 2008.

[22] M.L. Bell, D.J.N. Limebeer, and R.W.H. Sargent. Robust receding horizon optimal control. *Computers & Chemical Engineering*, 20:S781–S786, 1996.

[23] R. Bellman. Dynamic programming and lagrange multipliers. *Proceedings of the National Academy of Sciences*, 42(10):767–769, 1956.

[24] R. Bellman. Dynamic programming and stochastic control processes. *Information and Control*, 1(3):228–239, 1958.

[25] R. Bellman and K.J. Aström. On structural identifiability. *Mathematical biosciences*, 7(3):329–339, 1970.

[26] R.E. Bellman. *Adaptive Control Processes: A Guided Tour*, volume 4. Princeton university press, 1961.

[27] A. Bemporad, L. Bellucci, and T. Gabbriellini. Dynamic option hedging via stochastic model predictive control based on scenario simulation. *Quantitative Finance*, 14(10):1739–1751, 2014.

[28] A. Bemporad, F. Borrelli, and M. Morari. Robust model predictive control: Piecewise linear explicit solution. In *Proceedings of the 2001 European Control Conference (ECC)*, pages 939–944. IEEE, 2001.

[29] A. Bemporad, F. Borrelli, and M. Morari. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on automatic control*, 48(9):1600–1606, 2003.

[30] A. Bemporad and M. Morari. Robust model predictive control: A survey. In *Robustness in identification and control*, pages 207–226. Springer, 1999.

[31] A. Bemporad, L. Puglia, and T. Gabbriellini. A stochastic model predictive control approach to dynamic option hedging with transaction costs. In *Proceedings of the 2011 American Control Conference (ACC)*, pages 3862–3867. IEEE, 2011.

[32] A. Ben-Tal, D. Den Hertog, A. De Waegenaere, B. Melenberg, and G. Rennen. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357, 2013.

[33] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009.

[34] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[35] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–13, 1999.

[36] A. Ben-Tal, A. Nemirovski, and C. Roos. Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM Journal on Optimization*, 13(2):535–560, 2002.

[37] D. Bernardini and A. Bemporad. Energy-aware robust model predictive control based on noisy wireless sensors. *Automatica*, 48(1):36–44, 2012.

[38] D. Bernardini and A. Bemporad. Stabilizing model predictive control of stochastic constrained linear systems. *IEEE Transactions on Automatic Control*, 57(6):1468–1480, 2012.

[39] D.P. Bertsekas. Stochastic optimization problems with nondifferentiable cost functionals. *Journal of Optimization Theory and Applications*, 12(2):218–231, 1973.

[40] D.P. Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena Scientific, 1995.

[41] D.P. Bertsekas and I.B. Rhodes. On the minimax reachability of target sets and target tubes. *Automatica*, 7(2):233–247, 1971.

[42] D.P. Bertsekas and J.N. Tsitsiklis. *Introduction to probability*, volume 1. Athena Scientific, 2002.

[43] D. Bertsimas, D.B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011.

[44] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

[45] D. Bertsimas and M. Sim. Tractable approximations to robust conic optimization problems. *Mathematical Programming*, 107(1-2):5–36, 2006.

[46] D. Bertsimas and A. Thiele. Robust and data-driven optimization: Modern decision-making under uncertainty. *INFORMS*, 3, 2006.

[47] A.S. Bhadouria, M. Sorci, M. Gu, G. Belfort, and J. Hahn. Optimization of membrane separation processes for protein fractionation. *Industrial & Engineering Chemistry Research*, 53(13):5103–5109, 2013.

[48] D. Bhattacharyya, S. L. Back, R. I. Kermode, and M.C. Roco. Prediction of concentration polarization and flux behavior in reverse osmosis by numerical analysis. *Journal of Membrane Science*, 48(2):231–262, 1990.

[49] G.J. Bierman. A comparison of discrete linear filtering algorithms. *IEEE Transactions on Aerospace and Electronic Systems*, 9(1):28–37, 1973.

[50] L. Blackmore, A. Bektassov, M. Ono, and B.C. Williams. Robust, optimal predictive control of jump markov linear systems using particles. In *Hybrid Systems: Computation and Control*, pages 104–117. Springer, 2007.

[51] L. Blackmore and M. Ono. Convex chance constrained predictive control without sampling. In *Proceedings of the AIAA Guidance, Navigation and Control Conference*, pages 7–21, 2009.

[52] L. Blackmore, M. Ono, A. Bektassov, and B.C. Williams. A probabilistic particle-control approximation of chance-constrained stochastic predictive control. *IEEE Transactions on Robotics*, 26(3):502–517, 2010.

[53] L. Blackmore and B.C. Williams. Optimal, robust predictive control of nonlinear systems under probabilistic uncertainty using particles. In *Proceedings of the 2007 American Control Conference (ACC)*, volume 17. IEEE, 2007.

[54] F. Blanchini and S. Miani. *Set-theoretic methods in control*. Springer, 2007.

[55] B. Blankert, B.H.L. Betlem, and B. Roffel. Dynamic optimization of a dead-end filtration trajectory: blocking filtration laws. *Journal of membrane science*, 285(1):90–95, 2006.

[56] H-S. Bosch, R.C. Wolf, T. Andreeva, J. Baldzuhn, D. Birus, T. Bluhm, T. Bräuer, H. Braune, V. Bykov, and A. Cardella et al. Technical challenges in the construction of the steady-state stellarator wendelstein 7-x. *Nuclear Fusion*, 53(12):126001, 2013.

[57] W. R. Bowen and A. W. Mohammad. Characterization and prediction of nanofiltration membrane performance-a general assessment. *Chemical Engineering Research and Design*, 76(8):885–893, 1998.

[58] W. R. Bowen and H. Mukhtar. Characterisation and prediction of separation performance of nanofiltration membranes. *Journal of Membrane Science*, 112(2):263–274, 1996.

[59] A.J.B. Van Boxtel, Z.E.H. Otten, and H.J.L.J. Van der Linden. Evaluation of process models for fouling control of reverse osmosis of cheese whey. *Journal of membrane science*, 58(1):89–111, 1991.

[60] A.J.B. Van Boxtel, Z.E.H. Otten, and H.J.L.J. Van der Linden. Dynamic optimization of a one-stage reverse-osmosis installation with respect to membrane fouling. *Journal of membrane science*, 65(3):277–293, 1992.

[61] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[62] S.P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15. SIAM, 1994.

[63] G. P. P.W. Brans, C.G.P.H Schroën, R.G.M. Van der Sman, and R.M. Boom. Membrane fractionation of milk: state of the art and challenges. *Journal of Membrane Science*, 243(1):263–272, 2004.

[64] R. W. Brockett. Lie algebras and lie groups in control theory. In D. Q. Mayne and R. W. Brockett, editors, *Geometric Methods in System Theory: Proceedings of the NATO Advanced Study Institute*, pages 43–82. Springer, 1973.

[65] G.C. Calafiore. Random convex programs. *SIAM Journal on Optimization*, 20(6):3427–3464, 2010.

[66] G.C. Calafiore and M.C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

[67] G.C. Calafiore and L. Fagiano. Robust model predictive control via scenario optimization. *IEEE Transactions on Automatic Control*, 58(1):219–224, 2013.

[68] G.C. Calafiore and L. Fagiano. Stochastic model predictive control of lpv systems via scenario optimization. *Automatica*, 49(6):1861–1866, 2013.

[69] E.F. Camacho and C.B. Alba. *Model predictive control*. Springer Science & Business Media, 2013.

[70] M.C. Campi. Why is resorting to fate wise? a critical look at randomized algorithms in systems and control. *European Journal of Control*, 16(5):419–430, 2010.

[71] M.C. Campi and S. Garatti. Wait-and-judge scenario optimization. *Mathematical Programming*, pages 1–35, 2016.

[72] P.J Campo and M. Morari. Model predictive optimal averaging level control. *AIChE Journal*, 35(4):579–591, 1989.

[73] P.J. Campo and M. Morari. Robust control of processes subject to saturation nonlinearities. *Computers & Chemical Engineering*, 14(4-5):343–358, 1990.

[74] M. Cannon, J. Buerger, B. Kouvaritakis, and S.V. Raković. Robust tubes in nonlinear model predictive control. *IEEE Transactions on Automatic Control*, 56(8):1942–1947, 2011.

[75] M. Cannon, Q. Cheng, B. Kouvaritakis, and S.V. Raković. Stochastic tube mpc with state estimation. *Automatica*, 48(3):536–541, 2012.

[76] M. Cannon, B. Kouvaritakis, and X. Wu. Probabilistic constrained mpc for multiplicative and additive stochastic uncertainty. *IEEE Transactions on Automatic Control*, 54(7):1626–1632, 2009.

[77] X. Cao, M.B. Saltik, and S. Weiland. Optimal hankel norm model reduction for discrete-time descriptor systems. *Submitted to the Journal of the Franklin Institute*, -(-):–, -.

[78] X. Cao, M.B. Saltik, and S. Weiland. Hankel model reduction for descriptor systems. In *Proceedings of the 54th IEEE Conference on Decision and Control (CDC)*, pages 4668–4673. IEEE, 2015.

[79] X. Cao, M.B. Saltik, and S. Weiland. Optimal hankel norm approximation for continuous-time descriptor systems. In *Submitted to the American Control Conference (ACC)*, pages –, 2018.

[80] A. Casavola, D. Famularo, and G. Franzé. Robust constrained predictive control of uncertain norm-bounded linear systems. *Automatica*, 40(11):1865–1876, 2004.

[81] A.L. Cervantes, O.E. Agamennoni, and J.L. Figueroa. A nonlinear model predictive control system based on wiener piecewise linear models. *Journal of process control*, 13(7):655–666, 2003.

[82] A. Charnes and W.W. Cooper. Chance-constrained programming. *Management science*, 6(1):73–79, 1959.

[83] A. Charnes, W.W. Cooper, and G.H. Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3):235–263, 1958.

[84] D. Chatterjee, P. Hokayem, and J. Lygeros. Stochastic receding horizon control with bounded control inputs: a vector space approach. *IEEE Transactions on Automatic Control*, 56(11):2704–2710, 2011.

[85] D. Chatterjee and J. Lygeros. On stability and performance of stochastic predictive control techniques. *IEEE Transactions on Automatic Control*, 60(2):509–514, 2015.

[86] C.W. Chen, H.and Scherer and F. Allgöwer. A game theoretic approach to nonlinear robust receding horizon control of constrained systems. In *Proceedings of the 1997 American Control Conference (ACC)*, volume 5, pages 3073–3077, 1997.

[87] T. Cheng and H. Yeh. Complete momentum-balance analysis of permeate flux for ultrafiltration in hollow-fiber modules. *Tamkang Journal of Science and Engineering*, 11(3):239–246, 2008.

[88] Y-L. Chow and M. Pavone. A framework for time-consistent, risk-averse model predictive control: Theory and algorithms. In *Proceedings of the 2014 American Control Conference (ACC)*, pages 4204–4211. IEEE, 2014.

[89] E. Cinquemani, M. Agarwal, D. Chatterjee, and J. Lygeros. Convexity and convex approximations of discrete-time stochastic control problems with constraints. *Automatica*, 47(9):2082–2087, 2011.

[90] D.W. Clarke and C. Mohtadi. Properties of generalized predictive control. *Automatica*, 25(6):859–875, 1989.

[91] D.W Clarke, C. Mohtadi, and P.S. Tuffs. Generalized predictive control part i: The basic algorithm. *Automatica*, 23(2):137–148, 1987.

[92] P.D. Couchman, M. Cannon, and B. Kouvaritakis. Stochastic mpc with inequality stability constraints. *Automatica*, 42(12):2169–2174, 2006.

[93] E. Crisostomi, A. Lecchini-visintini, and J.M. Maciejowski. Combining monte carlo and worst-case methods for trajectory prediction in air traffic control: a case study. In *Journal of Guidance, Control and Dynamics*, 2008.

[94] J.B. Cruz Jr, M. Simaan, A. Gacic, and Y. Liu. Moving horizon game theoretic approaches for control strategies in a military operation. In *Proceedings of the IEEE 40th Conference on Decision and Control (CDC)*, volume 1, pages 628–633. IEEE, 2001.

[95] J.B. Cruz Jr, M.A. Simaan, A. Gacic, and Y. Liu. Moving horizon nash strategies for a military air operation. *IEEE Transactions on Aerospace and Electronic Systems*, 38(3):989–999, 2002.

[96] J.R. Cueli and C. Bordons. Iterative nonlinear model predictive control. stability, robustness and applications. *Control Engineering Practice*, 16(9):1023–1034, 2008.

[97] C.R. Cutler and B.L. Ramaker. Dynamic matrix control – a computer control algorithm. In *Proceedings of the joint automatic control conference*, number 17, page 72, 1980.

[98] F.A. Cuzzola, J.C. Geromel, and M. Morari. An improved approach for constrained robust model predictive control. *Automatica*, 38(7):1183–1189, 2002.

[99] M. Dalmau, N. Atanasova, S. Gabarrón, I. Rodriguez-Roda, and J. Comas. Comparison of a deterministic and a data driven model to describe MBR fouling. *Chemical Engineering Journal*, 260:300–308, 2015.

[100] A. Damen and S. Weiland. Robust control [lecture notes]. *Measurement and Control Group, Department of Electrical Engineering, Eindhoven University of Technology*, 2002.

[101] R. D'Andrea and F. Paganini. Interconnection of uncertain behavioral systems for robust contro [memorandum]. *California Institute of Technology*, 1993.

[102] R. D'Andrea, F. Paganini, and J.C. Doyle. Uncertain behavior [memorandum]. *California Institute of Technology*, 1993.

[103] M. L. Darby, M. Nikolaou, J. Jones, and D. Nicholson. Rto: An overview and assessment of current practice. *Journal of Process Control*, 21(6):874–884, 2011.

[104] E. Darnon, M.P. Belleville, and G.M. Rios. Modeling ultrafiltration of complex biological solutions. *AIChE journal*, 48(8):1727–1736, 2002.

[105] G. Daufin, J-P. Escudier, H. Carrere, S. Berot, L. Fillaudeau, and M. Decloux. Recent and emerging applications of membrane processes in the food and dairy industry. *Food and Bioproducts Processing*, 79(2):89–102, 2001.

[106] M.M. Davis, N.J. Aquilano, J. Balakrishnan, and R.B. Chase. *Fundamentals of Operations Management*. McGraw-Hill, 2005.

[107] D. M. De la Peña, T. Alamo, and A. Bemporad. A decomposition algorithm for feedback min-max model predictive control. In *Proceedings of the Joint Conference on Decision and Control and European Control Conference. CDC-ECC'05*, pages 5126–5131. IEEE, 2005.

[108] M. De la Pena, A. Bemporad, and T. Alamo. Stochastic programming applied to model predictive control. In *Proceedings of the Joint Conference on Decision and Control and European Control Conference CDC-ECC*, pages 1361–1366. IEEE, 2005.

[109] Roy De Maesschalck, Delphine Jouan-Rimbaud, and Désiré L Massart. The mahalanobis distance. *Chemometrics and intelligent laboratory systems*, 50(1):1–18, 2000.

[110] S. Di Cairano, D. Bernardini, A. Bemporad, and I.V. Kolmanovsky. Stochastic mpc with learning for driver-predictive vehicle control and its application to hev energy management. *IEEE Transactions on Control Systems Technology*, 22(3):1018–1031, 2014.

[111] M. Diehl and J. Bjornberg. Robust dynamic programming for min-max model predictive control of constrained uncertain systems. *IEEE Transactions on Automatic Control*, 49(12):2253–2257, 2004.

[112] A.D. Dimitriadis, N. Shah, and C.C. Pantelides. Rtn-based rolling horizon algorithms for medium term scheduling of multipurpose plants. *Computers & Chemical Engineering*, 21:S1061–S1066, 1997.

[113] D. Dochain. State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of process control*, 13(8):801–818, 2003.

[114] D. Dochain, N. Tali-Maamar, and J.P. Babary. On modelling, monitoring and control of fixed bed bioreactors. *Computers & chemical engineering*, 21(11):1255–1266, 1997.

[115] J.C. Doyle, K. Glover, P.P. Khargonekar, and B. A. Francis. State-space solutions to standard h/sub 2/and h/sub infinity/control problems. *IEEE Transactions on Automatic control*, 34(8):831–847, 1989.

[116] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, 1997.

[117] L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9(1):33–52, 1998.

[118] W. Eykamp. Microfiltration and ultrafiltration. *Membrane Science and Technology*, 2:1–43, 1995.

[119] M. Farina, L. Giulioni, L. Magni, and R. Scattolini. A probabilistic approach to model predictive control. In *Proceedings of the IEEE 52nd Conference on Decision and Control (CDC)*, pages 7734–7739. IEEE, 2013.

[120] M. Farina, L. Giulioni, L. Magni, and R. Scattolini. An approach to output-feedback mpc of stochastic linear discrete-time systems. *Automatica*, 55:140–149, 2015.

[121] M. Farina and R. Scattolini. Model predictive control of linear systems with multiplicative unbounded uncertainty and average constraints. *IFAC-PapersOnLine*, 48(23):266–271, 2015.

[122] M. Farina and R. Scattolini. Model predictive control of linear systems with multiplicative unbounded uncertainty and chance constraints. *Automatica*, 70:258–265, 2016.

[123] R. Findeisen, F. Allgöwer, and L.T. Biegler. *Assessment and future directions of nonlinear model predictive control*, volume 358. Springer, 2007.

[124] J. Fleming, M. Cannon, and B. Kouvaritakis. Stochastic tube mpc for lpv systems with probabilistic set inclusion conditions. In *Proceedings of the 53rd IEEE Conference on Decision and Control (CDC)*, pages 4783–4788. IEEE, 2014.

[125] M. Florian, J.K. Lenstra, and A.H.G. Rinnooy Kan. Deterministic production planning: Algorithms and complexity. *Management science*, 26(7):669–679, 1980.

[126] M.G. Forbes, R.S. Patwardhan, H. Hamadah, and R.B. Gopaluni. Model predictive control in industry: Challenges and opportunities. *IFAC-PapersOnLine*, 48(8):531–538, 2015.

[127] K. Fukuda and A. Prodon. Double description method revisited. In *Combinatorics and computer science*, pages 91–111. Springer, 1996.

[128] V. Gabrel, C. Murat, and A. Thiele. Recent advances in robust optimization: An overview. *European Journal of Operational Research*, 235(3):471–483, 2014.

[129] C.E. Garcia, D.M. Prett, and M. Morari. Model predictive control: theory and practice a survey. *Automatica*, 25(3):335–348, 1989.

[130] F. Giralt, F. de Transport, J. Giralt, and D. Libotean. Data driven neural network based models of RO desalination plant operation and RO membrane performance. Data analysis: Design, algorithms & applications. In *Proceedings of the 2008 AICHE Annual Meeting*, 2008.

[131] J.D. Glover and F.C. Schweppe. Control of linear dynamic systems with set constrained disturbances. *IEEE Transactions on Automatic Control*, 16(5):411–423, 1971.

[132] G.C. Goodwin, D.S. Carrasco, and M.M. Seron. Predictive control: a historical perspective. *International Journal of Robust and Nonlinear Control*, 22(12):1296–1313, 2012.

[133] G.C. Goodwin, H. Kong, G. Mirzaeva, and M.M. Seron. Robust model predictive control: reflections and opportunities. *Journal of Control and Decision*, 1(2):115–148, 2014.

[134] G.C. Goodwin and A.M. Medioli. Scenario-based, closed-loop model predictive control with application to emergency vehicle scheduling. *International Journal of Control*, 86(8):1338–1348, 2013.

[135] G.C. Goodwin, M.M. Seron, and J.A. De Doná. *Constrained control and estimation: an optimisation approach*. Springer Science & Business Media, 2006.

[136] P.J. Goulart and E.C. Kerrigan. Relationships between affine feedback policies for robust control with constraints. *IFAC Proceedings Volumes*, 38(1):608–613, 2005.

[137] P.J. Goulart and E.C. Kerrigan. Input-to-state stability of robust receding horizon control with an expected value cost. *Automatica*, 44(4):1171–1174, 2008.

[138] P.J. Goulart, E.C. Kerrigan, and J.M. Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 42(4):523–533, 2006.

[139] P.J. Goulart, E.C. Kerrigan, and D. Ralph. Efficient robust optimization for robust control with constraints. *Mathematical Programming*, 114(1):115–147, 2008.

[140] gPROMS. *Model Validation Guide Release v3.6*. Process System Enterprise Ltd., 2012.

[141] S. Grammatico, X. Zhang, K. Margellos, P. Goulart, and J. Lygeros. A scenario approach for non-convex control design. *IEEE Transactions on Automatic Control*, 61(2):334–345, 2016.

[142] M.S. Grewal and K. Glover. Identifiability of linear and nonlinear dynamical systems. *IEEE Transactions on automatic control*, 21(6):833–837, 1976.

[143] I.E. Grossmann. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE Journal*, 51(7):1846–1857, 2005.

[144] J.M. Grosso, C. Ocampo-Martínez, V. Puig, and B. Joseph. Chance-constrained model predictive control for drinking water networks. *Journal of Process Control*, 24(5):504–516, 2014.

[145] L Grüne and J. Pannek. Nonlinear model predictive control. In *Nonlinear Model Predictive Control*, pages 43–66. Springer, 2011.

[146] A. Guadix, E. Sørensen, L.G. Papageorgiou, and E.M. Guadix. Optimal design and operation of continuous ultrafiltration plants. *Journal of membrane science*, 235(1):131–138, 2004.

[147] W. Guo, H-H. Ngo, and J. Li. A mini-review on membrane fouling. *Bioresource technology*, 122:27–34, 2012.

[148] Y.A. Guzman, L.R. Matthews, and C.A. Floudas. New a priori and a posteriori probabilistic bounds for robust counterpart optimization: I. unknown probability distributions. *Computers & Chemical Engineering*, 84:568–598, 2016.

[149] J. Hahn, T. F. Edgar, and W. Marquardt. Controllability and observability covariance matrices for the analysis and order reduction of stable nonlinear systems. *Journal of process control*, 13(2):115–127, 2003.

[150] H. Haimovich, T. Perez, and G.C. Goodwin. On optimality and certainty equivalence in output feedback control of constrained uncertain linear systems. In *Proceedings of the 2003 European Control Conference (ECC)*, pages 1075–1080. IEEE, 2003.

[151] P. Hall. The distribution of means for samples of size n drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable. *Biometrika*, pages 240–245, 1927.

[152] C.A. Hans, P. Sopasakis, A. Bemporad, J. Raisch, and C. Reincke-Collon. Scenario-based model predictive operation control of islanded microgrids. In *Proceedings of the IEEE 54th Conference on Decision and Control (CDC)*, pages 3272–3277, 2015.

[153] C.A. Hans, P. Sopasakis, A. Bemporad, and C. et al. Reincke-Collon. Scenario-based model predictive operation control of islanded microgrids. In *Proceedings of the 54th IEEE Conference on Decision and Control (CDC)*, pages 3272–3277. IEEE, 2015.

[154] E.L. Haseltine and J.B. Rawlings. Critical evaluation of extended kalman filtering and moving-horizon estimation. *Industrial & engineering chemistry research*, 44(8):2451–2460, 2005.

[155] T. Hashimoto. Probabilistic constrained model predictive control for linear discrete-time systems with additive stochastic disturbances. In *Proceedings of the 52nd IEEE Conference on Decision and Control (CDC)*, pages 6434–6439. IEEE, 2013.

[156] A. Hausmann, P. Sanciolo, T. Vasiljevic, U. Kulozik, and M. Duke. Performance assessment of membrane distillation for skim milk and whey processing. *Journal of dairy science*, 97(1):56–71, 2014.

[157] A. Hausmann, P. Sanciolo, T. Vasiljevic, M. Weeks, K. Schroën, S. Gray, and M. Duke. Fouling mechanisms of dairy streams during membrane distillation. *Journal of Membrane Science*, 441:102–111, 2013.

[158] R. Henrion, P. Li, A. Möller, M.C. Steinbach, M. Wendt, and G. Wozny. Stochastic optimization for operating chemical processes under uncertainty. In *Online optimization of large scale systems*, pages 457–478. Springer, 2001.

[159] F. Herzog, G. Dondi, and H.P. Geering. Stochastic model predictive control and portfolio optimization. *International Journal of Theoretical and Applied Finance*, 10(02):203–233, 2007.

[160] J.P. Hespanha. *Linear systems theory*. Princeton university press, 2009.

[161] N. Hilal, O. O. Ogunbiyi, N. J. Miles, and R. Nigmatullin. Methods employed for control of fouling in mf and uf membranes: a comprehensive review. *Separation Science and Technology*, 40(10):1957–2005, 2005.

[162] P. Hokayem, D. Chatterjee, and J. Lygeros. On stochastic receding horizon control with bounded control inputs. In *Proceedings of the Joint IEEE Conference on Decision and Control and 28th Chinese Control Conference CDC/CCC*, pages 6359–6364. IEEE, 2009.

[163] P. Hokayem, E. Cinquemani, D. Chatterjee, F. Ramponi, and J. Lygeros. Stochastic receding horizon control with output feedback and bounded controls. *Automatica*, 48(1):77–88, 2012.

[164] S. Hovland, K. Willcox, and J.T. Gravdahl. Mpc for large-scale systems via model reduction and multiparametric quadratic programming. In *Proceedings of the 45th IEEE Conference on Decision and Control (CDC)*, pages 3418–3423. IEEE, 2006.

[165] P.J. Huxel and R.H. Bishop. Navigation algorithms and observability analysis for formation flying missions. *Journal of guidance, control, and dynamics*, 32(4):1218–1231, 2009.

[166] J.O. Irwin. On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to pearson's type ii. *Biometrika*, pages 225–239, 1927.

[167] B. J. James, Y. Jing, and X. D. Chen. Membrane fouling during filtration of milk-a microstructural study. *Journal of Food Engineering*, 60(4):431–437, 2003.

[168] M. Jelemensky, M. Klauco, R. Paulen, J. Lauwers, F. Logist, J. Van Impe, and M. Fikar. Time-optimal control and parameter estimation of diafiltration processes in the presence of membrane fouling. In *Proceedings of the 11th International IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS)*. IFAC, 2016.

[169] L. Ji, J. B. Rawlings, W. Hu, A. Wynn, and M. Diehl. Robust stability of moving horizon estimation under bounded disturbances. *IEEE Transactions on Automatic Control*, PP(99):1–1, 2016.

[170] D. Jia and B. Krogh. Min-max feedback model predictive control for distributed control with communication. In *Proceedings of the 2002 American Control Conference (ACC)*, volume 6, pages 4507–4512. IEEE, 2002.

[171] K.H. Johansson. *Relay feedback and multivariable control*. PhD thesis, Lund Institute of Technology (LTH), 1997.

[172] S. Judd. *The MBR book: principles and applications of membrane bioreactors for water and wastewater treatment*. Elsevier, 2010.

[173] N.M.P. Kakalis, V. Dua, V. Sakizlis, J.D. Perkins, and E.N. Pistikopoulos. A parametric optimisation approach for robust mpc. In *Proceedings of 15th IFAC World Congress*, 2002.

[174] R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *Journal of basic engineering*, 83(1):95–108, 1961.

[175] S. Kanev, C.W. Scherer, M. Verhaegen, and B. De Schutter. Robust output-feedback controller design via local bmi optimization. *Automatica*, 40(7):1115–1127, 2004.

[176] S. Kanev and M. Verhaegen. Robustly asymptotically stable finite-horizon mpc. *Automatica*, 42(12):2189–2194, 2006.

[177] N. Kantas, J.M. Maciejowski, and A. Lecchini-Visintini. Sequential monte carlo for model predictive control. In *Nonlinear Model Predictive Control*, pages 263–273. Springer, 2009.

[178] E.C. Kerrigan and J.M. Maciejowski. On robust optimization and the optimal control of constrained linear systems with bounded state disturbances. In *Proceedings of the 2003 European Control Conference (ECC)*, pages 1453–1458. IEEE, 2003.

[179] E.C Kerrigan and J.M. Maciejowski. Robustly stable feedback min-max model predictive control. In *Proceedings of the 2003 American Control Conference (ACC)*, volume 4, pages 3490–3495. IEEE, 2003.

[180] E.C. Kerrigan and J.M. Maciejowski. Feedback min-max model predictive control using a single linear program: robust stability and the explicit solution. *International Journal of Robust and Nonlinear Control*, 14(4):395–413, 2004.

[181] Eric Colin Kerrigan. *Robust constraint satisfaction: Invariant sets and predictive control*. PhD thesis, University of Cambridge, 2000.

[182] I. Kolmanovsky and E.G. Gilbert. Theory and computation of disturbance invariant sets for discrete-time linear systems. *Mathematical Problems in Engineering*, 4(4):317–367, 1998.

[183] M. Korda and J. Cigler. Nonquadratic stochastic model predictive control: A tractable approach. *Automatica*, 48(9):2352–2358, 2012.

[184] M. Korda, R. Gondhalekar, J. Cigler, and F. Oldewurtel. Strongly feasible stochastic model predictive control. In *Proceedings of the Joint IEEE Conference on Decision and Control and European Control Conference*, pages 1245–1251. IEEE, 2011.

[185] M.V. Kothare, V. Balakrishnan, and M. Morari. Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10):1361–1379, 1996.

[186] K.I. Kouramas, C. Panos, N.P. Faísca, and E.N. Pistikopoulos. An algorithm for robust explicit/multi-parametric model predictive control. *Automatica*, 49(2):381–389, 2013.

[187] B. Kouvaritakis and M. Cannon. *Non-linear Predictive Control: theory and practice*. Number 61. Iet, 2001.

[188] B. Kouvaritakis and M. Cannon. *Model Predictive Control*. Springer, 2016.

[189] B. Kouvaritakis, M. Cannon, S.V. Raković, and Q. Cheng. Explicit use of probabilistic distributions in linear predictive control. *Automatica*, 46(10):1719–1724, 2010.

[190] B. Kouvaritakis, M. Cannon, and R. Yadlin. On the centres and scalings of robust and stochastic tubes. In *Proceedings of the UKACC International Conference on Control*, pages 1–6. IET, 2010.

[191] F. Kozin. A survey of stability of stochastic systems. *Automatica*, 5(1):95–112, 1969.

[192] C. Kravaris, J. Hahn, and Y. Chu. Advances and selected recent developments in state and parameter estimation. *Computers & chemical engineering*, 51:111–123, 2013.

[193] P. Kühl, M. Diehl, T. Kraus, J.P. Schlöder, and H.G. Bock. A real-time algorithm for moving horizon state and parameter estimation. *Computers & chemical engineering*, 35(1):71–83, 2011.

[194] W. Kühn. Rigorously computed orbits of dynamical systems without the wrapping effect. *Computing*, 61(1):47–67, 1998.

[195] S. Kumar and J.H. Seinfeld. Optimal location of measurements in tubular reactors. *Chemical engineering science*, 33(11):1507–1516, 1978.

[196] V.M. Kuntsevich and B.N. Pshenichnyi. Minimal invariant sets of dynamic systems with bounded disturbances. *Cybernetics and Systems Analysis*, 32(1):58–64, 1996.

[197] W.H. Kwon and S.H. Han. *Receding horizon control: model predictive control for state models*. Springer Science & Business Media, 2006.

[198] S. Lall, J.E. Marsden, and S. Glavaški. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *International journal of robust and nonlinear control*, 12(6):519–535, 2002.

[199] W. Langson, I. Chryssochoos, S.V. Raković, and D.Q. Mayne. Robust model predictive control using tubes. *Automatica*, 40(1):125–133, 2004.

[200] M. Lazar, D. M. De La Peña, W. P. M. H. Heemels, and T. Alamo. Min-max nonlinear model predictive control with guaranteed input-to-state stability. In *Proceedings of the 17th Symposium on Mathematical Theory for Networks and Systems*, 2006.

[201] M. Lazar, D. M. De La Peña, W. P. M. H. Heemels, and T. Alamo. On input-to-state stability of min–max nonlinear model predictive control. *Systems & Control Letters*, 57(1):39–48, 2008.

[202] M. Lazar and R.H. Gielen. On parameterized dissipation inequalities and receding horizon robust control. *International journal of robust and nonlinear control*, 22(12):1314–1329, 2012.

[203] M. Lazar, W.P.M.H. Heemels, A. Bemporad, and S. Weiland. On the stability and robustness of non-smooth nonlinear model predictive control. In *Proceedings of the Workshop on Assessment and Future Directions of NMPC*, 2005.

[204] H. Lee, J. M. Pinto, I.E. Grossmann, and S. Park. Mixed-integer linear programming model for refinery short-term scheduling of crude oil unloading with inventory management. *Industrial & Engineering Chemistry Research*, 35(5):1630–1641, 1996.

[205] J. H. Lee and N. L. Ricker. Extended kalman filter based nonlinear model predictive control. *Industrial & Engineering Chemistry Research*, 33(6):1530–1541, 1994.

[206] J.H. Lee. Model predictive control: review of the three decades of development. *International Journal of Control, Automation and Systems*, 9(3):415–424, 2011.

[207] J.H. Lee. From robust model predictive control to stochastic optimal control and approximate dynamic programming: A perspective gained from a personal journey. *Computers & Chemical Engineering*, 70:114–121, 2014.

[208] J.H. Lee and Z. Yu. Worst-case formulations of model predictive control for systems with bounded parameters. *Automatica*, 33(5):763–781, 1997.

[209] J.H. Lee and Z.H. Yu. Tuning of model predictive controllers for robust performance. *Computers & Chemical Engineering*, 18(1):15–37, 1994.

[210] P. Li, H. Arellano-Garcia, and G. Wozny. Chance constrained programming approach to process optimization under uncertainty. *Computers & Chemical Engineering*, 32(1):25–45, 2008.

[211] P. Li, M. Wendt, and G. Wozny. Robust model predictive control under chance constraints. *Computers & Chemical Engineering*, 24(2):829–834, 2000.

[212] R. Li, M. Henson, and M. J. Kurtz. Selection of model parameters for off-line parameter estimation. *IEEE Transactions on control systems technology*, 12(3):402–412, 2004.

[213] Z. Li and M.G. Ierapetritou. Integrated production planning and scheduling using a decomposition framework. *Chemical Engineering Science*, 64(16):3585–3597, 2009.

[214] Furong Liu, Taoand Gao. *Cascade Control System*, pages 321–347. Springer London, London, 2012.

[215] L. Ljung. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1):36–50, 1979.

[216] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, 1999.

[217] J. Löfberg. Approximations of closed-loop minimax mpc. In *Proceedings of the IEEE 43rd Conference on Decision and Control (CDC)*, volume 2, pages 1438–1442. IEEE, 2003.

[218] J. Löfberg. Modeling and solving uncertain optimization problems in yalmip. In *Proceedings of the 17th IFAC World Congress*, pages 1337–1341, 2008.

[219] C. Løvaas, M.M. Seron, and G.C. Goodwin. A dissipativity approach to robustness in constrained model predictive control. In *Proceedings of the IEEE 46th Conference on Decision and Control (CDC)*, pages 1180–1185. IEEE, 2007.

[220] C. Løvaas, M.M. Seron, and G.C. Goodwin. Robust model predictive control of input-constrained stable systems with unstructured uncertainty. In *Proceedings of the 2007 European Control Conference (ECC)*, pages 3303–3310. IEEE, 2007.

[221] C. Løvaas, M.M. Seron, and G.C. Goodwin. Robust output-feedback model predictive control for systems with unstructured uncertainty. *Automatica*, 44(8):1933–1943, 2008.

[222] C. Løvaas, M.M. Seron, and G.C. Goodwin. Robust output-feedback mpc with integral action. *IEEE Transactions on Automatic Control*, 55(7):1531–1543, 2010.

[223] Y. Lu and Y. Arkun. Quasi-min-max mpc algorithms for lpv systems. *Automatica*, 36(4):527–540, 2000.

[224] J.H.A Ludlage. *Controllability Analysis of Industrial Processes Towards the Industrial Application*. PhD thesis, Eindhoven University of Technology, 1997.

[225] B.F. Lund and B.A. Foss. Parameter ranking by orthogonalization-applied to nonlinear mechanistic models. *Automatica*, 44(1):278–281, 2008.

[226] J.M. Maciejowski. *Predictive control: with constraints*. Pearson education, 2002.

[227] L. Magni, D. Pala, and R. Scattolini. Stochastic model predictive control of constrained linear systems with additive uncertainty. In *Proceedings of the 2009 European Control Conference (ECC)*, pages 2235–2240. IEEE, 2009.

[228] M. Mahmood and P. Mhaskar. Lyapunov-based model predictive control of stochastic nonlinear systems. *Automatica*, 48(9):2271–2276, 2012.

[229] H.M. Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.

[230] H.M. Markowitz. Foundations of portfolio theory. *The journal of finance*, 46(2):469–477, 1991.

[231] J.I. Marriott, S. Sørensen, and I.D.L. Bogle. Detailed mathematical modelling of membrane modules. *Computers and Chemical Engineering*, 25:693–700, 2001.

[232] D.L. Marruedo, T. Alamo, and E.F. Camacho. Input-to-state stable mpc for constrained discrete-time nonlinear systems with bounded additive uncertainties. In *Proceedings of the IEEE 41st Conference on Decision and Control (CDC)*, volume 4, pages 4619–4624. IEEE, 2002.

[233] D.L. Marruedo, T. Alamo, D.M. Raimondo, D.M. de la Pena, J.M. Bravo, A. Ferramosca, and E.F. Camacho. Input-to-state stability: a unifying framework for robust model predictive control. In *Nonlinear model predictive control*, pages 1–26. Springer, 2009.

[234] J. C. Maxwell. On governors. *Proceedings of the Royal Society of London*, 16:270–283, 1867.

[235] D.Q. Mayne. Model predictive control: Recent developments and future promise. *Automatica*, 50(12):2967–2986, 2014.

[236] D.Q. Mayne. Robust and stochastic model predictive control: Are we going in the right direction? *Annual Reviews in Control*, 2016.

[237] D.Q. Mayne and E.C. Kerrigan. Tube-based robust nonlinear model predictive control. In *Nonlinear Control Systems*, volume 7, pages 36–41, 2007.

[238] D.Q. Mayne, E.C. Kerrigan, and P. Falugi. Robust model predictive control: advantages and disadvantages of tube-based methods. In *Proceedings of the 18th IFAC World Congress*, pages 191–196, 2011.

[239] D.Q. Mayne, E.C. Kerrigan, E.J. Van Wyk, and P. Falugi. Tube-based robust nonlinear model predictive control. *International Journal of Robust and Nonlinear Control*, 21(11):1341–1353, 2011.

[240] D.Q. Mayne, S.V. Raković, R. Findeisen, and F. Allgöwer. Robust output feedback model predictive control of constrained linear systems. *Automatica*, 42(7):1217–1222, 2006.

[241] D.Q. Mayne, S.V. Raković, R. Findeisen, and F. Allgöwer. Robust output feedback model predictive control of constrained linear systems: Time varying case. *Automatica*, 45(9):2082–2087, 2009.

[242] D.Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O.M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

[243] D.Q. Mayne, M.M. Seron, and S.V. Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, 2005.

[244] A. Maysoon. Assessment of mathematical models for ultrafiltration of multi-solute continuous cross-flow process. Master's thesis, University of Waterloo, Ontario, Canada, 2012.

[245] M.H. McAndrew. On the product of directed graphs. *Proceedings of the American Mathematical Society*, 14(4):600–606, 1963.

[246] A. Mesbah. Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems Magazine*, 36(6):30–44, 2016.

[247] P. Mhaskar, N.H. El-Farra, and P.D. Christofides. Stabilization of nonlinear systems with state and control constraints using lyapunov-based predictive control. *Systems & Control Letters*, 55(8):650–659, 2006.

[248] H. Michalska and D.Q Mayne. Robust receding horizon control of constrained nonlinear systems. *IEEE transactions on automatic control*, 38(11):1623–1633, 1993.

[249] D.A. Mindell. Opening black's box: Rethinking feedback's myth of origin. *Technology and Culture*, 41(3):405–434, 2000.

[250] M. Morari and J.H. Lee. Model predictive control: past, present and future. *Computers & Chemical Engineering*, 23(4):667–682, 1999.

[251] P.C. Müller and H.I. Weber. Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems. *Automatica*, 8(3):237–246, 1972.

[252] K. R. Muske, J. B. Rawlings, and J. H. Lee. Receding horizon recursive state estimation. In *Proceedings of the 1993 American Control Conference (ACC)*, pages 900–904. IEEE, 1993.

[253] Z.K. Nagy and R.D. Braatz. Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis. *Journal of Process Control*, 14(4):411–422, 2004.

[254] S.I. Nakao, T. Nomura, and S. Kimura. Characteristics of macromolecular gel layer formed on ultrafiltration tubular membrane. *AIChE Journal*, 25(4):615–622, 1979.

[255] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2006.

[256] U. Nieken and U. Tuttlies. Chemical reaction engineering [lecture notes]. *Institut für Chemische Verfahrenstechnik, Universität Stuttgart, Germany*, 2014.

[257] M. O. Nigam, B. Bansal, and X.D. Chen. Fouling and cleaning of whey protein concentrate fouled ultrafiltration membranes. *Desalination*, 218(1):313–322, 2008.

[258] H. Nijmeijer. Observability of autonomous discrete time non-linear systems: a geometric approach. *International journal of control*, 36(5):867–874, 1982.

[259] F. Oldewurtel, C.N. Jones, and M. Morari. A tractable approximation of chance constrained stochastic mpc based on affine disturbance feedback. In *Proceedings of the IEEE 47th Conference on Decision and Control (CDC)*, pages 4731–4736. IEEE, 2008.

[260] F. Oldewurtel, A. Parisio, C. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth. Energy efficient building climate control using stochastic model predictive control and weather predictions. In *Proceedings of the 2010 American control conference (ACC)*, number EPFL-CONF-169733, pages 5100–5105. IEEE, 2010.

[261] S. Omatu, S. Koide, and T. Soeda. Optimal sensor location problem for a linear distributed parameter system. *IEEE Transactions on automatic control*, 23(4):665–673, 1978.

[262] M. Ono, L. Blackmore, and B.C. Williams. Chance constrained finite horizon optimal control with nonconvex constraints. In *Proceedings of the 2010 American Control Conference (ACC)*, pages 1145–1152. IEEE, 2010.

[263] L. Özkan, M.V. Kothare, and C. Georgakis. Model predictive control of nonlinear systems using piecewise linear models. *Computers & Chemical Engineering*, 24(2):793–799, 2000.

[264] L. Özkan, M.V. Kothare, and C. Georgakis. Control of a solution copolymerization reactor using multi-model predictive control. *Chemical engineering science*, 58(7):1207–1221, 2003.

[265] G. Pannocchia. Robust model predictive control with guaranteed setpoint tracking. *Journal of Process Control*, 14(8):927–937, 2004.

[266] C.C. Pantelides and J.G. Renfro. The online use of first-principles models in process operations: review, current status and future needs. *Computers & Chemical Engineering*, 51:136–148, 2013.

[267] A. Papoulis and S.U. Pillai. *Probability, random variables, and stochastic processes*. McGraw-Hill, 2002.

[268] J-H. Park, T-H. Kim, and T. Sugie. Output feedback model predictive control for lpv systems based on quasi-min–max algorithm. *Automatica*, 47(9):2052–2058, 2011.

[269] X. Pascual, H. Gu, A.R. Bartman, A. Zhu, A. Rahardianto, J. Giralt, R. Rallo, P.D. Christofides, and Y. Cohen. Data-driven models of steady state and transient operations of spiral-wound ro plant. *Desalination*, 316:154–161, 2013.

[270] R. Paulen and M. Fikar. *Optimal Operation of Batch Membrane Processes*. Springer, 2016.

[271] R. Paulen, M. Fikar, G. Foley, Z. Kovács, and P. Czermak. Optimal feeding strategy of diafiltration buffer in batch membrane processes. *Journal of Membrane Science*, 411:160–172, 2012.

[272] R. Paulen, G. Foley, M. Fikar, Z. Kovács, and P. Czermak. Minimizing the process time for ultrafiltration/diafiltration under gel polarization conditions. *Journal of Membrane Science*, 380(1):148–154, 2011.

[273] R. Paulen, M. Jelemenskỳ, Z. Kovács, and M. Fikar. Economically optimal batch diafiltration via analytical multi-objective optimal control. *Journal of Process Control*, 28:73–82, 2015.

[274] D. Peaucelle and D. Arzelier. Robust performance analysis with lmi-based methods for real parametric uncertainty via parameter-dependent lyapunov functions. *IEEE Transactions on Automatic Control*, 46(4):624–630, 2001.

[275] T. Pérez and G.C. Goodwin. Stochastic output feedback model predictive control. In *Proceedings of the 2001 American Control Conference (ACC)*, volume 3, pages 2412–2417. IEEE, 2001.

[276] K.B. Petersen and M.S. Pedersen. The matrix cookbook. *Technical University of Denmark*, 7:15, 2008.

[277] G. Pin, L. Magni, T. Parisini, and D.M. Raimondo. Robust receding-horizon control of nonlinear systems with state dependent uncertainties: an input-to-state stability approach. In *Proceedings of the 2008 American Control Conference (ACC)*, pages 1667–1672. IEEE, 2008.

[278] B. Pluymers, J.A. Rossiter, J.A.K. Suykens, and B. De Moor. A simple algorithm for robust mpc. In *Proceedings of the 16th IFAC World Congress*, 2005.

[279] B. Pluymers, J.A.K. Suykens, and B. De Moor. Min–max feedback mpc using a time-varying terminal constraint set and comments on efficient robust constrained model predictive control with a time-varying terminal constraint set. *Systems & Control Letters*, 54(12):1143–1148, 2005.

[280] Y. Pochet and L.A. Wolsey. *Production Planning by Mixed Integer Programming*. Springer Science & Business Media, 2006.

[281] B.T. Polyak and E.N. Gryazina. Billiard walk-a new sampling algorithm for control and optimization. *Proceedings of the IFAC Worl Congress*, 47(3):6123–6128, 2014.

[282] L. Pontryagin. On the evasion process in differential games. *Applied Mathematics & Optimization*, 1(1):5–19, 1974.

[283] Y. Pouliot. Membrane processes in dairy technology—from a simple idea to worldwide panacea. *International Dairy Journal*, 18(7):735–740, 2008.

[284] M. Prandini, S. Garatti, and J. Lygeros. A randomized approach to stochastic model predictive control. In *Proceedings of the IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 7315–7320. IEEE, 2012.

[285] A. Prékopa. Logarithmic concave measures with application to stochastic programming. *Acta Scientiarum Mathematicarum*, 32(3-4):301, 1971.

[286] A. Prékopa. *Stochastic programming*. Springer, 1995.

[287] J.A. Primbs and C.H. Sung. Stochastic receding horizon control of constrained linear systems with state and control multiplicative noise. *IEEE Transactions on Automatic Control*, 54(2):221–230, 2009.

[288] S.J. Qin and T.A. Badgwell. An overview of industrial model predictive control technology. In *AIChE Symposium Series*, volume 93, pages 232–256. AIChE, 1997.

[289] S.J. Qin and T.A. Badgwell. An overview of nonlinear model predictive control applications. In *Nonlinear model predictive control*, pages 369–392. Springer, 2000.

[290] S.J. Qin and T.A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.

[291] D.M. Raimondo, D. Limon, M. Lazar, L. Magni, and E.F. Camacho. Regional input-to-state stability of min-max model predictive control. In *Nonlinear Control Systems*, volume 7, pages 42–47, 2007.

[292] D.M. Raimondo, D.M. Limon, M. Lazar, L. Magni, and E.F. Camacho. Min-max model predictive control of nonlinear systems: A unifying overview on stability. *European Journal of Control*, 15(1):5–21, 2009.

[293] S.V. Rakovic. Minkowski algebra and banach contraction principle in set invariance for linear discrete time systems. In *Proceedings of 46th IEEE Conference on Decision and Control (CDC)*, pages 2169–2174. IEEE, 2007.

[294] S.V. Raković, E.C. Kerrigan, K.I. Kouramas, and D.Q. Mayne. Invariant approximations of the minimal robust positively invariant set. *IEEE Transactions on Automatic Control*, 50(3):406–410, 2005.

[295] S.V. Raković, B. Kouvaritakis, M. Cannon, and C. Panos. Fully parameterized tube model predictive control. *International Journal of Robust and Nonlinear Control*, 22(12):1330–1361, 2012.

[296] S.V. Raković, B. Kouvaritakis, M. Cannon, C. Panos, and R. Findeisen. Parameterized tube model predictive control. *IEEE Transactions on Automatic Control*, 57(11):2746–2761, 2012.

[297] S.V. Raković, B. Kouvaritakis, R. Findeisen, and M. Cannon. Homothetic tube model predictive control. *Automatica*, 48(8):1631–1638, 2012.

[298] S.V. Rakovic and D.Q. Mayne. A simple tube controller for efficient robust model predictive control of constrained linear discrete time systems subject to bounded disturbances. In *Proceedings of the 16th IFAC World Congress*, 2005.

[299] C.V. Rao. *Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems*. PhD thesis, University of Wisconsin, 2000.

[300] C.V. Rao, J.B. Rawlings, and D.Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, 48(2):246–258, 2003.

[301] J.B. Rawlings. Tutorial overview of model predictive control. *IEEE Control Systems Magazine*, 20(3):38–52, 2000.

[302] J.B. Rawlings and D.Q. Mayne. *Model predictive control: Theory and design*. Nob Hill Pub., 2009.

[303] L.R. Ray and R.F. Stengel. A monte carlo approach to the analysis of control system robustness. *Automatica*, 29(1):229–236, 1993.

[304] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen. Stochastic stability of the continuous-time extended kalman filter. *IEE Proceedings - Control Theory and Applications*, 147(1):45–52, 2000.

[305] A.P. Reverberi, V.P. Meshalkin, C. Cerrato, and Y. O. Savina. Dynamics of a reverse osmosis unit with application to pulsating regimes for process optimization. *Theoretical Foundations of Chemical Engineering*, 45(2):190–197, 2011.

[306] J. Richalet. Industrial applications of model based predictive control. *Automatica*, 29(5):1251–1274, 1993.

[307] J. Richalet, A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: Applications to industrial processes. *Automatica*, 14(5):413–428, 1978.

[308] R.T. Rockafellar. Coherent approaches to risk in optimization under uncertainty. *Tutorials in operations research*, 3:38–61, 2007.

[309] R.T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42, 2000.

[310] R.T. Rockafellar, S. Uryasev, and M. Zabarankin. Generalized deviations in risk analysis. *Finance and Stochastics*, 10(1):51–74, 2006.

[311] M.A. Rodrigues and D. Odloak. Output feedback mpc with guaranteed robust stability. *Journal of Process Control*, 10(6):557–572, 2000.

[312] S.S. Sablani, M.F.A. Goosen, R. Al-Belushi, and M. Wilf. Concentration polarization in ultrafiltration and reverse osmosis: a critical review. *Desalination*, 141(3):269–289, 2001.

[313] D.R. Saffer and F.J. Doyle. Analysis of linear programming in model predictive control. *Computers & chemical engineering*, 28(12):2749–2763, 2004.

[314] N.V. Sahinidis. Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering*, 28(6):971–983, 2004.

[315] V. Sakizlis, N.M.P. Kakalis, V. Dua, J.D. Perkins, and E.N. Pistikopoulos. Design of robust model-based controllers via parametric programming. *Automatica*, 40(2):189–201, 2004.

[316] M.B. Saltik, L. Özkan, M. Jacobs, and A. van der Padt. Optimal start-up and operation policy for an ultrafiltration membrane unit in whey separation. In *Proceedings of the 26th European Symposium on Computer Aided Process Engineering (ESCAPE)*, volume 38, pages 1599–1604. Elsevier BV, 2016.

[317] M.B. Saltik, L. Özkan, S. Weiland, and J.H.A Ludlage. Moment based model predictive control for
systems with additive uncertainty. In *Proceedings of American Control Conference (ACC)*. IEEE, 2017.

[318] M.B. Saltik, L. Özkan, S. Weiland, and P. M. J. Van den Hof. Sensor configuration problem: application to a membrane separation unit. *IFAC-PapersOnLine*, 49(7):189–194, 2016.

[319] S. Sarykalin, G. Serraino, and S. Uryasev. Value-at-risk vs. conditional value-at-risk in risk management and optimization. *INFORMS*, pages 270–294, 2008.

[320] A. Saxena, B.P. Tripathi, M. Kumar, and V.K. Shahi. Membrane-based techniques for the separation and purification of proteins: an overview. *Advances in Colloid and Interface Science*, 145(1):1–22, 2009.

[321] C.W. Scherer. Lmi relaxations in robust control. *European Journal of Control*, 12(1):3–29, 2006.

[322] J. Scherpen and W.S. Gray. Minimality and local state decompositions of a nonlinear state space realization using energy functions. *IEEE Transactions on automatic control*, 45(11):2079–2086, 2000.

[323] G. Schildbach, G.C. Calafiore, L. Fagiano, and M. Morari. Randomized model predictive control for stochastic linear systems. In *Proceedings of the 2012 American Control Conference (ACC)*, pages 417–422. IEEE, 2012.

[324] G. Schildbach and M. Morari. Scenario mpc for linear time-varying systems with individual chance constraints. In *Proceedings of the 2015 American Control Conference (ACC)*, pages 415–421. IEEE, 2015.

[325] A. Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.

[326] A.T. Schwarm and M. Nikolaou. Chance-constrained model predictive control. *AIChE Journal*, 45(8):1743–1752, 1999.

[327] P.O.M. Scokaert and D.Q. Mayne. Min-max feedback model predictive control for constrained linear systems. *IEEE Transactions on Automatic Control*, 43(8):1136–1142, 1998.

[328] N. Shah. Mathematical programming techniques for crude oil scheduling. *Computers & Chemical Engineering*, 20:S1227–S1232, 1996.

[329] A. Shapiro. Monte carlo sampling methods. *Handbooks in operations research and management science*, 10:353–425, 2003.

[330] A. Sharma, M. Jelemensky, R. Paulen, and M. Fikar. Estimation of membrane fouling parameters for concentrating lactose using nanofiltration. In *Proceedings of the 26th European Symposium on Computer Aided Process Engineering (ESCAPE)*, 2016.

[331] W.F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964.

[332] R. Sheikholeslami. Fouling mitigation in membrane processes: report on a workshop held january 26-29, 1999, technion. *Desalination*, 123(1):45–53, 1999.

[333] A.K. Singh and J. Hahn. Determining optimal sensor locations for state and parameter estimation for stable nonlinear systems. *Industrial & engineering chemistry research*, 44(15):5645–5659, 2005.

[334] A.K. Singh and J. Hahn. On the use of empirical gramians for controllability and observability analysis. In *Proceedings of the 2005 American Control Conference*, volume 1, page 140, 2005.

[335] A.K. Singh and J. Hahn. Sensor location for stable nonlinear dynamic systems: Multiple sensor case. *Industrial & engineering chemistry research*, 45(10):3615–3623, 2006.

[336] M.M. Siraj. *Reducing the effect of uncertainty in robust optimization for oil recovery*. PhD thesis, Eindhoven University of Technology, 2017.

[337] E.D. Sontag. Input to state stability: Basic concepts and results. In *Nonlinear and optimal control theory*, pages 163–220. Springer, 2008.

[338] M. Soroush. State and parameter estimations and their applications in process control. *Computers & Chemical Engineering*, 23(2):229–245, 1998.

[339] B. Srinivasan, D. Bonvin, E. Visser, and S. Palanki. Dynamic optimization of batch processes: II. role of measurements in handling uncertainty. *Computers & chemical engineering*, 27(1):27–44, 2003.

[340] D. Sui, L. Feng, and M. Hovd. Robust explicit moving horizon control and estimation: A batch polymerization case study. *Modeling, Identification and Control*, 30(1):17–25, 2009.

[341] D. Sui, L. Feng, C.J. Ong, and M. Hovd. Robust explicit model predictive control for linear systems via interpolation techniques. *International Journal of Robust and Nonlinear Control*, 20(10):1166–1175, 2010.

[342] C. Sumana and C.H. Venkateswarlu. Optimal selection of sensors for state estimation in a reactive distillation process. *Journal of process control*, 19(6):1024–1035, 2009.

[343] H.J. Sussmann and J.C. Willems. 300 years of optimal control: from the brachystochrone to the maximum principle. *IEEE Control Systems Magazine*, 17(3):32–44, 1997.

[344] K. Sutherland. *Profile of the International Membrane Industry-Market Prospects to 2008*. Elsevier, 2003.

[345] N. Tali-Maamar, J.P. Babary, and D. Dochain. Influence of the sensor location on the practical observability of distributed parameter bioreactors. In *Proceedings of the International conference on control*, volume 1, pages 255–260. IET, 1994.

[346] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005.

[347] D. Telen, M. Vallerio, L. Cabianca, B. Houska, J. Van Impe, and F. Logist. Approximate robust optimization of nonlinear systems under parametric uncertainty and process noise. *Journal of Process Control*, 33:140–154, 2015.

[348] R. Tempo, G. Calafiore, and F. Dabbene. *Randomized algorithms for analysis and control of uncertain systems: with applications*. Springer Science & Business Media, 2012.

[349] J.M.K. Timmer and H.C. Van der Horst. Whey processing and separation technology: state-of-the-art and new developments. In *Proceedings of the 1997 International Whey Conference*. International Dairy Federation, 1998.

[350] N Quang Tran. *Tuning model-based controllers for autonomous maintenance*. PhD thesis, Technische Universiteit Eindhoven, 2015.

[351] T. Tran, K-V. Ling, and J.M. Maciejowski. Model predictive control via quadratic dissipativity constraint. In *Proceedings of the IEEE 53rd Conference on Decision and Control (CDC)*, pages 6689–6694. IEEE, 2014.

[352] F.W.J. Van den Berg, H.C.J. Hoefsloot, H.F.M. Boelens, and A.K. Smilde. Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria. *Chemical engineering science*, 55(4):827–837, 2000.

[353] A. van der Padt. Membrane fractionation: Dairy process chain. Lecture Presentation of Course FPE-20306 - Food Process Engineering, 2013.

[354] R.G.M. van der Sman and H.M. Vollebregt. Transient critical flux due to coupling of fouling mechanisms during crossflow microfiltration of beer. *Journal of Membrane Science*, 435:21–37, 2013.

[355] D.H. Van Hessem and O.H. Bosgra. A conic reformulation of model predictive control including bounded and stochastic disturbances under state and input constraints. In *Proceedings of the 41st IEEE Conference on Decision and Control*, volume 4, pages 4643–4648. IEEE, 2002.

[356] D.H. Van Hessem and O.H. Bosgra. A full solution to the constrained stochastic closed-loop mpc problem via state and innovations feedback and its receding horizon implementation. In *Proceedings of the 42nd IEEE Conference on Decision and Control (CDC)*, volume 1, pages 929–934. IEEE, 2003.

[357] D.H. Van Hessem and O.H. Bosgra. Stochastic closed-loop model predictive control of continuous nonlinear chemical processes. *Journal of Process Control*, 16(3):225–241, 2006.

[358] L.G. van Willigenburg, H.M. Vollebregt, and R.G.M. van der Sman. Optimal adaptive scheduling and control of beer membrane filtration. *Control Engineering Practice*, 34:77–87, 2015.

[359] J.A.W. Vissers. *Model-based estimation and control methods for batch cooling crystallizers*. PhD thesis, Technische Universiteit Eindhoven, 2012.

[360] M.P. Vitus and C.J. Tomlin. On feedback design and risk allocation in chance constrained control. In *Proceedings of the Joint IEEE Conference on Decision and Control and European Control Conference*, pages 734–739. IEEE, 2011.

[361] A. Voelker, K. Kouramas, and E.N. Pistikopoulos. Simultaneous design of explicit/multi-parametric constrained moving horizon estimation and robust model predictive control. *Computers & Chemical Engineering*, 54:24–33, 2013.

[362] W. Waldraff, D. Dochain, S. Bourrel, and A. Magnus. On the use of observability measures for sensor location in tubular reactor. *Journal of process control*, 8(5):497–505, 1998.

[363] Z. Wan and M.V. Kothare. Robust output feedback model predictive control using off-line linear matrix inequalities. *Journal of Process Control*, 12(7):763–774, 2002.

[364] Z. Wan and M.V. Kothare. Efficient robust constrained model predictive control with a time varying terminal constraint set. *Systems & Control Letters*, 48(5):375–383, 2003.

[365] Z. Wan, B. Pluymers, M.V. Kothare, and B. De Moor. Comments on: efficient robust constrained model predictive control with a time varying terminal constraint set by wan and kothare. *Systems & Control Letters*, 55(7):618–621, 2006.

[366] L. Wang, G. Libert, and P. Manneback. Kalman filter algorithm based on singular value decomposition. In *Proceedings of the 31st IEEE Conference on Decision and Control (CDC)*, pages 1224–1229. IEEE, 1992.

[367] Y.J. Wang and J.B. Rawlings. A new robust model predictive control method i: theory and computation. *Journal of Process Control*, 14(3):231–247, 2004.

[368] H.S. Witsenhausen. A minimax control problem for sampled linear systems. *IEEE Transactions on Automatic Control*, 13(1):5–21, 1968.

[369] L. Xie, P. Li, and G. Wozny. Chance constrained nonlinear model predictive control. In *Assessment and Future Directions of Nonlinear Model Predictive Control*, pages 295–304. Springer, 2007.

[370] X. Yang, D.W. Griffith, and L.T. Biegler. Nonlinear programming properties for stable and robust nmpc. *IFAC-PapersOnLine*, 48(23):388–397, 2015.

[371] X. Yang and J.M. Maciejowski. Risk-sensitive model predictive control with gaussian process models. *IFAC-PapersOnLine*, 48(28):374–379, 2015.

[372] K.W.K. Yee, A. Alexiadis, J. Bao, and D.E. Wiley. Effects of multiple-stage membrane process designs on the achievable performance of automatic control. *Journal of Membrane Science*, 320(1):280–291, 2008.

[373] K.W.K. Yee, A. Alexiadis, J. Bao, and D.E. Wiley. Effects of recycle ratios on process dynamics and operability of a whey ultrafiltration stage. *Desalination*, 236(1):216–223, 2009.

[374] K.W.K. Yee, D.E. Wile, and J. Bao. A unified model of the time dependence of flux decline for the long-term ultrafiltration of whey. *Journal of Membrane Science*, 332(1):69–80, 2009.

[375] M.S. Yorgun, I.A. Balcioglu, and O. Saygin. Performance comparison of ultrafiltration, nanofiltration and reverse osmosis on whey treatment. *Desalination*, 229(1):204–216, 2008.

[376] S. Yu, T. Qu, F. Xu, and H. Chen. Model predictive control of linear stochastic systems with constraints. In *Proceedings of the 2015 American Control Conference (ACC)*, pages 950–955. IEEE, 2015.

[377] E. Zafiriou. Robust model predictive control of processes with hard constraints. *Computers & Chemical Engineering*, 14(4-5):359–371, 1990.

[378] X. Zhang, S. Grammatico, G. Schildbach, P. Goulart, and J. Lygeros. On the sample size of randomized mpc for chance-constrained systems with application to building climate control. In *Proceedings of the 2014 European Control Conference (ECC)*, pages 478–483. IEEE, 2014.

[379] X. Zhang, S. Grammatico, G. Schildbach, P. Goulart, and J. Lygeros. On the sample size of random convex programs with structured dependence on the uncertainty. *Automatica*, 60:182–188, 2015.

[380] X. Zhang, K. Margellos, P. Goulart, and J. Lygeros. Stochastic model predictive control using a combination of randomized and robust optimization. In *Proceedings of the IEEE 52nd Conference on Decision and Control (CDC)*, pages 7740–7745. IEEE, 2013.

[381] Y-J. Zhao, K-F. Wu, Z-J. Wang, L. Zhao, and S-S. Li. Fouling and cleaning of membrane-a literature review. *Journal of Envionmental Sciences*, 12(2):241–251, 2000.

[382] E. Zondervan and B. Roffel. A multi-level approach for the optimization of an ultrafiltration plant processing surface water. *Membrane Water Treatment*, 1(1):61, 2010.

[383] S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198, 2013.

Acknowledgements

Here I would like to thank all the people who contributed directly to the accomplishment of this work, and indirectly to my life.

First of all, I thank Leyla Özkan for her scientific supervision and crucial suggestions always pointing to a possible improvement. Four year ago, she had some doubts whether I would be fitting to the research project yet still provided me the opportunity to pursue this dream. Looking back from today, I have gained a lot from the open discussions that are always available on academic or personal life and dear friendship. Next, I would like to thank my promotor Siep Weiland for his almost-unbounded technical and personal skills on guiding the research activities, keen interest on the subjects that I have enjoyed dealing with and, I believe most importantly, sustaining an environment in which tough questions lead to beautiful observations. His critical eye led to perfection in the way we have expressed our scientific claims, which is one crucial asset for me that is not common even in various levels of academia. My second promoter, Paul Van den Hof, contributed greatly to my personal development, although I always felt that I should have been more proactive about discussing the conducted research to a greater extend with him. Apart from just being the leader that CS group needs, I have observed how to master academic ethics and integrity from Paul. I will always appreciate the support and education I have received from all of you and I hope I will carry them to future.

I am grateful to the members of my defense committee for going through a long and diverse dissertation. Their suggestions highlighting the drawbacks are highly appreciated.

I believe Jobert Ludlage, 'my best Dutch friend', would not be happy with a couple of intricate words and bare thanks; he deserves a nice chat with a smiling friend and no extra obligations. Also, deeply, thank you for your efforts trying to explain me what does robustness mean till the very last months of the research. I would like to thank Marcella Porru and Alejandro Marquez for their dear friendship and constructive discussions happened within the IMPROVISE project. I would also like to present my thanks to Albert van der Padt and Marc Jacobs for the communication and support. Each one of you has been crucial for the successful completion of this dissertation.

Among the staff members in CS group, I would like to specifically thank Will, Diana, Tijs and Lucia for the constant positive attitude. Many grey days became bearable with your smiles. I deeply miss Barbara since the day we had our last chat, she would have made a huge fun of the situation that I put myself at the end of the PhD. I would like to thank Maurice, Duarte and Mircea for the short yet helpful discussions that we have had during the PhD trip. I am greatly in depth for the technical inquires or discussions on the

research problems with Sofie, Mohsin, Edwin and Veaceslav; and tried my best, with great enjoyment, for the ones initiated by Alrianes, Yanin and Amritam. During the wonderful Japan trip I truly met with some of the colleagues and would like to thank Constantijn and Alina, specifically. I would also like to thank Can, Emre, Cansu and Canan for their warm friendship in various conferences and trips.

The presented (or skipped) research subjects are conducted with the unimaginable effort of great B.Sc. or M.Sc. students. I would like to present my warmest regards to Xingang, Elise, Sander, Willem, Bardia, Marco, Ronkai and Dilge for their extensive work and wish them success throughout their careers. Many times I felt I was not dedicating myself for your works in a sufficient manner, yet all of you achieved remarkable results. I would like to extend my thanks also to my collaborator and long-time-missing friend Nikolaos for his sharp observation, to inspire a chapter's worth of research out of a lunch-meeting presentation.

I assume that some of this research would not be existing without the friends at the smoking corner, at Gemini, Potential and Flux. Most of this experience was with Sinan and Tom, I suppose, whom I learned a lot about joyful cycle-trips. Nikolaos, Decabal and Elena can easily find out their substantial contributions in this thesis that flourished within the tight smoking breaks. Though not that a great smoker he is, İlhan has always received my interest for his never ending energy to come up with arguments against today's academia. Furthermore, I have had great experiences with my housemates, Dennis, Harm, Hernan and Lara. Being able to tolerate to my late night habits was already enough as a sign of our dear friendship. I also thank to Ceylan and Başar for allowing their places for me when I was in need.

Furthermore, I would like to thank all my colleagues in Eindhoven for the chats in corridors, the coffee breaks and lunches and the social events, which have made the last four years most joyful. I was surrounded with great friends, more specifically Ruxandra, Dhruv, Pepijn, Ioannis, Guiseppe, and tried my best to reflect this back to them along these years. I am in depth to Kendal for the extensive discussions on the thesis cover and his dedicated work. I also tried my best to sustain a healthy connection with the loved ones in Turkey, namely Haydar, Deniz, Serhat, Yasemin and Cağrı, and would like to thank for their unconditional support. Moreover, I would like to present my deepest gratitude to my family for their acceptance and help, İmge, Handan, Zerrin and Şenol.

Lastly, I would like to express my devotion to Sena, my wife and the person I fully admire. I am thankful to her for every moment we have had together. We have been each others bridge heading to where we are able to find some indescribable warmth in this cold world.

M. Bahadır Saltık,
Eindhoven, The Netherlands,
November 7-17, 2017.

About the Author

M. Bahadır Saltık was born on 06-09-1989 in Lüleburgaz, Turkey. He received his B.Sc. degree in Electrical and Electronics Engineering, with specialization on control area, from the Middle East Technical University, Ankara, Turkey in 2011. In the same year, he started his M.Sc. studies in Systems and Control program at the Eindhoven University of Technology. He completed this program with the graduation project titled “Output Feedback Control of Linear Systems with Event-Triggered Measurements" in Control Systems Technology group at Mechanical Engineering Department in 2013.

Thereafter, he started working towards a Ph.D. degree in Control Systems group of the Electrical Engineering Department at the same institution, within the project ‘Improving Process Operation via Rigorous Models’ (IMPROVISE) under the guidance of dr. L. Özkan, prof.dr.ir. P.M.J. Van den Hof and prof.dr. S. Weiland, and supported by Institute of Sustainable Process Technology (ISPT). The research focus of IMPROVISE project is to elevate the use of rigorous models in online operation of complex processes with a close interest in first principles based modelling, synthesis of model based observer-controller pairs that are acting under uncertain effects. Some of the main results of this research project are presented in this thesis.

Bahadır’s research interests include fundamental aspects of system theory and optimization based control theory. He has a keen interest in algebraic geometry, astronomy and socioeconomic structures in current or historic societies. He is married and lives in Eindhoven.