Chance Constrained Model Predictive Control and Application to SAGD Process

by

Wenhan Shen

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Process Control

Department of Chemical and Materials Engineering

University of Alberta

Abstract

Model Predictive Control (MPC) is widely applied in the process industry nowadays. Chemical processes are corrupted by all kinds of uncertainties, such as measurement noises, disturbances and parameter uncertainties. Without consideration of uncertainties, conventional MPC will cause various problems, for example, violation of constraints and sub-optimal results. Chance constrained MPC (CCMPC) is introduced to generate a safe and optimal control strategy to minimize the effect of uncertainties.

In this thesis, two types of uncertainties in the state space model are considered: system noises and parameter uncertainties. Robust optimization (RO) approximation, a novel method dealing with joint chance constraints, is investigated to solve CCMPC problem. This method leads to results close to the true optimal and is not restricted to certain types of distribution. This work is further applied on the steam assisted gravity drainage (SAGD) process. Constraint violations are greatly reduced by using the RO method.

For system noises, the RO method can be directly applied with the inclusion of uncertainty sets. The type of uncertainty set is selected based on the distribution. Two-layer optimization is proposed, one layer guarantees probability satisfaction and the other layer deals with optimizing the cost. Compared with traditional analytical methods, RO method is not limited to specific distribution and shows better performance in objective function.

For parameter uncertainties, random variables are multiplied with each other, increasing the difficulty to solve the problem. Stochastic tubes help to get rid of multiplicative uncertainty by requiring the tubes to satisfy the constraints. The problem is solved by the RO based tube method. Exponential growth in computation time is avoided. It also guarantees recursive feasibility and stability but results in conservative solutions.

Finally, the RO method is applied to CCMPC of the SAGD process. A state space model is obtained as a proxy model from the production data. The control strategy is calculated based on the proxy model. The developed control strategy is then applied to a reservoir simulator and gives satisfactory results. RO based CCMPC greatly reduces violations due to uncertainties. The SAGD performance is improved by avoiding operating close to critical conditions.

Acknowledgements

It has been an amazing journey over the past two years as a graduate student. I am grateful for the precious time in life when I can concentrate on conducting research and exploring new areas.

I would like to express my sincere gratitude to Professor Biao Huang, Professor Zukui Li and Professor Fraser Forbes for their continuous support and guidance throughout my graduate study and research. It is Dr. Huang who brought me into the Computer Process Control (CPC) group where I deeply enjoyed the research atmosphere. I appreciate his in-depth knowledge, wisdom, enthusiasm and patience. I am thankful for various opportunities in research and project and I could never learn so much without Dr. Huang's help. I enjoyed discussing problems with Dr. Li. His encouragement, insightful comments and expertise in optimization guided me through tough times in research. Dr. Forbes set up an example for me in how to organize graduate life, manage time and face difficulties. I could still remember him teaching me how to present speech, do time management and conduct research.

I would like to give my special thanks to Nabil Magbool Jan, Yuan Yuan, Ruomu Tan, Su Liu and Yanjun Ma for their help and suggestions in my research work. I also want to appreciate my colleagues for their help: Yaojie Lu, Ming Ma, Nima Sammknejad, Alireza Fatehi, Mohammad Rashedi, Anahita Sadeghian, Fadi Ibrahim, Shekhar Sharma, Xiaodong Xu, Rishik Ranjan, Yujia Zhao, Ouyang Wu, Ruoxia Li, Chao Shang, Kangkang Zhang, Zheyuan Liu, Rahul Raveendran, Shabnam Sedghi, Hariprasad Kodamana, Shunyi Zhao, Seraphina Kwak, Hossein Shahandeh, Jayson McAllister, Tianrui An, Mengqi Fang, Gary Wang, Lei Fan, Chaoqun Li, Guoyang Yan and many others from CPC group. I would like to thank my friends who make

a difference in my life: Dian Sheng, Jiarao Huang, Rui Pan, Iohann Oro, Oliver Philipp, Gloria Ohin, Linda London Nguyen, Yongbeom Kim, Xunyuan Yin, An Zhang, Weiran Wang, Li Lin, Leo Li, Lang Liu, Boon Chew, Zhankun Xi, Yaqing Li, Yongzan Liu, Shuo Zhang, Jing Zhang, Tianbo Liu, Shuning Li, Juan Li, Hanbo Li, Nan Shang, Shiqi Lai, Ying Xiong, Qi Zeng, Alanna Spicer, Dean Spicer, Edythe Tham, Nils Reuter, Zhijiang Lou, Hemant Bansal, Khushaal Popli, Di Zhu and many others.

Financial support from NSERC and Alberta Innovates Technology Futures is greatly appreciated.

Lastly, I want to give thanks to my parents, Zengxin Shen and Ling Wang, for their endless love, support and encouragement. It is family who is always there to hold my hands and support me in tough times.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

MPC is widely used in the process industries. MPC has advantages in taking care of multiple inputs and outputs, incorporating constraints and optimizing system performance. In conventional MPC, however, uncertainties are not taken into account. This raises various issues, including in-feasibility and sub-optimality. Robust MPC (RMPC) and stochastic MPC (SMPC) are the two main approaches dealing with uncertainty problem in MPC. RMPC assumes that uncertainties lie in a bounded set and gives control actions based on the worst case [1], [2]. SMPC, however, incorporates the distribution of uncertainty. Constraint types in SMPC include chance constraints, maximum violation and expectation on constraints. Chance constrained MPC (CCMPC), includes a general case and gives user freedom to tune the risk. It exploits probability guarantees which give a trade-off between control performance and fulfillment of constraints [3]. In terms of SMPC approaches, three categories can be classified: stochastic programming based approach [4], [5], [6], [7], [8], [9], stochastic tube approaches [10], [11], and affine parameterization of control policy [12], [13]. Stochastic programming approaches, including analytical approaches [4], [5], [6], [7] and sampling methods [8], [9], are naturally adopted to CCMPC. By separating system states into deterministic components and random components, stochastic tube approaches help solve the problem of intractability generally existed in CCMPC. Affine parameterization in feedback control policy offers better solution

than open-loop method. The affine structure helps reducing calculation effort to a limited amount. We will focus on all three methods in the thesis. By the structure of chance constraints, it can be separated into individual chance constrained problem [10], [11] and joint chance constrained problem [6]. Individual chance constrained problem is much easier to solve and can be included in the joint chance constrained problem. We focus on solving joint chance constrained problem in the thesis.

CCMPC problems do not have a general solution. Certain assumptions have to be made (i.e. Gaussian distribution) for a lot of methods [14], [15]. In robust optimization (RO), constraints are guaranteed regardless of the uncertainty distribution, as long as the random variables are in a bounded set. The advantage it shows in dealing with arbitrary distributions brings us to utilize the RO method in CCMPC.

Modeling uncertainty comes not only in additive uncertainty but also in multiplicative uncertainty. RO based method can deal with additive noise in a state space model. However, multiplicative uncertainty induced by unknown parameters poses difficulty in using the RO based method. Propagation of uncertainty over the prediction horizon introduces significant conservativeness and computational complexity [16]. By defining regions which include the predicted states and propagating these regions from one time step to the next via one-step-ahead calculation, this problem can be solved. Stochastic tube approaches are thus introduced [10], [11], [16]. Tube methods also provide guarantees for recursive feasibility and stability [3]. We use stochastic tubes to convert multiplicative uncertainty into forms which can be solved by RO method.

In SAGD process, PID controllers are widely used for sub-cool control. However, advanced process control methods (i.e. MPC) are still rarely used. One important reason is that we do not have an accurate model describing SAGD behavior. Reservoir simulators allow us to do various tests and get rich data for modeling. This leads to an interest of building proxy models and applying advanced CCMPC in SAGD process.

1.2 Thesis contributions

The main contribution of this thesis is developing robust optimization approximation based chance constrained model predictive control (RO based CCMPC) method, which is not restricted by certain types of uncertainty distributions compared to other methods. The method can be applied to state space models with system noise or parametric uncertainty. Specifically, the contributions of this thesis are summarized as follows:

1. Developed an RO based CCMPC method for state space model with system noise;
2. Combined stochastic tube MPC with RO based methods to solve parametric uncertainty problem;
3. Built data-driven proxy models from simulation of SAGD process;
4. Implemented RO based CCMPC in the reservoir simulator of SAGD process.

1.3 Thesis outline

The layout of this thesis is as follows:

Chapter 2 proposes RO based CCMPC method and compares it with other methods. A closed-loop structure is introduced to improve RO method performance. RO method is tested in two case studies: CSTR and hydrodesulfuration process.

Chapter 3 deals with parameter uncertainty in CCMPC problems. Stochastic tube methods are introduced to convert chance constraints into the forms which can be solved by RO based CCMPC method.

Chapter 4 aims to implement RO based CCMPC in SAGD process. A state space proxy model is built from simulation data. The control input from RO based CCMPC method is calculated by treating the proxy model as the real process model. The control strategy is tested on the reservoir simulator.

Chapter 5 concludes the thesis.

Chapter 2

Chance Constrained Model Predictive Control in Dealing with System Noises

2.1 Introduction

A critical challenge for planning and control in the process industry is the presence of uncertainty. Due to imperfect knowledge of the process, disturbances and measurement noises, uncertainty exists in both model parameters and system states. Various issues arise due to the uncertainties, including poor performance, instability or violation of the constraints [17], [18], [19]. The focus in our work is on the restriction of violations while minimizing the cost in the presence of uncertainties.

Generally, there are two kinds of constraints: hard constraints and soft constraints. Hard constraints are required to be satisfied arbitrarily. Examples are those constraints related to operational safety, product quality or equipment capacity. For example, the flow rate in a pipeline cannot exceed its capacity. These constraints are usually critical requirements which have zero tolerance for the violation. Soft constraints are not required to be guaranteed 100%, which means that the system can tolerate a certain degree of violations. Examples are those related to process performance, like temperature in building climate control [13]. The temperature is supposed to be kept in the comfort zone, but not required to be always satisfied. As soft constraint is closely related to the performance, it is usually user-defined.

For the soft constraints, based on measures of the constraint violation, it can be separated into three groups: constraints on the expectation, constraints on the worst case and probabilistic constraints.

$$\mathbb{E}[g(x,\xi)] \leq 0$$

$$\sup_{\xi} g(x,\xi) \leq 0$$

$$\text{Prob}(g(x,\xi) \leq 0) \geq 1-\delta$$

where, x is the decision variable and ξ is the random variable.

These three measures have different purposes in dealing with soft constraints. It is interesting to note that in probabilistic constraints (δ is called violation risk), if $\delta = 0.5$, the constraints on expectation hold. If $\delta = 0$, the constraints on worst case hold. The probabilistic constraints define a more general case. Probabilistic constraint is usually referred to as chance constraint in literature.

Over the nominal operation which ignores uncertainties, the inclusion of chance constraints provides several advantages:

1. A guarantee on the satisfaction of constraints under uncertainties: given the tolerance on violation, we can guarantee a probability that the constraints will be satisfied;

2. Better performance: in optimizing the profit, it usually happens when optimal point is close to or on the boundary of the constraints. By tuning the risk of violation, the user can freely determine whether to achieve better performance or to have less risk.

Chance constrained problems are difficult to solve because it is hard to check feasibility of the chance constraints and most of the feasible regions are non-convex [7]. Special cases which can be equivalently converted into convex problems can be found in [14], [20]. Generally speaking, most chance constrained problems are solved through approximation methods. There are two main categories: sampling based methods and analytical approximations [6]. The sampling methods include sample average approximation which requires large number of samples and computational effort [9], [6], and

scenario approach which has great conservativeness [20]. Analytical approximations try to approximate the chance constraints with deterministic formulations. We will discuss in details about the analytical methods, which include ellipsoidal relaxation [6], iterative risk allocation [5] and robust optimization approximation method [7].

The process model in the thesis is discrete-time state space model. In this chapter only system noises added to the states are considered as uncertainties.

This chapter is organized as follows: preliminaries of CCMPC are given in Section 2.2, problem formulation is described in Section 2.3, Section 2.4 presents different methods for CCMPC, and lastly in Section 2.5, two case studies are given.

2.2 Preliminaries

Definition 2.2.1. (Joint Chance Constraint) *A joint chance constraint is defined when several constraints have to be jointly fulfilled with probability* $1-\delta$*:*

$$\text{Prob}(\psi_j(x,\xi) \leq 0,\ \forall j) \geq 1-\delta \tag{2.1}$$

Definition 2.2.2. (Individual Chance Constraint) *An individual chance constraint is defined when a single constraint is required to be fulfilled with probability* $1-\delta$*:*

$$\text{Prob}(\psi(x,\xi) \leq 0) \geq 1-\delta \tag{2.2}$$

Definition 2.2.3. (Active Constraint)

$$\min_x f(x)$$
$$\textit{subject to } g_i(x) \leq 0,\ i = 1,\ldots,m$$

The i_{th} *constraint is called active constraint at optimal solution* x^* *if* $g_i(x^*) = 0$*.*

Theorem 2.2.1. (Boole's Inequality) *For a countable set of events* $E_1, E_2, E_3, \cdots$*, the probability of the union of these events is smaller than the summation of the probability of each individual one.*

$$Prob\left(\bigcup_i E_i\right) \leq \sum_i P(E_i)$$

2.3 Problem Formulation

2.3.1 System Description

The discrete state space model is described as:

$$x_{k+1} = Ax_k + Bu_k + W\omega_k \tag{2.3}$$

with current time step $k \in \mathbb{N}$, system state $x_k \in \mathbb{R}^n$, control input $u_k \in \mathbb{R}^m$, and process noise $\omega_k \in \mathbb{R}^q$ is a stationary process which follows a certain distribution.

Model predictive control, as its name suggests, includes two parts: prediction and control. At each time instant k, a measurement of x_k is taken, based on the model, and the prediction of state in N steps is formulated. Prediction cost is minimized under system constraints. Then only the input calculated at first step is applied to the system. The prediction and control will be repeated at next time instant $k+1$.

The objective function is defined as:

$$J^*(x_k) = \min_{u_{k|k},\ldots,u_{k+N-1|k}} V_f(x_{k+N|k}) + \sum_{i=0}^{N-1} l_k(x_{k+i|k}, u_{k+i|k}) \tag{2.4}$$

The constraints are defined as:

$$l_h \leq u_{k+i-1|k} \leq g_h, \qquad 1 \leq i \leq N \tag{2.5}$$

$$Prob(l_s \leq x_{k+i|k} \leq g_s) \geq 1 - \delta_i, \qquad 1 \leq i \leq N \tag{2.6}$$

where k is current time instant, $k+i$ is future time instant, N is prediction horizon, $u_{k+i-1|k}$ is future input at time $k+i-1$, $x_{k+i|k}$ is predicted state at time $k+i$, $l_k : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_{+0}$ is stage cost, and $V_f : \mathbb{R}^n \to \mathbb{R}_{+0}$ is terminal cost, δ_i is allowed risk for violating the constraint at time instant $k+i$, l_h and g_h are lower and upper bound of the hard constraint respectively, l_s and g_s are lower and upper bound of the soft constraint respectively.

With open-loop policy, the inputs are given arbitrarily. The inputs are required to satisfy the hard constraints. Future states are required to satisfy the soft constraints.

After introducing MPC, the predicted states will be reformulated in a vector form based on 2.3:

$$\mathbf{x} = Gx_k + H\mathbf{u} + P\boldsymbol{\xi}$$

where,

$$G = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix}, H = \begin{bmatrix} B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix}, P = \begin{bmatrix} W & 0 & \cdots & 0 \\ AW & W & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}W & A^{N-2}W & \cdots & W \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_{k+1|k} \\ x_{k+2|k} \\ \vdots \\ x_{k+N|k} \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+N-1} \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \omega_k \\ \omega_{k+1} \\ \vdots \\ \omega_{k+N-1} \end{bmatrix}$$

The constraints are written in vector form based on 2.5, 2.6:

$$\mathbf{l}_h \leq \mathbf{u} \leq \mathbf{g}_h$$

$$\text{Prob}(l_s \leq R_i\mathbf{x} \leq g_s) \geq 1 - \delta_i, \qquad 1 \leq i \leq N$$

where,

$$R_i = \begin{bmatrix} 0_{n\times n} & 0_{n\times n} & \cdots & I_n & \cdots & 0_{n\times n} & 0_{n\times n} \end{bmatrix}_{n\times nN}, \mathbf{l}_h = \begin{bmatrix} l_h \\ \vdots \\ l_h \end{bmatrix}_{mN\times 1}, \mathbf{g}_h = \begin{bmatrix} g_h \\ \vdots \\ g_h \end{bmatrix}_{mN\times 1}$$

In R_i, I_n is $n \times n$ identity matrix in i_{th} block.

The objective function in 2.4 is usually formulated as a quadratic form of cost:

$$\min_{\mathbf{u}} \mathbb{E}[\mathbf{x}^T Q_{obj}\mathbf{x} + \mathbf{u}^T R_{obj}\mathbf{u}]$$

Denote: $Q_{qua} = H^T Q_{obj} H + R_{obj}, Q_{lin} = x(k)^T G^T Q_{obj} H, c = x(k)^T G^T Q_{obj} G x(k) + trace(P^T Q_{obj} P)$. Thus we have

$$\min_{\mathbf{u}}[\mathbf{u}^T Q_{qua}\mathbf{u} + 2Q_{lin}\mathbf{u} + c] \tag{2.7}$$

Then the constraints can be equivalently written into:

$$\text{Prob}(S_i(Gx(k) + H\mathbf{u} + P\boldsymbol{\xi}) \leq \mathbf{s}) \geq 1 - \delta_i, \qquad 1 \leq i \leq N \tag{2.8}$$

$$\mathbf{l}_h \leq \mathbf{u} \leq \mathbf{g}_h \tag{2.9}$$

Where,

$$S_i = \begin{bmatrix} 0_{2n\times n} & 0_{2n\times n} & \cdots & \begin{bmatrix} I_n \\ -I_n \end{bmatrix} & \cdots & 0_{2n\times n} & 0_{2n\times n} \end{bmatrix}, \mathbf{s} = \begin{bmatrix} g_s \\ -l_s \end{bmatrix}$$

In S_i, $\begin{bmatrix} I_n \\ -I_n \end{bmatrix}$ is $2n \times n$ matrix in i_{th} block.

2.3.2 Chance Constrained Problem

From 2.8, we know that there are N joint chance constraints. As the terms inside the probability equation 2.8 are in a linear form, we can further simplify it into the following:

$$\text{Prob}(\mathbf{y}_{0i} + \mathbf{y}_i\boldsymbol{\xi} \leq \mathbf{0}) \geq 1 - \delta_i, \qquad \forall i \tag{2.10}$$

where,

$$\mathbf{y}_{0i} = S_i(Gx(k) + H\mathbf{u}) - \mathbf{s}$$
$$\mathbf{y}_i = S_iP$$

We can also write the joint chance constraint in an explicit way:

$$\text{Prob}(\mathbf{y}_{0i}^j + \mathbf{y}_i^j\boldsymbol{\xi} \leq 0, j = 1, 2, \cdots, 2n) \geq 1 - \delta_i, \qquad \forall i \tag{2.11}$$

where $\mathbf{y}_{0i}^j, \mathbf{y}_i^j$ denotes j_{th} row of $\mathbf{y}_{0i}, \mathbf{y}_i$.

$$\mathbf{y}_{0i}^j = F_j\mathbf{y}_{0i}, \mathbf{y}_i^j = F_j\mathbf{y}_i \qquad 1 \leq j \leq 2n$$

$$F_j = \begin{bmatrix} 0 & 0 & \cdots & 1_{j_{th}} & \cdots & 0 & 0 \end{bmatrix}_{1\times 2n}$$

Thus 2.10 is equivalent to:

$$\text{Prob}(\mathbf{y}_{0i}^j + \mathbf{y}_i^j\boldsymbol{\xi} \leq 0, \quad \forall j) \geq 1 - \delta_i, \qquad \forall i \tag{2.12}$$

where,

$$\mathbf{y}_{0i}^j = F_jS_i(Gx(k) + H\mathbf{u}) - F_j\mathbf{s}$$
$$\mathbf{y}_i^j = F_jS_iP$$

2.4 Chance Constrained MPC Methods

As we have introduced before, there is no closed-form solution of chance constrained problem generally. Most methods deal with chance constraint by approximation, and formulate it into a deterministic problem with conservatism introduced [6], [13], [7].

2.4.1 Ellipsoidal Relaxation Method

Gaussian distribution is most widely used in statistics. By estimating the mean value and covariance of the process noise from collected data, most uncertainties in real process can be approximated by Gaussian distribution. Assume the process noise in 2.3 follows Gaussian distribution $\boldsymbol{\xi} \sim N(\mu, \Sigma_{\boldsymbol{\xi}})$. In order to evaluate the chance constraints, one needs to calculate the integral of multivariate Gaussian distribution. However, there is no analytical form of multivariate Gaussian integral over arbitrary region. But there does exist analytical solution over ellipsoidal region. Ellipsoidal method can relax the problem into a tractable form.

The ellipsoidal region is defined as the set $\mathbb{S}(\boldsymbol{\xi}) = \{\boldsymbol{\xi} | (\boldsymbol{\xi} - \mu)^T \Sigma_{\boldsymbol{\xi}}^{-1} (\boldsymbol{\xi} - \mu) \leq r^2\}$. The probability under the ellipsoidal region is:

$$P_{\boldsymbol{\xi}} = \frac{1}{\sqrt{|\Sigma_{\boldsymbol{\xi}}|(2\pi)^{n_{\boldsymbol{\xi}}}}} \int\limits_{(\boldsymbol{\xi}-\mu)^T \Sigma_{\boldsymbol{\xi}}^{-1} (\boldsymbol{\xi}-\mu) \leq r^2} \exp\left(-\frac{1}{2} (\boldsymbol{\xi} - \mu)^T \Sigma_{\boldsymbol{\xi}}^{-1} (\boldsymbol{\xi} - \mu) \right) d\boldsymbol{\xi} \tag{2.13}$$

From [21], we know that the integral is equivalent to one-dimensional integral with χ^2 distribution:

$$P_{\boldsymbol{\xi}} = \frac{1}{2^{\frac{n_{\boldsymbol{\xi}}}{2}} \Gamma(\frac{n_{\boldsymbol{\xi}}}{2})} \int_0^{r^2} \chi^{n_{\boldsymbol{\xi}}/2-1} \exp(-\chi/2) d\chi$$

By letting $P_{\boldsymbol{\xi}} = 1 - \delta_i$, the confidence radius r can be calculated from χ^2 distribution table. Thus the probability requirement in 2.10 can be replaced by requiring $(\boldsymbol{\xi} - \mu)^T \Sigma_{\boldsymbol{\xi}}^{-1} (\boldsymbol{\xi} - \mu) \leq r^2$. If for $\forall \boldsymbol{\xi} \in \mathbb{S}(\boldsymbol{\xi})$, the inequality $\mathbf{y}_{0i} + \mathbf{y}_i \boldsymbol{\xi} \leq \mathbf{0}$ is satisfied, then the chance constraint 2.10 is certainly satisfied.

It is equivalent to satisfying the following condition for the above inequality to

hold:

$$\sup_{\boldsymbol{\xi}\in\mathbb{S}(\boldsymbol{\xi})} \{\mathbf{y}_{0i} + \mathbf{y}_i\boldsymbol{\xi}\} \leq \mathbf{0}$$

To calculate the upper bound of the above equation, we need to use Cauchy-Schwarz Inequality: $|\langle \mathbf{u}, \mathbf{v}\rangle|^2 \leq \langle \mathbf{u}, \mathbf{u}\rangle \cdot \langle \mathbf{v}, \mathbf{v}\rangle$. By letting $\sigma = \boldsymbol{\xi} - \mu$, we have $\mathbf{y}_i^j\sigma \leq \sqrt{\mathbf{y}_i^j\Sigma_{\boldsymbol{\xi}}\mathbf{y}_i^{jT}} \cdot \sqrt{\sigma^T\Sigma_{\boldsymbol{\xi}}^{-1}\sigma}$. As we know $\sigma^T\Sigma_{\boldsymbol{\xi}}^{-1}\sigma \leq r^2$, we have $\mathbf{y}_i^j\sigma \leq r\sqrt{\mathbf{y}_i^j\Sigma_{\boldsymbol{\xi}}\mathbf{y}_i^{jT}}$.

Thus, by doing the above approximations, the deterministic constraints are obtained:

$$\mathbf{y}_{0i}^j + \mathbf{y}_i^j\mu + r\sqrt{\mathbf{y}_i^j\Sigma_{\boldsymbol{\xi}}\mathbf{y}_i^{jT}} \leq 0, \qquad \forall j \tag{2.14}$$

To summarize, the ellipsoidal method has two steps:

1. Calculate the equivalent radius r from χ^2 distribution based on the probability requirement $1 - \delta_i$ and uncertainty dimension $n_{\boldsymbol{\xi}}$;

2. Formulate the approximate deterministic constraints 2.14

No slack variables are introduced in the formulated constraints. As decision variables $\mathbf{u}$ are in $\mathbf{y}_{0i}^j$, the problem falls in the form of linear programming, which is solvable in polynomial time. However, the guarantee of probability satisfaction does not take in consideration of the structure of constraints, which will yield a rather conservative result. As the dimension of uncertainty $\boldsymbol{\xi}$ increases, the confidence radius r will increase causing the conservatism to increase. In MPC problem, the dimension of $\boldsymbol{\xi}$ is usually large due to the long prediction horizon, thus the conservatism in ellipsoidal relaxation method is severe.

2.4.2 Boole's Inequality and Iterative Risk Allocation

Similar to Ellipsoidal Relaxation, the Boole's method also assumes the process noise can be approximated by Gaussian distribution $\boldsymbol{\xi} \sim N(\mu, \Sigma_{\boldsymbol{\xi}})$. Before introducing Boole's method, we first take a look at how to solve individual chance constraint with Gaussian distribution.

Assume the individual risk is given as: δ_{ij}. We have the following formulation of individual chance constraint:

$$\text{Prob}(\mathbf{y}_{0i}^j + \mathbf{y}_i^j \boldsymbol{\xi} \leq 0) \geq 1 - \delta_{ij} \tag{2.15}$$

It is equivalent to forming it into constraint of violation restriction:

$$\text{Prob}(\mathbf{y}_{0i}^j + \mathbf{y}_i^j \boldsymbol{\xi} \geq 0) \leq \delta_{ij} \tag{2.16}$$

To calculate the univariate integral in 2.16, the random variable $\mathbf{y}_i^j \boldsymbol{\xi}$ needs to be normalized first. We have:

$$\boldsymbol{\xi}_{STD} = \frac{\mathbf{y}_i^j \boldsymbol{\xi} - \mathbf{y}_i^j \mu}{\sqrt{\mathbf{y}_i^j \sigma_{\boldsymbol{\xi}} \mathbf{y}_i^{jT}}} \sim N(0,1)$$

Thus equation 2.16 can be formulated as:

$$\text{Prob}(\boldsymbol{\xi}_{STD} \geq \frac{-\mathbf{y}_{0i}^j - \mathbf{y}_i^j \mu}{\sqrt{\mathbf{y}_i^j \sigma_{\boldsymbol{\xi}} \mathbf{y}_i^{jT}}}) \leq \delta_{ij}$$

As we know the cumulative distribution function (CDF) of the standard normal distribution is denoted as: $\Phi(x)$.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

We have,

$$1 - \Phi(x) = \int_x^{\infty} \boldsymbol{\xi}_{STD} d\boldsymbol{\xi} \leq \delta_{ij}, \quad x = \frac{-\mathbf{y}_{0i}^j - \mathbf{y}_i^j \mu}{\sqrt{\mathbf{y}_i^j \sigma_{\boldsymbol{\xi}} \mathbf{y}_i^{jT}}}$$

As $\Phi(x)$ is a monotonically increasing function, the above inequalities can be formulated as:

$$\mathbf{y}_{0i}^j + \mathbf{y}_i^j \mu + \Phi^{-1}(1 - \delta_{ij}) \sqrt{\mathbf{y}_i^j \Sigma_{\boldsymbol{\xi}} \mathbf{y}_i^{jT}} \leq 0 \tag{2.17}$$

Thus, we know that for individual chance constraint 2.16 with Gaussian distribution, it is equivalent to the second-order cone constraint 2.17. This becomes rather easy to solve if we can covert the joint chance constraint into individual chance constraints.

The joint chance constraint 2.11 is converted to the form of probability of violation here.

$$\text{Prob}\Big(\{\mathbf{y}_{0i}^1+\mathbf{y}_i^1\boldsymbol{\xi}\geq 0\}\cup\ldots\cup\{\mathbf{y}_{0i}^j+\mathbf{y}_i^j\boldsymbol{\xi}\geq 0\}\cup\ldots\cup\{\mathbf{y}_{0i}^{2n}+\mathbf{y}_i^{2n}\boldsymbol{\xi}\geq 0\}\Big)\leq\delta_i \quad (2.18)$$

As a matter of simplification, we use $\text{Prob}(\mathbf{y}_{0i}^j+\mathbf{y}_i^j\boldsymbol{\xi}\geq 0, j=1\cup 2\cup\cdots\cup 2n)$ to denote the left-hand side of 2.18.

Boole's inequality is used to bound the probability violation. So if we require the summation of the probability of violation smaller than the required risk, then the joint probability of violation must be satisfied.

$$\text{Prob}(\mathbf{y}_{0i}^j+\mathbf{y}_i^j\boldsymbol{\xi}\geq 0, j=1\cup 2\cup\cdots\cup 2n)\leq\sum_{j=1}^{n}\text{Prob}(\mathbf{y}_{0i}^j+\mathbf{y}_i^j\boldsymbol{\xi}\geq 0)\leq\delta_i \quad (2.19)$$

Equivalently, we can write:

$$\begin{aligned}\text{Prob}(\mathbf{y}_{0i}^j+\mathbf{y}_i^j\boldsymbol{\xi}\geq 0)&\leq\delta_{ij}\\ \sum_{j=1}^{2n}\delta_{ij}&\leq\delta_i\end{aligned} \quad (2.20)$$

where, δ_{ij} is the individual risk.

But still, there remains one question of how to allocate the joint risk δ_i on each individual chance constraint. A straight forward idea is to average the risk on all the individual chance constraints. By averaging the risk, $\delta_{ij}=\dfrac{\delta_i}{n}$, we have:

$$\mathbf{y}_{0i}^j+\mathbf{y}_i^j\mu+\Phi^{-1}(1-\frac{\delta_i}{n})\sqrt{\mathbf{y}_i^j\Sigma_{\boldsymbol{\xi}}\mathbf{y}_i^{jT}}\leq 0\quad\forall i,j \quad (2.21)$$

However, we know there are some constraints which are more likely to have violations and some constraints which tend to have less or zero chance to violate. Having equal risk for them will lead to conservative solutions. Therefor, we need better allocation method to yield better optimization result.

The risk allocation problem can then be formulated as:

$$\min_{\mathbf{U},\delta}E[\mathbf{u}^TQ_{qua}\mathbf{u}+2Q_{lin}\mathbf{u}+c] \quad (2.22)$$

subject to

$$\mathbf{l}_h \leq \mathbf{u} \leq \mathbf{u}_h \tag{2.23}$$

$$\mathbf{y}_{0i}^j + \mathbf{y}_i^j \mu + \Phi^{-1}(1-\delta_{ij})\sqrt{\mathbf{y}_i^j \Sigma_{\boldsymbol{\xi}} \mathbf{y}_i^{jT}} \leq 0 \quad \forall i,j \tag{2.24}$$

$$\sum_{j=1}^{2n} \delta_{ij} \leq \delta_i \quad \forall i \tag{2.25}$$

$$\delta_{ij} \geq 0 \quad \forall i,j \tag{2.26}$$

$$\mathbf{y}_{0i}^j = F_j S_i (G x_k + H\mathbf{u}) - F_j \mathbf{s} \quad \forall i,j \tag{2.27}$$

$$\mathbf{y}_i^j = F_j S_i P \quad \forall i,j \tag{2.28}$$

We can prove that the above problem is a convex optimization problem. It is known that convex optimization is defined on minimizing convex or maximizing concave functions over convex sets. The constraints 2.23, 2.25 - 2.28 are all linear constraints. It leaves to us to prove 2.24 is convex.

Denote $f_{ij}(\mathbf{u}) = \mathbf{y}_{0i}^j + \mathbf{y}_i^j \mu$ and $g_{ij}(\delta_{ij}) = \Phi^{-1}(1-\delta_{ij})\sqrt{\mathbf{y}_i^j \Sigma_{\boldsymbol{\xi}} \mathbf{y}_i^{jT}}$ and let $\mathbf{u}^* = \lambda \mathbf{u}_1 + (1-\lambda)\mathbf{u}_2$ and $\delta_{ij}^* = \lambda \delta_{ij}^1 + (1-\lambda)\delta_{ij}^2$. It is easy to know that $f_{ij}(\mathbf{u}^*) = \lambda f_{ij}(\mathbf{u}_1) + (1-\lambda) f_{ij}(\mathbf{u}_2)$. Also we know that $\Phi^{-1}(x)$ is convex function for $x > 0.5$. Affine function preserves convexity. So $\Phi^{-1}(1-x)$ is convex function for $x < 0.5$. Thus,

$$\lambda(\Phi^{-1}(1-\delta_{ij}^1)) + (1-\lambda)(\Phi^{-1}(1-\delta_{ij}^2)) \geq \Phi^{-1}(1-\delta_{ij}^*)$$

So we have,

$$g_{ij}(\delta_{ij}^*) \leq \lambda g_{ij}(\delta_{ij}^1) + (1-\lambda) g_{ij}(\delta_{ij}^2)$$

Thus, if

$$f_{ij}(\mathbf{u}_1) + g_{ij}(\delta_{ij}^1) \leq 0, \quad f_{ij}(\mathbf{u}_2) + g_{ij}(\delta_{ij}^2) \leq 0$$

we know,

$$f_{ij}(\mathbf{u}^*) + g_{ij}(\delta_{ij}^*) \leq \lambda\Big(f_{ij}(\mathbf{u}_1) + g_{ij}(\delta_{ij}^1)\Big) + (1-\lambda)\Big(f_{ij}(\mathbf{u}_2) + g_{ij}(\delta_{ij}^2)\Big) \leq 0$$

But the problem exists in the fact there is no analytical form of inverse cumulative distribution function (icdf) Φ^{-1}, which makes the problem hard to solve.

Here we introduce a sub-optimal method to solve it.

Knowing that, by loosening active constraint (increasing the feasible region), we have more chance to search for a better result. Iterative risk allocation (IRA) follows the idea that, by giving more risk from inactive constraints to active constraints, a better result will be obtained.

The method includes two parts: searching for better allocation of risk and solving the optimization problem with the allocated risk. A two-stage optimization is formulated here:

Upper-stage Optimization (Risk Allocation):

$$\begin{aligned} \min_{\delta} \quad & J^*(\delta) \\ \text{s.t.} \quad & \sum_{j=1}^{2n} \delta_{ij} \leq \delta \quad \forall i \\ & \delta_{ij} \geq 0 \quad \forall i, j \\ & \delta \in \{\delta | \exists \mathbf{u} \text{ that satisfies } 2.23, 2.24, 2.27, 2.28\} \end{aligned} \tag{2.29}$$

Lower-stage Optimization (Individual Chance Constraints):

$$\begin{aligned} J^*(\delta) = \min_{\mathbf{u}} \quad & E[\mathbf{u}^T Q_{qua} \mathbf{u} + 2 Q_{lin} \mathbf{u} + c] \\ \text{s.t.} \quad & \mathbf{u} \text{ satisfies } 2.23, 2.24, 2.27, 2.28 \end{aligned} \tag{2.30}$$

The lower-stage optimization obtains the allocated risk from the upper-stage and solves individual chance constrained problem. Based on the result calculated from lower-stage optimization, upper-stage will continuously allocate risk from inactive constraints to active constraints. We will illustrate further how to allocate the risk.

Let δ and δ' both be risk assignments, $\mathcal{R}(\delta)$ and $\mathcal{R}(\delta')$ be the corresponding feasible region of $\mathbf{u}$ defined by 2.23, 2.24, 2.27 and 2.28. For most applicable cases, we can assume $\delta_i \leq 0.5$, thus we have $\delta_{ij} \leq 0.5$. If $\delta_{ij} \leq \delta'_{ij} \leq 0.5$, then $0 \leq \Phi^{-1}(1 - \delta'_{ij}) \leq \Phi^{-1}(1 - \delta_{ij})$. Noticing the constraint 2.24, we know that $\mathcal{R}(\delta) \subseteq \mathcal{R}(\delta')$. Therefore we know $J^*(\delta) \geq J^*(\delta')$. So the value of J^* decreases or keeps the same as δ increases.

Also, another property is that, for inactive constraint 2.24 with δ_{ij}, we can always

Table 2.1: *Algorithm for Iterative Risk Allocation*

Algorithm for Iterative Risk Allocation
1. Initial risk given: $\delta_{(0)ij} = \dfrac{\delta_i}{2n} \quad \forall i, j,\ n = 0;$
2. Solve lower-stage optimization 2.30 with $\delta_{(n)}$, the optimal result J^*;
3. Calculate the number of active individual chance constraint of joint chance constraint i: $N_{active,i}\ \forall i$;
4. If $\sum_{i=1}^{N} N_{active,i} = 0$ or $2nN$, then stop the iteration with U^*, δ^*, J^*;
5. Let $\delta'_n = \delta_n$. $\forall i, j$, if the constraint 2.24 is inactive, tighten associated risk: $$\delta'_{(n)ij} = \alpha\delta_{(n)ij} + (1-\alpha)\left(1 - \Phi(-\frac{\mathbf{y}_{0i}^{j} + \mathbf{y}_i^j \mu}{\sqrt{\mathbf{y}_i^j \Sigma_{\boldsymbol{\xi}} \mathbf{y}_i^{jT}}})\right)$$
6. Calculate the residual risk as: $\delta_{residual,i} = \delta_{(n)i} - \sum_{j=1}^{2n} \delta'_{(n)ij} \quad \forall i$
7. Let $\delta_{(n+1)} = \delta'_{(n)}$. $\forall i, j$, if the constraint 2.24 is active, loosen the associated risk: $$\delta_{(n+1)ij} = \delta_{(n)ij} + \delta_{residual,i}/N_{active,i}\ \forall i, j$$
8. Let $n = n + 1$. For $n > 1$, check if $\lvert J^* - J_{prev} * \rvert < \epsilon$, if so, stop the iteration with U^*, δ^*, J^*. If not, continue with the 2_{nd} step.

find another δ'_{ij} which satisfies,

$$\delta_{ij} > \delta'_{ij} \tag{2.31}$$

$$\mathbf{y}_{0i}^{j\,*} + \mathbf{y}_i^{j*}\mu + \Phi^{-1}(1-\delta_{ij})\sqrt{\mathbf{y}_i^{j*}\Sigma_{\boldsymbol{\xi}}\mathbf{y}_i^{j*T}} < \mathbf{y}_{0i}^{j\,*} + \mathbf{y}_i^{j*}\mu + \Phi^{-1}(1-\delta'_{ij})\sqrt{\mathbf{y}_i^{j*}\Sigma_{\boldsymbol{\xi}}\mathbf{y}_i^{j*T}} < 0 \quad \forall i, j \tag{2.32}$$

where $\mathbf{y}_{0i}^{j\,*}$ and $\mathbf{y}_i^{j*}$ are optimal values from $\mathbf{u}^*$ given δ.

It means if we can find any δ'_{ij} which satisfies the above inequalities, the optimal value does not change while some extra risk $\delta_{ij} - \delta'_{ij}$ has been freed. The extra risk can be allocated to the active constraints. The IRA algorithm is listed in Table 2.1.

In Fig 2.1, it shows that the objective function value is getting smaller in IRA

iteration by iteration.

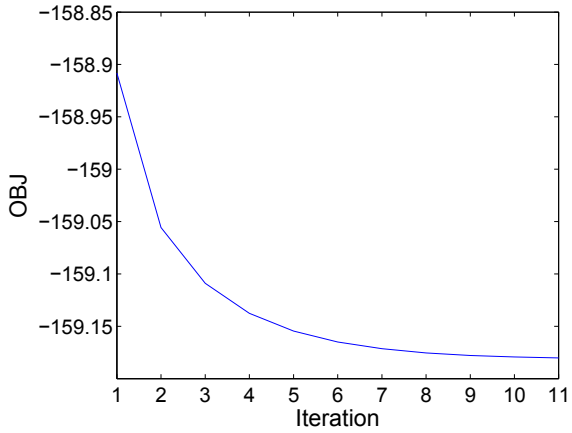


Figure 2.1: *Objective Value after Each Iteration*

In most cases, the IRA method is less conservative than the ellipsoidal relaxation because the constraint structure is taken into consideration together with the probability satisfaction. The lower-stage optimization is second-order cone programming. The total calculation time depends on the upper stage: the number of iterations it takes for the objective value to converge. It depends on case to case, but in general, the IRA method is moderate in time consumption in solving CCMPC.

For the conservatism, it mainly comes from the Boole's inequality. We will illustrate first from the example with two random events E_1 and E_2. The probability of the union of these two events equals: $\text{Prob}(E_1 \cup E_2) = \text{Prob}(E_1) + \text{Prob}(E_2) - \text{Prob}(E_1 \cap E_2)$. The Boole's inequality upper bound the joint chance constraint with $\text{Prob}(E_1 \cup E_2) \leq \text{Prob}(E_1) + \text{Prob}(E_2)$. The conservatism comes from the part where both E_1, E_2 happens, which is $\text{Prob}(E_1 \cap E_2)$. An illustration is shown in the Fig 2.2.

However, to quantify the conservatism is hard. One reason is that it is difficult to calculate $\text{Prob}(E_i \cap \ldots \cap E_j)$, the other reason is that multiple cases will increase the difficulty in the mathematical formulation of the conservatism in probability. To help quantify the conservatism, we assume the random events $E_1, \ldots, E_j, \ldots, E_{2n}$ is independent of each other, with the true probability of violation δ_{ij}. It happens in the case when each state has its independent source of disturbance or noise, which

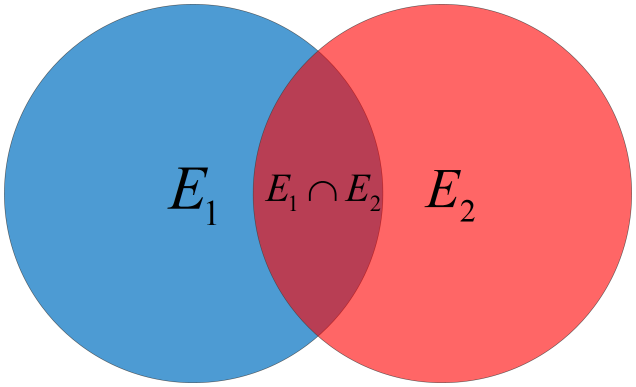


Figure 2.2: *Illustration of Conservatism in Boole's Inequality*

does not affect other states.

So we have the probability of union of these events:

$$\begin{aligned}\text{Prob}(E_1 \cup \ldots \cup E_{2n}) &= 1 - \text{Prob}(\overline{E}_1 \cap \ldots \cap \overline{E}_{2n}) \\ &= 1 - \text{Prob}(\overline{E}_1) \cap \ldots \cap \text{Prob}(\overline{E}_{2n}) \\ &= 1 - \prod_{j=1}^{2n}(1 - \delta_{ij})\end{aligned}$$

Thus the probability conservatism can be evaluated as:

$$\delta_i - \text{Prob}(E_1 \cup \ldots \cup E_{2n}) = \prod_{j=1}^{2n}(1 - \delta_{ij}) - (1 - \delta_i)$$

From arithmetic-geometric mean inequality, we know that:

$$(1 - \frac{\delta_i}{2n})^{2n} \geq \prod_{j=1}^{2n}(1 - \delta_{ij}) \geq 1 - \delta_i$$

The left-hand equality establishes when $\delta_{ij_1} = \delta_{ij_2},\ \forall j_1, j_2 \in [1, 2n]$. The right-hand equality establishes when $\delta_{ij_k} = \delta_i, \delta_{ij} = 0\ if j \neq j_k$. So the conservatism has the upper and lower bound as:

$$(1 - \frac{\delta_i}{2n})^{2n} - (1 - \delta_i) \geq \prod_{j=1}^{2n}(1 - \delta_{ij}) - (1 - \delta_i) \geq 0$$

So under the independence condition and given δ_i, n, the conservatism is minimized when the joint chance constraint can be converted into one individual chance

constraint, which means that there is only one $\delta_{ij_k} = \delta_i$. It happens often when there is only one active constraint on the system state. The conservatism is maximized when each individual chance constraint shares the same risk.

Here is an evaluation in Table 2.2 on the maximum conservatism from the effect of δ_i, n. It shows that the conservatism increases significantly when δ_i increases. The conservatism only increases slightly when n increases. It can be seen that when $\delta_i \leq 0.2$, the conservatism maintains a moderate value. When $\delta_i > 0.2$, the conservatism increases significantly.

Table 2.2: *Conservatism Evaluation*

$2n$	δ_i	Conservatism	Percentage%
10	0.1	0.0044	4.38
20	0.1	0.0046	4.61
30	0.1	0.0047	4.69
40	0.1	0.0047	4.72
50	0.1	0.0047	4.75
20	0.1	0.0046	4.61
20	0.2	0.0179	8.95
20	0.3	0.0391	13.05
20	0.4	0.0676	16.90
20	0.5	0.1027	20.54

2.4.3 Robust Optimization Approximation Method

Robust counterpart optimization has been developed by a lot of researchers. In robust counterpart optimization, the best solution is searched for in the uncertainty-induced set [22]. It is required in robust optimization that the uncertainty-induced set covers the whole uncertainty space. Such a condition guarantees that, if there is a solution, it is feasible for all the possible uncertainties. In stochastic programming, however, the uncertainty-induced set does not necessarily need to cover the whole uncertainty space [23]. We look for the set which is large enough to satisfy the probability and satisfy the constraints. It is illustrated in Fig 2.3.

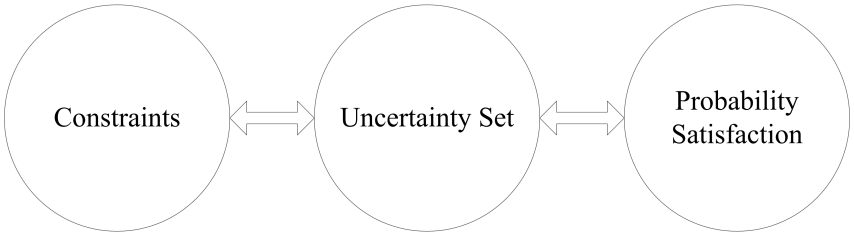


Figure 2.3: *Uncertainty Set Relation*

It is similar to the idea in ellipsoidal relaxation method. The difference is that in ellipsoidal method, uncertainty region is built without consideration of the constraint structure, while in RO method, constraints are fully considered to decide on the uncertainty set. Another difference is that ellipsoidal method is based on Gaussian distribution, while RO method can be applied on arbitrary distributions. The distribution information is used but not restricted to certain types. A selection of uncertainty set shape is important in getting less conservative result.

The general robust optimization can be written as:

$$\begin{aligned} &\min_{x} f(x, \boldsymbol{\xi}) \\ &\text{s.t. } \sup_{\boldsymbol{\xi} \in \mathbb{U}} g(x, \boldsymbol{\xi}) \leq 0 \end{aligned} \tag{2.33}$$

while the chance constraint problem is:

$$\begin{aligned} &\min_{x} f(x, \boldsymbol{\xi}) \\ &\text{s.t. } \text{Prob}(g(x, \boldsymbol{\xi}) \leq 0) \geq 1 - \delta \end{aligned} \tag{2.34}$$

The RO approximation method focuses on how to convert the constraints in 2.34 to 2.33. We will give the details of how to derive the RO approximated constraints in the following.

Consider the joint linear chance constraints in 2.11, while the uncertainties are assumed to have any kind of distribution.

By using maxima operator, the joint chance constraint can be transformed into individual chance constraints. The equivalent expression is:

$$\text{Prob}(\max_{j} \{\mathbf{y}_{0i}^{j} + \mathbf{y}_{i}^{j} \boldsymbol{\xi}\} \leq 0) \geq 1 - \delta_i \quad \forall i \tag{2.35}$$

Or equivalently,

$$\text{Prob}(\max_j\{\mathbf{y}_{0i}^j + \mathbf{y}_i^j\boldsymbol{\xi}\} > 0) \leq \delta_i \quad \forall i \tag{2.36}$$

For further derivation, define

$$\eta_i = \max_j\{\mathbf{y}_{0i}^j + \mathbf{y}_i^j\boldsymbol{\xi}\} \tag{2.37}$$

The inequality constraint can be rewritten as

$$\text{Prob}(\eta_i > 0) \leq \delta_i \tag{2.38}$$

Then the probability function can be replaced by expectation and indicator function,

$$\text{Prob}\{\eta_i > 0\} = E[1_{(0,+\infty)}(\eta_i)] \tag{2.39}$$

An upper bound in figure is applied to indicator function. It is required with $t_i > 0$, $[u]^+$ takes value u if it is positive, 0 otherwise.

$$1_{(0,+\infty)}(\eta_i) \leq \frac{1}{t_i}\{[\eta_i + t_i]^+\} \tag{2.40}$$

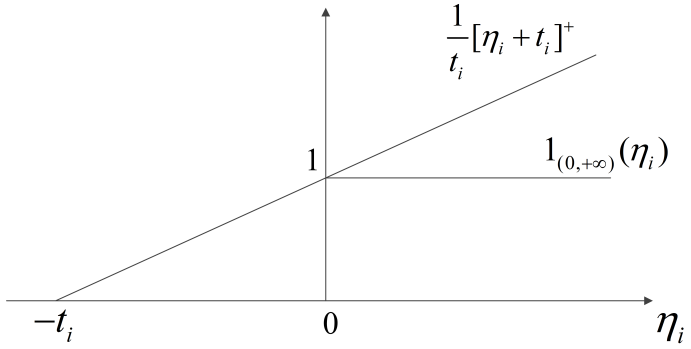


Figure 2.4: *Upper Bound on Indicator Function*

Apply the relation in the indicator to the previous constraints. We have,

$$\text{Prob}(\max_j\{y_{0i}^j + y_i^j\boldsymbol{\xi}\} > 0) \leq \frac{1}{t_i}E[(\max_j\{y_{0i}^j + y_i^j\boldsymbol{\xi}\} + t_i)^+] \tag{2.41}$$

Next, the right-hand-side of the above equation is further approximated with the Meilijson and Nadas inequality on the expected maximum [24]:

$$E[(\max_j X_j + t)^+] \leq E[(Y + t)^+] + \sum_j E[(X_j - Y)^+] \quad \text{for any r.v. Y} \tag{2.42}$$

where X_j are random variables.

Let $Y = w_{0i} + w_i^T \boldsymbol{\xi}$, where $w = [w_1 w_2 \dots w_m]^T$, and apply the above inequality for the expectation term in the previous inequality.

$$E[(\max_j\{y_{0i}^j + y_i^j \boldsymbol{\xi}\} + t_i)^+] \leq E[(\omega_{0i} + \omega_i^T \boldsymbol{\xi} + t_i)^+] + \sum_j [(y_{0i}^j + {y_i^j}^T \boldsymbol{\xi} - \omega_{0i} - \omega_i^T \boldsymbol{\xi})^+] \tag{2.43}$$

Define $u_{0i} = \omega_{0i} + t_i, u_i = \omega_i, v_{0i}^j = y_{0i}^j - \omega_{0i}$ and $v_i^j = y_i^j - \omega_i$. Substitute the new variables into equation 2.43, we have:

$$E[(\max_j\{y_{0i}^j + y_i^j \boldsymbol{\xi}\} + t_i)^+] \leq E[(u_{0i} + u_i^T \boldsymbol{\xi})^+] + \sum_j E[(v_{0i}^j + {v_i^j}^T \boldsymbol{\xi})^+] \tag{2.44}$$

As discussed before, the uncertainty-induced set $\mathbb{U}$ is introduced here. $\mathbb{U}_i$ is the uncertainty set for each joint chance constraint i. The upper bound of the expectation is described in the following:

$$E[(\mu_0 + \boldsymbol{\xi}^T \mu)^+] \leq (\mu_0 + \max_{\boldsymbol{\xi} \in \mathbb{U}} \boldsymbol{\xi}^T \mu)^+ \tag{2.45}$$

Based on the previous inequalities 2.44 and 2.45, the following relation holds:

$$E[(\max_j\{y_{0i}^j + y_i^j \boldsymbol{\xi}\} + t_i)^+] \leq (u_{0i} + \max_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T u_i)^+ + \sum_j (v_{0i}^j + \max_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T v_i^j)^+ \tag{2.46}$$

Put it back to inequality 2.35:

$$\text{Prob}(\max_j\{y_{0i}^j + y_i^j \boldsymbol{\xi}\} > 0) \leq \frac{1}{t_i}\{(u_{0i} + \max_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T u_i)^+ + \sum_j (v_{0i}^j + \max_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T v_i^j)^+\} \tag{2.47}$$

The approximation then becomes:

$$\frac{1}{t_i}\{(u_{0i} + \max_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T u_i)^+ + \sum_j (v_{0i}^j + \max_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T v_i^j)^+\} \leq \delta_i \tag{2.48}$$

Let

$$\phi_i = (u_{0i} + \max_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T u_i)^+ \quad \gamma_i^j = \sum_j (v_{0i}^j + \max_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T v_i^j)^+$$

We have:

$$\phi_i + \sum_j \gamma_i^j \leq \delta_i t_i \tag{2.49}$$

Then remove the $(\cdot)^+$ terms, the following equivalent formulation is derived:

$$\begin{cases} \phi_i + \sum_j \gamma_i^j \leq \delta_i t_i \\ \phi_i \geq u_{0i} + \max\limits_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T u_i \\ \phi_i \geq 0 \\ \gamma_i^j \geq v_{0i}^j + \max\limits_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T v_i^j \\ \gamma_i^j \geq 0 \end{cases} \tag{2.50}$$

To summarize, the approximation to multiple joint chance constraints is given by:

$$\begin{cases} \phi_i + \sum_j \gamma_i^j \leq \delta_i t_i \\ \phi_i \geq u_{0i} + \max\limits_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T \mathbf{u}_i \\ \phi_i \geq 0 \\ \gamma_i^j \geq v_{0i}^j + \max\limits_{\boldsymbol{\xi} \in \mathbb{U}_i} \boldsymbol{\xi}^T v_i^j \quad \forall j \\ \gamma_i^j \geq 0 \quad \forall j \\ u_{0i} = \omega_{0i} + t_i \\ \mathbf{u}_i = \boldsymbol{\omega}_i \\ v_{0i}^j = y_{0i}^j - \omega_{0i} \quad \forall j \\ \mathbf{v}_i^j = \mathbf{y}_i^j - \boldsymbol{\omega}_i \quad \forall j \end{cases} \quad \forall i \tag{2.51}$$

Let us introduce different kinds of robust uncertainty set, including box, polyhedral and ellipsoidal.

For the joint chance constraint, with a box type uncertainty set defined as $\mathbb{U}_i = \{\boldsymbol{\xi} || \boldsymbol{\xi}_k | \leq \Delta_i, k = 1, \ldots, m\}$, the corresponding robust optimization approximation

model is equivalent to,

$$\begin{cases} \phi_i + \sum_j \gamma_i^j \le \delta_i t_i \\ \phi_i \ge u_{0i} + \Delta_i \sum_k p_i^k \\ -p_i^k \le u_i^k \le p_i^k \quad \forall k \\ \phi_i \ge 0 \\ \gamma_i^j \ge v_{0i}^j + \Delta_i \sum_k q_i^{j,k} \quad \forall j \\ -q_i^{j,k} \le v_i^{j,k} \le q_i^{j,k} \quad \forall j,k \\ \gamma_i^j \ge 0 \quad \forall j \\ u_{0i} = \omega_{0i} + t_i \\ \mathbf{u}_i = \boldsymbol{\omega}_i \\ v_{0i}^j = y_{0i}^j - \omega_{0i} \quad \forall j \\ \mathbf{v}_i^j = \mathbf{y}_i^j - \boldsymbol{\omega}_i \quad \forall j \end{cases} \qquad \forall i \tag{2.52}$$

For polyhedral type uncertainty set defined as $\mathbb{U}_i = \{\sum_{k=1}^m |\boldsymbol{\xi}| \le \Delta_i\}$, the corre-

sponding robust optimization approximation model is equivalent to,

$$\begin{cases} \phi_i + \sum_j \gamma_i^j \le \delta_i t_i \\ \phi_i \ge u_{0i} + \Delta_i z_{0i} \\ z_{0i} \ge p_i^k \quad \forall k \\ -p_i^k \le u_i^k \le p_i^k \quad \forall k \\ \phi_i \ge 0 \\ \gamma_i^j \ge v_{0i}^j + \Delta_i z_i^j \quad \forall j \\ z_i^j \le q_i^{j,k} \quad \forall j,k \\ -q_i^{j,k} \le v_i^{j,k} \le q_i^{j,k} \quad \forall j,k \\ \gamma_i^j \ge 0 \quad \forall j \\ u_{0i} = \omega_{0i} + t_i \\ \mathbf{u}_i = \boldsymbol{\omega}_i \\ v_{0i}^j = y_{0i}^j - \omega_{0i} \quad \forall j \\ \mathbf{v}_i^j = \mathbf{y}_i^j - \boldsymbol{\omega}_i \quad \forall j \end{cases} \quad \forall i \tag{2.53}$$

For ellipsoidal type uncertainty set defined as $\mathbb{U}_i = \{\sum_{k=1}^m \boldsymbol{\xi}_k^2 \le \Delta_i^2\}$, the corresponding robust optimization approximation model is equivalent to,

$$\begin{cases} \phi_i + \sum_j \gamma_i^j \le \delta_i t_i \\ \phi_i \ge u_{0i} + \Delta_i \sqrt{\sum_k \left(u_i^k\right)^2} \\ \phi_i \ge 0 \\ \gamma_i^j \ge v_{0i}^j + \Delta_i \sqrt{\sum_k \left(v_i^{j,k}\right)^2} \quad \forall j \\ \gamma_i^j \ge 0 \quad \forall j \\ u_{0i} = \omega_{0i} + t_i \\ \mathbf{u}_i = \boldsymbol{\omega}_i \\ v_{0i}^j = y_{0i}^j - \omega_{0i} \quad \forall j \\ \mathbf{v}_i^j = \mathbf{y}_i^j - \boldsymbol{\omega}_i \quad \forall j \end{cases} \quad \forall i \tag{2.54}$$

Still we need to bring the question of designing the set size Δ_i which satisfies the probability satisfaction. The uncertainty set which covers the whole uncertainty area is surely too conservative. It is given in [23] that two methods are given to evaluate the probability guarantees. One is prior probability bound, and the other is posterior probability bound. For prior probability bound, probability guarantee with set size is calculated first before optimization. For posterior probability bound, probability is calculated directly from optimization result, which can be viewed as a probability check.

Apparently, the result from prior probability method can be conservative or even infeasible. The posterior probability method will lead to closer result to the true optimal, but it takes more iterations.

We choose the posterior probability for closer approximation. A straightforward and efficient way to calculate the probability is using Monte Carlo method. For most of the cases, with smaller set size, probability of satisfaction will decrease and objective becomes less conservative and vice versa [23], [7]. Bisection method has been widely used in root finding for monotonic functions. So to get closer to the probability satisfaction, bisection method is used to find appropriate set size Δ_i.

Thus, the RO method can be separated into two stages: upper stage for finding appropriate set size and lower stage for optimization with fixed set size.

Upper stage optimization:

$$\begin{aligned} &\min_{\mathbf{\Delta}} \sum_{i=1}^{N} |p_i - (1-\delta_i)| \\ \text{s.t.} \quad & \\ & p_i = p_{i,satisfaction}(\mathbf{\Delta}) \quad \forall i \end{aligned} \tag{2.55}$$

Lower stage optimization with fixed $\mathbf{\Delta}$:

$$\begin{aligned} &\min_{U,\text{all other slack variables}} \{\mathbf{U}^T Q_{qua} \mathbf{U} + 2 Q_{lin} \mathbf{U} + c\} \\ \text{s.t.} \quad & \\ & \mathbf{L_h} \leq \mathbf{U} \leq \mathbf{U_h} \\ & 2.52 \quad \text{or} \quad 2.53 \quad \text{or} \quad 2.54 \quad \text{based on the uncertainty type chosen} \end{aligned} \tag{2.56}$$

Remark 1. The lower stage optimization will be solved with fixed set size $\mathbf{\Delta}$. The result of optimal $\mathbf{U}^*, J^*, t_{cal}$ and probability of satisfaction $p_{i,satisfaction}(\mathbf{\Delta})$ will be given to the upper stage. The upper stage tries to find the closest result to the required probability iteration by iteration.

Remark 2. The probability of satisfaction is estimated with Monte Carlo sampling technique: NM samples are drawed to give the estimation of probability of multiple joint chance constraints:

$$p_{i,satisfaction}(\mathbf{\Delta}) = \frac{1}{M}\sum_{s=1}^{M} 1_{(0,\infty)}\left(\max_j\{y_{0i}^j + \mathbf{y}_i^j \boldsymbol{\xi}^{(s)}\} \leq 0\right) \quad \forall i \tag{2.57}$$

Remark 3. Here are some issues which need to be dealt with in the bisection method. These issues happen due to the specific properties of multiple joint chance constraints in model predictive control problems. The first one is to consider the time progress in model predictive control. There will be certain occasions when the states are far away from the boundary so that almost all possible uncertainties will not drive the states to violate the constraints. The probability of satisfaction will always be larger than the required one, then set size will equal to the lower bound.

Another issue we need to consider is that for each joint chance constraint, the solution will be affected by the set size, and the solution will affect the probability of other joint chance constraints. When changing all the set sizes simultaneously, some joint chance constraints can fall into the zone where their probability can never be satisfied. One solution is after a certain times of iterations, the upper bound will be increased if the probability is still pretty small or the lower bound will be decreased if the probability is still pretty large. The algorithm for optimization over the set size is listed in Table 2.3.

Remark 4. The RO method approximates the joint chance constraint with linear constraints. Uncertainty-induced set is brought to approximate the original uncertainties. In this way, the distribution can be ignored to obtain a general method for all kinds of distributions. However, due to the loss of the distribution information, the uncertainty set shape becomes a key in deciding the performance. We give the derivation for three kinds of uncertainty set here: box, ellipsoidal and polyhe-

Table 2.3: *Algorithm for Optimization over Set Size*

Algorithm for Optimization over Set Size
1. Solve the lower stage problem with $\Delta_i = \Delta_{LB_{ini}},\ \forall i$ If $p_{i,satisfaction}(\boldsymbol{\Delta}) \geq 1 - \delta_i,\ \forall i,$ Then $\Delta_i = \Delta_{LB_{ini}},\ \forall i$, Return End
2. $\Delta_i = (\Delta_{LB_{ini}} + \Delta_{UB_{ini}})/2,\ \forall i$, solve the lower stage optimization with $\Delta_i,\ \forall i$.
3. If last calculation is infeasible, Then set $p_i = 1,\ \forall i$
4. If $p_i < 1 - \delta_i - p_{tol}$ or $p_i > 1 - \delta_i + p_{tol} \quad \forall i$ If $p_i < 1 - \delta_i - p_{tol},$ Then $\Delta_{LB}(i) = \Delta_i$ If $p_i > 1 - \delta_i + p_{tol},$ Then $\Delta_{UB}(i) = \Delta_i$ $\Delta_i = 0.5(\Delta_{LB}(i) + \Delta_{UB}(i))$ End
5. If iteration = ITERtol If $p_i < 1 - \delta_i - p_{tol},$ Then $\Delta_{UB}(i) = \Delta_{UB_{ini}} \quad \forall i$ If $p_i > 1 - \delta_i + p_{tol},$ Then $\Delta_{LB}(i) = \Delta_{LB_{ini}} \quad \forall i$ End
6. Solve the lower stage optimization with $\boldsymbol{\Delta}$
7. CONDITION i: If $p_i \in [1 - \delta_i - p_{tol}, 1 - \delta_i + p_{tol}]$ or $p_i \geq 1 - \delta_i + p_{tol}$ with $\Delta_i \in [\Delta_{LB_{ini}}, \Delta_{LB_{ini}} + \Delta_{tol}]$, then CONDITION i is satisfied. If all CONDITION i is satisfied, then CONDITION is satisfied. If CONDITION is not satisfied, then come back to step 3. If satisfied, then return.

dral. The choice really depends on the fit to the distribution shape. Thus they can have drawbacks when dealing with complicated distributions, like Gaussian mixture models.

2.4.4 Comparison Case

To test and compare different methods, control is applied to the system defined by:

$$x_{k+1} = Ax_k + Bu_k + W\omega_k$$

$$A = \begin{bmatrix} 1.2 & 0 & 0 \\ 0.17 & 0.87 & 0 \\ -0.13 & 0.87 & 0 \end{bmatrix}, B = \begin{bmatrix} 0.17 \\ 7.2 \times 10^{-3} \\ -0.1628 \end{bmatrix} W = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \\ -0.01 & 0.01 \end{bmatrix}$$

The constraints are only soft constraints:

$$\text{Prob}(L_s \leq x_{k+i} \leq U_s) \geq 1 - \delta \quad i = 1, 2, \ldots, N$$

$$L_s = \begin{bmatrix} -0.3 \\ -0.2 \\ -0.81 \end{bmatrix}, U_s = \begin{bmatrix} 1.01 \\ 1.10 \\ 0.42 \end{bmatrix}$$

The other parameters are: $N = 10, \delta = 0.025, \Delta\delta_{tol} = 0.001$. The system noise follows standard Gaussian distribution: $\omega_k \sim N(\mathbf{0}, I)$.

The initial state and reference state are respectively:

$$x_{ini} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, x_{ref} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

For each time step, the objective function is formulated as:

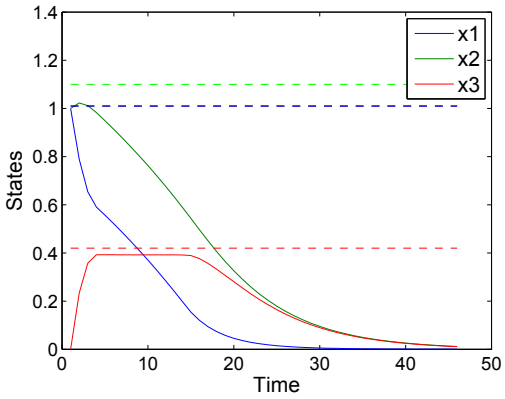
$$OBJ = \min_{u_i} \sum_{i=1}^{N} (x_i^T Q x_i + u_i R u_i)$$

s.t

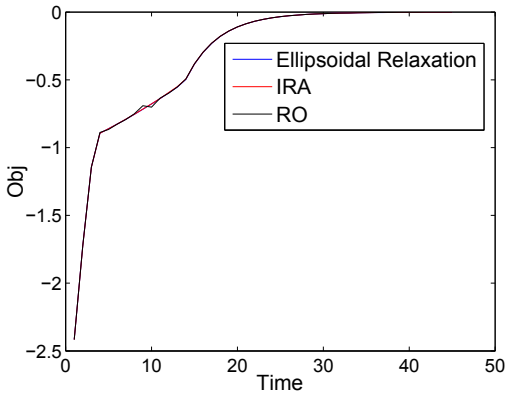
$$\text{Prob}(L_s \leq x_i \leq U_s) \geq 1 - \delta \quad i = 1, 2, \ldots, N$$

Then equivalently, the optimization problem can be formulated as in 2.7 and 2.8. A more general form of chance constraint is then stated as 2.10. The problem is then solved through 2.51 with RO method.

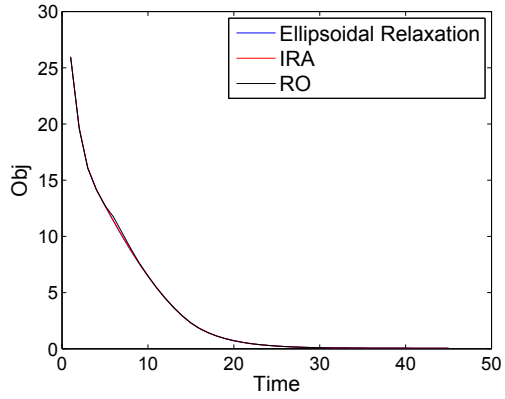
The result is shown in Fig 2.5.



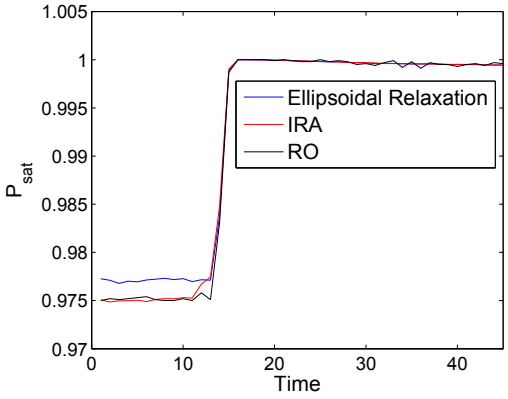
(a) *States Trajectory and Constraints*



(b) *System Inputs*



(c) *Objective Value*

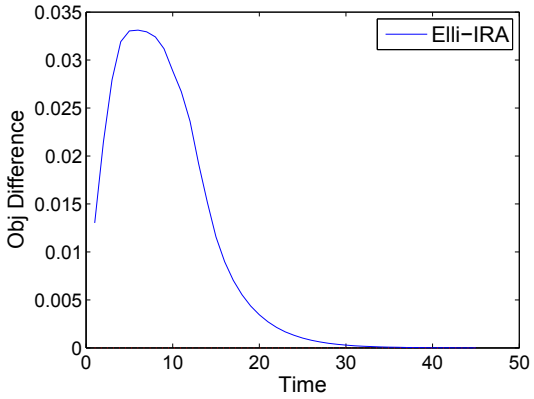


(d) *Probability Satisfaction*

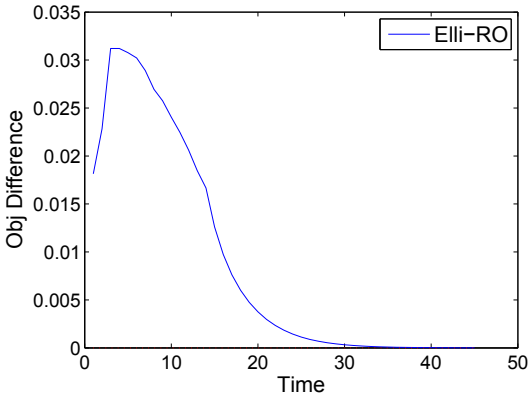
Figure 2.5: *Comparison of Different Methods*

The system evolves with the realization which is the mean value of the disturbances. From Fig 2.5a we can see that x_3 keeps a safe distance from the boundary in order not to violate the constraint in probability sense. The states eventually converge to the steady states and the objective value decreases to zero. For the probability of satisfaction, ellipsoidal method does not make it close enough to the required probability due to its inherit conservatism. IRA and RO method have a continuous transition from required probability to 100%.

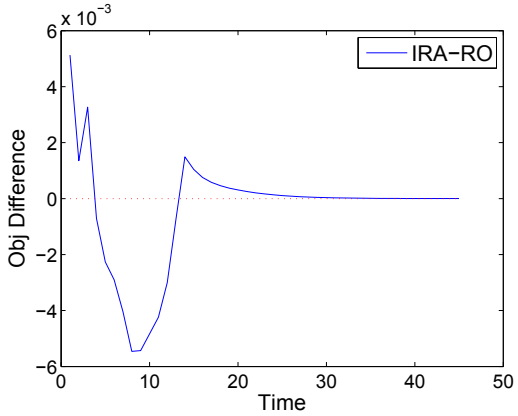
The performance in objective value is compared by calculating difference between each method, which are shown in the Fig 2.6a, 2.6b, 2.6c. It is obvious to see that the ellipsoidal method remains the most conservative one. IRA and RO are close in



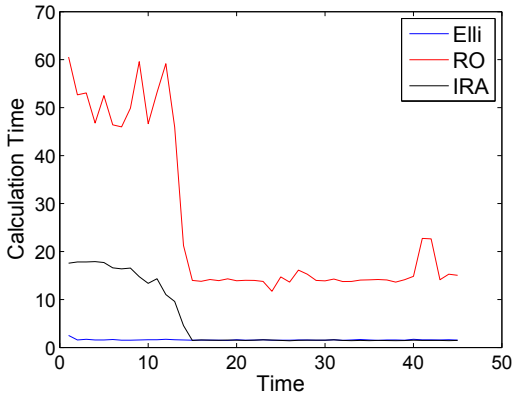
(a) *OBJ Difference between Ellipsoidal and IRA*



(b) *OBJ Difference between Ellipsoidal and RO*



(c) *OBJ Difference between IRA and RO*



(d) *Calculation Time*

Figure 2.6: *Detailed Comparison of Different Methods*

performance while IRA is better in time period $[4, 13]$ and RO is better in time $[1, 3]$ and $[14, 36]$. When considering the calculation time, ellipsoidal method takes the least time as expected, IRA is modest and RO takes the longest time. RO does not show significant advantages over IRA when dealing with Gaussian noise. However, when considering arbitrary distribution, RO is a good choice as a distribution-free method.

2.4.5 Closed-loop MPC

The performance in objective function can be further improved by including closed-loop structure. We need to bring in game theory for further explanation. Suppose we have two individuals to attend one game, one needs to predict the other's movement

in the future and prepares for the worst circumstance. It is similar situation in the open-loop MPC. The controlled input is calculated based on all possible future disturbances. The "controller" is the player who does not know any plan of the other player called "disturbance". However, closed-loop MPC comes from the idea that after player "disturbance" plays in the future, the player "controller" employs control policy $u = u(x, \omega)$ based on all the available knowledge of the "disturbance" movement [25]. Thus the negative effect from disturbance is greatly reduced.

To restrict the control policies to a finite-dimensional subspace, affine disturbance feedback is chosen. Define $\mathbf{u} = M\boldsymbol{\xi} + h$ for the closed-loop control, we have:

$$u_{t+k|t} = \sum_{j=0}^{k-1} M_{(t+k,t+j)|t}\omega_{t+j|t} + h_{t+j|t}$$

$$M = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ M_{(t+1,t)|t} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ M_{(t+2,t)|t} & M_{(t+2,t+1)|t} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M_{(t+N-1,t)|t} & M_{(t+N-1,t+1)|t} & M_{(t+N-1,t+2)|t} & \cdots & \mathbf{0} \end{bmatrix} \tag{2.58}$$

The objective function is formulated as,

$$\min_{\mathbf{u}}[\mathbf{h}^T Q_{qua}\mathbf{h} + 2Q_{lin}\mathbf{h} + c] \tag{2.59}$$

where,

$$Q_{qua} = H^T Q_{obj} H + R_{obj},\ Q_{lin} = {x_k}^T G^T Q_{obj} H$$
$$c = {x_k}^T G^T Q_{obj} G x_k$$

The chance constraints 2.8 are reformulated as,

$$\text{Prob}\{S_i(Gx_k + Hh + (P + HM)\boldsymbol{\xi}) \leq \mathbf{s}\} \geq 1 - \delta_i, \qquad 1 \leq i \leq N \tag{2.60}$$

Equivalently, we have,

$$\text{Prob}(\mathbf{y}_{0i}^j + \mathbf{y}_i^j \boldsymbol{\xi} \geq 0) \leq \delta_{ij}, \qquad \forall i, j \tag{2.61}$$

$$\mathbf{y}_{0i}^j = F_j(S_i(Gx_k + Hh) - \mathbf{s}), \quad \mathbf{y}_i^j = F_j S_i(P + HM) \qquad \forall i, j \tag{2.62}$$

Thus, for the ellipsoidal method, the chance constraints are formulated as:

$$\mathbf{y}_{0i}^{j} + \mathbf{y}_{i}^{j}\mu + r\sqrt{\mathbf{y}_{i}^{j}\Sigma_{\boldsymbol{\xi}}\mathbf{y}_{i}^{j^T}} \leq 0, \qquad \forall i, j \tag{2.63}$$

$$\mathbf{y}_{0i}^{j} = F_j(S_i(Gx_k + Hh) - \mathbf{s}), \quad \mathbf{y}_{i}^{j} = F_j S_i(P + HM) \qquad \forall i, j \tag{2.64}$$

Compared to the open-loop case, the deterministic part becomes a second-oder cone constraint. It can be solved with great efficiency by interior point methods.

For IRA method, the chance constraints are formulated as:

$$\mathbf{y}_{0i}^{j} + \mathbf{y}_{i}^{j}\mu + \Phi^{-1}(1 - \delta_{ij})\sqrt{\mathbf{y}_{i}^{j}\Sigma_{\boldsymbol{\xi}}\mathbf{y}_{i}^{j^T}} \leq 0 \qquad \forall i, j \tag{2.65}$$

$$\mathbf{y}_{0i}^{j} = F_j(S_i(Gx_k + Hh) - \mathbf{s}), \quad \mathbf{y}_{i}^{j} = F_j S_i(P + HM) \qquad \forall i, j \tag{2.66}$$

The diverging property of IRA method is preserved while the difference is the same with ellipsoidal method: linear constraints are changed into second-oder cone constraints.

For RO method, the deterministic constraints 2.52 2.53 2.54 are the same except for the definition of $\mathbf{y}_{0i}^{j}, \mathbf{y}_{i}^{j}$. The difference is the same to the above two methods.

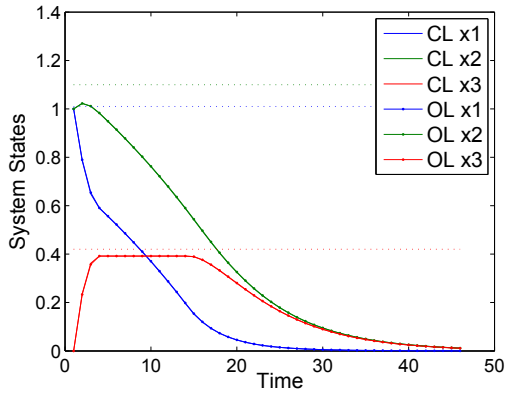
We can compare the improvement for a simple case study with RO method. The previous example is still used here.

From Fig 2.7, we can tell that the closed-loop with disturbance affine feedback behaves better than the open-loop in objective value. The linear constraints are now second-oder cone constraints with more slack variables included. The total computational effort required by second-oder cone constraints is greater than linear programming problems.

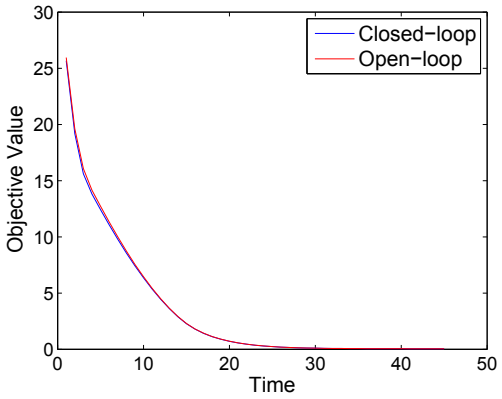
2.5 Case Study

2.5.1 Hydrodesulfuration Process

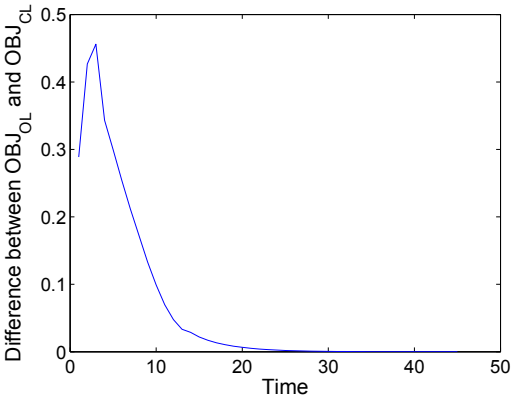
In Petroleum refineries the hydrodesulfuration process is used to remove sulfur from the hydrocarbons to fulfill environmental policies. To perform the desulfuration reaction, hydrogen is put in contact with hydrocarbon in fixed bed reactors with a specific catalyst.



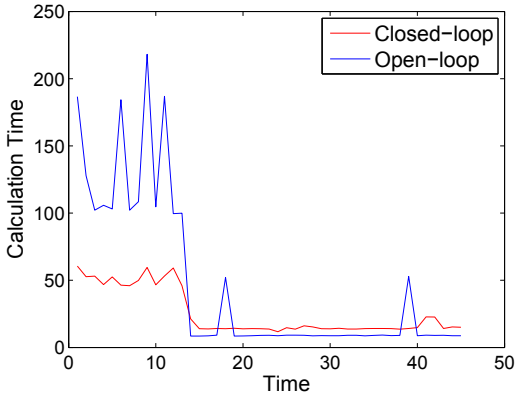
(a) *States Trajectory and Constraints*



(b) *Objective Function Value*



(c) *Difference between Objective Value*



(d) *Calculation Time*

Figure 2.7: *Comparison of Open-loop and Closed-loop Using RO Method*

Fig 2.8 shows a hydrodesulfuration process. Several key variables are explained here: F_1, F_3, F_5, F_{10} are feed-in molar flow rate of stream 1,3,5 and 10, stream 2 is shut down due to high operation cost. F_1, F_{10} are the manipulated variables. X_1, X_5 are the molar fraction of hydrogen in stream 1 and 5. ξ_2 is the molar fraction of hydrogen in stream 3, which is a random variable following Gaussian distribution $N(\mu_2, \sigma_2^2)$. F_X^{H2} is the hydrogen consumption rate inside the reactors, X_{H2} is the hydrogen molar fraction correspondingly. X_5, X_{H2} are the controlled variables. F_{HC} is the flow rate of hydrocarbon to be desulfurized. ξ_1 is the hydrogen consumption rate and follows Gaussian distribution $N(\mu_1, \sigma_1^2)$. The parameters of the process are shown in Table 2.4.

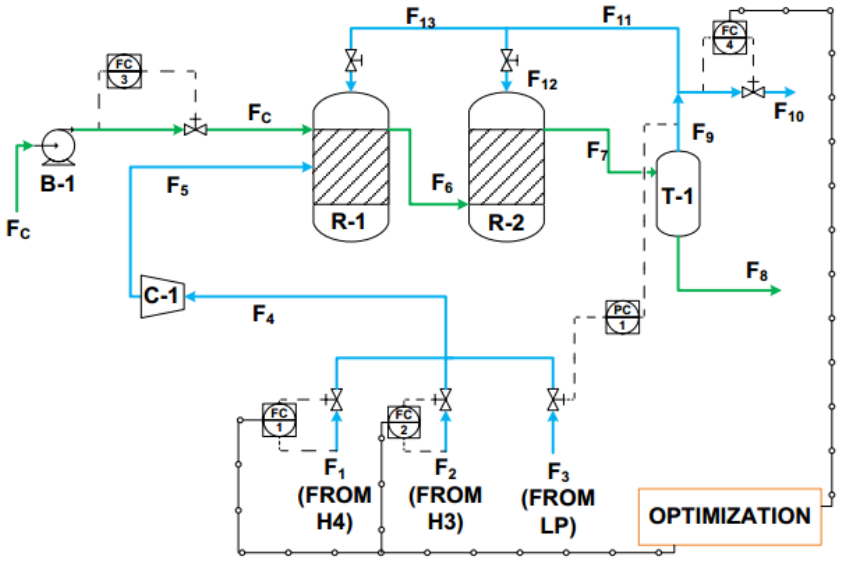


Figure 2.8: *Hydrodesulfuration Process*

We have the following equations describing model of the process:

$$F_1 X_1 + F_3 \xi_2 = F_5 X_5$$

$$F_1 + F_3 = F_5$$

$$\tau_2 \frac{dX_{H2}}{dt} = F_5 X_5 - F_X^{H2} - F_{10} X_{H2}$$

$$\tau_1 \frac{dF_X^{H2}}{dt} + F_X^{H2} = F_{HC} \xi_1$$

where,

$$\tau_2 = \frac{VP}{ZRT} \tag{2.67}$$

The constraints on manipulated variables:

$$F_1^L \leq F_1 \leq F_1^U$$

The constraints on controlled variables:

$$\text{Prob}(X_5 \geq X_5^L, X_{H2} \geq X_{H2}^L) \geq \alpha$$

The system model is linearized around the steady states F_{Xss}^{H2}, X_{5ss} and steady

Table 2.4: *Parameters in Hydro-desulfuration Process*

Parameter	Value	Unit	Parameter	Value	Unit
X_1	0.991	1	μ_1	12.5	$kmol/m^3$
τ_1	0.3	h	μ_2	0.85	1
τ_2	132.9041	h	σ_1	0.4	$kmol/m^3$
F_1^L	0	$kmol/h$	σ_2	0.013	1
F_1^U	1400	$kmol/h$	X_5^L	0.9	1
F_{HC}	102	m^3/h	X_{H2}^L	0.7	1

inputs $F_{1si}, F_{10si}, \xi_{1si}, \xi_{2si}$, then discretized:

$$x_d(k+1) = A_d x_d(k) + B_d u_d(k) + W_d \xi_d(k)$$

$$y_d(k) = C x_d(k) + D u_d(k) + P \xi_d(k)$$

$$x_d(k) = \begin{bmatrix} \Delta F_X^{H2}(k) \\ \Delta X_{H2}(k) \end{bmatrix}, u_d(k) = \begin{bmatrix} \Delta F_1(k) \\ \Delta F_{10}(k) \end{bmatrix}, \xi_d(k) = \begin{bmatrix} \Delta \xi_1(k) \\ \Delta \xi_2(k) \end{bmatrix}, y_d(k) = \Delta X_5(k)$$

where the values of the parameters are in the Appendix. Δ represents the difference between the real value and the steady states.

The constraints are then formulated as,

$$F_1^L - F_{1si} \leq \Delta F_1(k) \leq F_1^H - F_{1si}$$

$$\text{Prob}(\Delta X_5 \geq X_5^L - X_{5ss}, \Delta X_{H2} \geq X_{H2}^L - X_{H2ss}) \geq \alpha$$

The objective function in MPC problem is formulated as,

$$OBJ = \min_{u_d(i)} \sum_{i=1}^{N} \left(x_d(i)^T Q x_d(i) + u_d(i)^T R u_d(i) \right)$$

The solution procedures follow similar steps as in comparison case study.

In Fig 2.9, the system states and output are $x_{d1}(\Delta F_X^{H2})$: hydrogen consumption rate inside the reactor, $x_{d2}(\Delta X_{H2})$: hydrogen composition in the reactor, $y_d(\Delta X_5)$: hydrogen composition inside stream 5. The constraints are plotted as dashed lines.

From the comparison case we can see that nominal MPC cannot deal with the uncertainties in feasibility. The violation on constraints cannot be regulated through

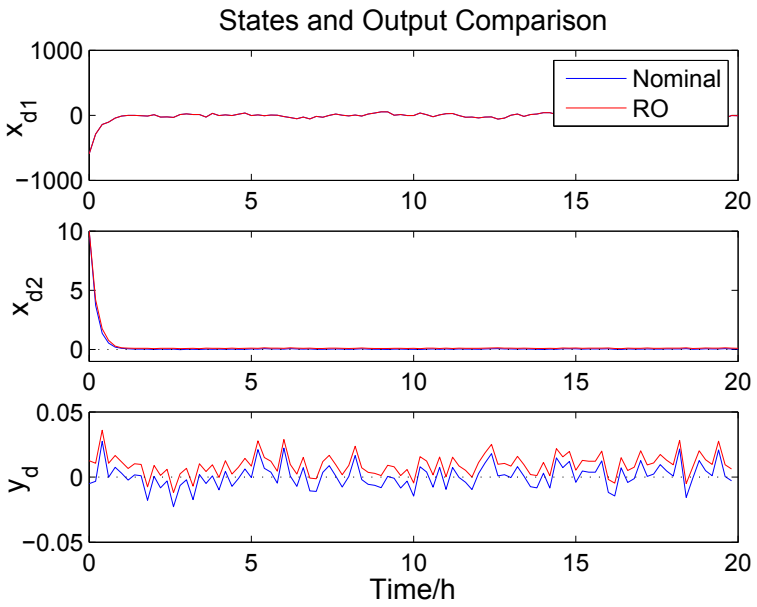


Figure 2.9: *Hydrodesulfuration Operation in Different Method*

Table 2.5: *Comparison of Probability Violation in Hydrodesulfuration Process*

Different Methods in Probability Comparisons		
	RO Method	Nominal MPC
Total Violation Percentage %	10	49
X2 Violation Percentage %	0	4
Y Violation Percentage %	10	46

the method of nominal MPC. However, RO method shows significant improvement over nominal MPC. The violation of the constraints is controlled at 10%.

2.5.2 Continuous Stirred-tank Reactor Process

The second case study is the continuous stirred-tank reactor (CSTR) process:

Consider a well-mixed, non-isothermal continuous stirred tank reactor in Fig 2.10 where three parallel irreversible elementary exothermic reactions take place of the form $A \to B$, $A \to C$, $A \to D$. B is the desired product and C, D are by-products.

The feed to the reactor consists of pure A at flow rate F, temperature T_{A0} and

molar concentration C_{A0}. Due to the non-isothermal nature of the reactor, a jacket is used to remove/provide heat Q to the reactor. The product temperature and concentration are desired to be controlled. The parameters of CSTR are shown in Table 2.6.

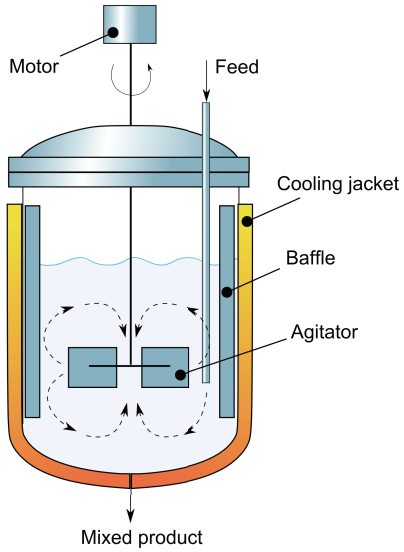


Figure 2.10: *CSTR Illustration*

The system model is described as the following,

$$\frac{dT}{dt} = \frac{F}{V_r}(T_{A0} - T) - \sum_{i=1}^{3} \frac{\Delta H_i}{\sigma c_p} k_{i0} e^{\frac{-E_i}{RT}} C_A + \frac{Q}{\sigma c_p V_r}$$

$$\frac{dC_A}{dt} = \frac{F}{V_r}(C_{A0} - C_A) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT}} C_A$$

The system is linearized at unsteady operation states and discretized:

$$x_d(k+1) = A_d x_d(k) + B_d u_d(k)$$

$$x_m(k) = x_d(k) + W_d \omega_d(k)$$

$$x_d(k) = \begin{bmatrix} \Delta T(k) \\ \Delta C_A(k) \end{bmatrix}, u_d(k) = \begin{bmatrix} \Delta C_{A0}(k) \\ \Delta Q(k) \end{bmatrix}$$

where $x_m(k)$ is the measured states at time k, $W_d \omega_d(k)$ is the measurement noise. $W_d = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.001 \end{bmatrix}, \omega \sim N(0, I)$.

Table 2.6: *Parameters in CSTR Process*

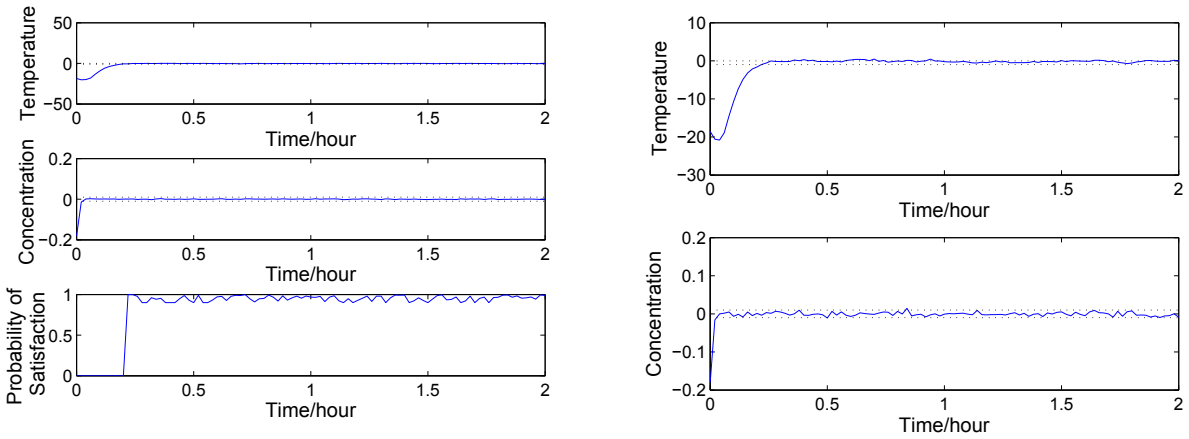
Parameter	Value	Unit	Parameter	Value	Unit
F	4.998	m^3/h	k_{10}	3×10^6	h^{-1}
V_r	1	m^3	k_{20}	3×10^5	h^{-1}
R	8.314	$KJ/kmol \cdot K$	k_{30}	3×10^5	h^{-1}
T_{A0}	300	K	E_1	5×10^4	$KJ/kmol$
C_{A0}	4	$kmol/m^3$	E_2	7.53×10^4	$KJ/kmol$
ΔH_1	-5.0×10^4	$KJ/kmol$	E_3	7.53×10^4	$KJ/kmol$
ΔH_2	-5.2×10^4	$KJ/kmol$	σ	1000	kg/m^3
ΔH_3	-5.4×10^4	$KJ/kmol$	c_p	0.231	$KJ/kg \cdot K$

The constraints are defined as,

$$u_{min} \leq u_d(k) \leq u_{max} \quad \forall k$$

$$\text{Prob}(x_{min} \leq x_d(k) \leq x_{max}) \geq \alpha \quad \forall k \in \text{steady state region}$$

where α is the probability satisfaction. The chance constraints are imposed on the system after states come into the "steady states region": $[x_{min}, x_{max}]$.



(a) *State Trajectory with RO CCMPC* **(b)** *State Trajectory with Nominal MPC*

Figure 2.11: *State Trajectory with Different Methods*

The solution procedures follow similar steps as in comparison case study.

Table 2.7: *Comparison of Probability Violation in CSTR*

Different Methods in Probability Comparisons		
	RO method	Nominal MPC
Total violation percentage %	5.49	37.36
Temperature violation percentage %	5.49	37.36
Concentration violation percentage %	0	0

In the Fig 2.11a, 2.11b, blue line represents the system real response, red line represents prediction of the states. As the system itself is nonlinear, the prediction based on linear system is calculated to approximate the real states. In Fig 2.11a, it also shows the probability satisfaction of the predicted states. The probability of states satisfies the requirement of 90%. The average violation in RO method is 5.49% while in nominal MPC violations raised by the uncertainty is as high as 37.36%. The RO method takes in consideration of uncertainty and guarantees the violation to be lower than the required amount.

2.6 Conclusions

In this chapter, we proposed RO method to solve chance constrained MPC problem with additive system noises. A joint chance constrained problem under MPC framework is formulated at first. Then ellipsoidal relaxation method and iterative risk allocation are introduced but both show their drawbacks in sub-optimality and restrictions of uncertainty distributions. RO method is proposed as a distribution-free method with results close to optimal solutions. A closed-loop structure can be further included to improve the optimization performance. RO based chance constrained MPC is also applied to two simulation case studies: hydrodesulfuration process and CSTR process. Both case studies show a great reduction in constraint violation compared to traditional MPC.

Chapter 3

Chance Constrained Model Predictive Control in Dealing with Parameter Uncertainties

3.1 Introduction

The previous chapter deals with feasibility and optimality issues of system with additive noise. However, as all physical systems are inherently nonlinear and time-varying, not all of them can be linearized with fixed parameters. Linear Parameter Varying (LPV) systems are concerned with linear dynamics with exogenous non-stationary parameters.

$$x_{k+1} = A(\theta_k)x_k + B(\theta_k)u_k + W\omega_k \tag{3.1}$$

One problem that rises in the LPV system is the computational complexity. Computational complexity grows exponentially with the prediction horizon N: given an initial polyhedral uncertainty set with s vertices, the N step prediction will lie in a polytope with s^N vertices. Facing the problem of computational complexity, the predictive control of LPV system is restricted to systems with small prediction horizon. Another issue with the LPV system is the multiplicative uncertainty. Consider the

system with only parameter uncertainty:

$$x_{k+1|k} = A(\omega_k)x_k + B(\omega_k)u_k$$
$$x_{k+2|k} = A(\omega_{k+1})A(\omega_k)x_k + A(\omega_{k+1})B(\omega_k)u_k + B(\omega_{k+1})u_{k+1}$$
$$\dots$$

With further predictions in time, the uncertainty terms are multiplied by each other. It is difficult to preserve the linear form of chance constraints on which we can implement the methods we have introduced before.

Tube MPC is first introduced to solve robust MPC problem for additive noise with reduced computational complexity. The constraint tightening approach [26], [27], [28] or the tube MPC method [29] is done through tightening constraints on nominal states with reduced calculation. The stochastic tube containing the noise is constructed by propagating uncertainty set in the system dynamics. The constraints on the nominal states are then tightened by excluding the stochastic tube. As the prediction horizon increases, the stochastic tube will expand drastically, leading to the shrinking of the nominal state constraint set. With a large prediction horizon, it is hard to guarantee the feasibility of this method, which will also rise instability problems.

Dual mode MPC is introduced in [26], [30] to provide a guarantee for recursive feasibility and stability. State feedback parameter K is designed off-line for unconstrained system to guarantee stability: after the system comes into the terminal set, the system can be steered into the robust invariant set autonomously. Terminal constraint is applied in the prediction horizon to guarantee recursive feasibility. Notice that with additive noise, the system cannot be kept at the steady states, a robust invariant set is defined of which the states will never come out when purely driven by bounded disturbance.

With the success of tube based method in robust MPC, stochastic MPC becomes more attractive to be studied as the distribution information is being used in stochastic MPC, which helps to achieve a better performance. In [10], Mark Cannon proposes one way of generating probabilistic tubes with ellipsoidal shape in LTI system with

additive noise, using fixed cross-section and variable scaling. The scaling is varied to guarantee the probability satisfaction. Ellipsoidal shape can be replaced by polyhedral shape to approximate more general distributions. However, as the probabilistic constraints are calculated to give the scaling before solving optimization problem, it is easy to get conservative results. [31] proposes a related method for a more general case when both additive and multiplicative uncertainty are included. The system is divided into nominal dynamics and error dynamics, and the error is divided into two parts: one with simple additive random variables, the other part which consists of both additive and multiplicative random variables. The additive random variables are required to satisfy the chance constraints similar to the method in [10], while the other is constructed with robust tubes with bounding facets of fixed orientation. The conservatism is reduced by optimizing tube size and facet number of the tubes. However, the conservatism introduced by scaling before solving the optimization problem cannot be avoided. [32] proposes a different method in dealing with uncertainties. Instead of excluding uncertainty tube from constraints, the uncertainty tube is added to the nominal prediction in solving the optimization problem. The tube size is treated as another optimized variable, reducing the conservatism in calculating scaling at first. Nested tube methods are introduced in [16], where chance constraints on predicted states are relaxed to constraints on the cross sections of the tubes and probabilistic transitions between the tubes. By satisfying the constraints recursively, dealing with multiplicative uncertainties is avoided. But it is computationally demanding to invoke constraints on all the nested tubes. In [33], scenario approach is proposed in dealing with arbitrary distribution in LPV system. However, scenario approach suffers from having large number of samples and high conservativeness. In [34], both hard constraint and chance constraint are guaranteed by having states in the tubes and the tubes satisfying the constraints. However, computation time is severely increased by using mixed integer programming to solve chance constraints. In [35], it follows the same idea of constructing a tube (an invariant set under control). The difference is that explicit MPC is in use in [35] with all the calculations off-line.

A robust optimization approximation method based on polynomial tubes is pro-

posed in this chapter. Feasibility is guaranteed by restraining the states in the tubes and the tubes satisfying the constraints. The linear form is preserved using vertices of polyhedral tubes. Robust optimization approximation is implemented on the relaxed form of linear chance constraints, with the guarantee that the method can deal with arbitrary distribution. Recursive feasibility and stability is also guaranteed by terminal constraint and feedback policy.

3.2 Notation and Preliminary

3.2.1 Polyhedral Sets

The state tube cross-sections are defined as polyhedral sets. Compared to other convex sets (like ellipsoidal sets), the polyhedral sets have the advantage of approximating different shapes of convex sets and keeping linearity to most of the operations. The difficulty is that the complexity of the polyhedral sets is not fixed by the space dimension [34].

Definition 3.2.1. (Convex Polyhedral Set)

$$\mathcal{P}(V, g) = \{x | Vx \leq g\} = \{x | V_i x \leq g_i,\ i = 1, 2, \ldots, s\} \tag{3.2}$$

where V_i denotes the i-th row of the $s \times n$ matrix V and g_i the i-th component of the $s \times 1$ vector g.

A polyhedral set includes the origin as an interior point if and only if $g > 0$. A polyhedral set including the origin can be always represented as:

$$\mathcal{P}(V, \mathbf{1}) = \mathcal{P}(V) = \{x | Vx \leq \mathbf{1}\} = \{x | V_i x \leq 1,\ i = 1, 2, \ldots, s\} \tag{3.3}$$

where $\mathbf{1} = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}^T$ represents the column vector filled with one.

The above definition can be achieved from 3.2 by dividing both sides of the inequality by $g_i > 0$.

Its dual form is,

$$\mathcal{V}(X) = \{x | x = Xz = \sum_{i=1}^{s} X_i z_i,\ \mathbf{1}^T z = 1,\ z \geq 0\} \tag{3.4}$$

where X_i is the i-th column of the $n \times s$ matrix X, and z_i is the i-th component of the $s \times 1$ vector z. The column vector X_i can be regarded as vertices of the polyhedral set.

Definition 3.2.2. (Polytope) *A bounded polyhedral set is called polytope. We give the definition of* 0*-symmetric convex polytope as:*

$$\overline{\mathcal{P}}(V) = \{x| -\mathbf{1} \leq Vx \leq \mathbf{1}\} = \{x| -1 \leq V_i x \leq 1,\ i = 1, 2, \ldots, s\} \tag{3.5}$$

3.2.2 Polyhedral Invariant Set and Contractive Set

Definition 3.2.3. (Positive Invariant Set) *Given the system $x_{k+1} = \Phi(\omega)x_k$, a set $\mathcal{X} \subset \mathcal{R}^n$ is said to be positive invariant if for every value of $x(t) \in \mathcal{X}$ such that $x(t+1) \in \mathcal{X}$.*

Definition 3.2.4. (Pre-image Set) *For a stable system $x_{k+1} = \Phi(\omega)x_k$, consider a time instant k and the set of all states at the previous time $k-1$ for which, the condition $x(k) \in \mathcal{X}$ is satisfied for all $\omega \in W$. This set, named pre-image set, is*

$$\mathrm{Pre}_{AS}(\mathcal{X}) = \{x|\Phi(\omega)x \in \mathcal{X},\ \forall \omega \in W\} \tag{3.6}$$

Definition 3.2.5. (λ-contractive Set) *For system $x_{k+1} = \Phi(\omega)x_k$, if the following condition holds:*

$$\Psi_S(\Phi(\omega)x) \leq \lambda, \quad \forall \omega \in W,\ \forall x \in S \tag{3.7}$$

where $\Phi_S(x)$ is the Minkowski function of S, for some $0 \leq \lambda < 1$, the set S is called λ-contractive.

Definition 3.2.6. (Minkowski Function) *Given a convex and compact set S, its Minkowski function is*

$$\Psi_S(x) = \inf\{\lambda \geq 0 : x \in \lambda S\} \tag{3.8}$$

Definition 3.2.7. (λ-contractive Pre-image Set)

$$\mathrm{Pre}_{AS}^{\lambda\text{-}contractive}(\mathcal{X}) = \{x|\Psi_{\mathcal{X}}\{\Phi(\omega)x\} \leq \lambda,\ \forall \omega \in W,\ \forall x \in \mathcal{X}\} \tag{3.9}$$

Definition 3.2.8. (Spectral Radius) *Given a square matrix A its spectral radius is defined as the largest modulus of its eigenvalues $\Sigma(A) = \max\{|\lambda : \lambda \in eig(A)|\}$.*

Definition 3.2.9. (Joint Spectral Radius) *For a set of matrices the joint spectral radius of the set is defined as the supreme of the spectral radius of all possible products of the generating matrices.*

With the definitions given, we can construct the positive invariant set.

First, set notations are given here,

$$\mathcal{X}_{-1} = \text{Pre}_{AS}(\mathcal{X}) \cap \mathcal{X} \tag{3.10}$$

$$\mathcal{X}_{-k-1} = \text{Pre}_{AS}(\mathcal{X}_{-k}) \cap \mathcal{X}_{-k} \tag{3.11}$$

$$\overline{\mathcal{X}}_{-\infty} = \lim_{k\to\infty} \mathcal{X}_{-k} \tag{3.12}$$

For system $x_{k+1} = \Phi(\omega)x_k$, $\mathcal{X}$ is the convex feasible set that requires the states to satisfy. The condition that there exists a convex set $\mathcal{X}_{ini}$ such that for any $x(0) \in \mathcal{X}_{ini}, x(k) \in \mathcal{X}$ for all $k \geq 0$ is that $\overline{\mathcal{X}}_{-\infty}$ is nonempty. Moreover, any such initial set which satisfies the condition must be a subset of $\overline{\mathcal{X}}_{-\infty}$, namely:

$$\mathcal{X}_{ini} \in \overline{\mathcal{X}}_{-\infty}$$

where, $\overline{\mathcal{X}}_{-\infty}$ defines the polyhedral positive invariant set:

$$\overline{\mathcal{X}}_{-\infty} = \{x | Vx \leq \mathbf{1}\} \tag{3.13}$$

Furthermore, we want contractive property of the polyhedral set. We have,

$$\mathcal{X}_{-1} = \text{Pre}_{AS}^{\lambda\text{-contractive}}(\mathcal{X}) \cap \mathcal{X} \tag{3.14}$$

$$\mathcal{X}_{-k-1} = \text{Pre}_{AS}^{\lambda\text{-contractive}}(\mathcal{X}_{-k}) \cap \mathcal{X}_{-k} \tag{3.15}$$

$$\overline{\mathcal{X}}_{-\infty} = \lim_{k\to\infty} \mathcal{X}_{-k} \tag{3.16}$$

If, for some k,

$$\mathcal{X}_{-k} = \mathcal{X}_{-k-1} \tag{3.17}$$

Then,

$$\overline{\mathcal{X}}_{-\infty} = \mathcal{X}_{-k} \tag{3.18}$$

Table 3.1: *Algorithm for Calculation of Contractive Set*

Algorithm for calculation of λ-contractive set
1. Use the conservative feasible region as the initial polyhedral set: $$\mathcal{X}_0 = \{x \vert F^{(0)} \leq g^{(0)}\} = \mathcal{X} = \{x \vert GKx \leq b, F_s x \leq b_s\},\ k = 0$$
2. Compute the λ-contractive pre-image set: $$\mathrm{Pre}_{AS}^{\lambda\text{-contractive}}(\mathcal{X}) = \{x \vert F^{(k)} \Phi_i x \leq \lambda g^{(k)},\ \forall i\}$$
3. Compute the polyhedral set $\mathcal{X}_{-k-1} = \mathrm{Pre}_{AS}^{\lambda\text{-contractive}}(\mathcal{X}_{-k}) \cap \mathcal{X}_{-k}$ and let matrices F^{k+1} and g^{k+1} be those associated to the constraints representing set $\mathcal{X}_{-k-1}$, say $\mathcal{X}_{-k-1} = \{x \vert F^{(k+1)} x \leq g^{(k+1)}\}$
4. Check if $\mathcal{X}_{-k-1} = \mathcal{X}_{-k}$, if so, we can get $\mathcal{P}(V) = \mathcal{X}_{-k-1}$, else, let $k = k+1$ and come back to step 2.

Thus, we have the λ-contractive polyhedral set as:

$$\mathcal{P}(V) = \overline{\mathcal{X}}_{-\infty} \tag{3.19}$$

The algorithm of calculation of λ-contractive set is given in Table 3.1.

Proposition 3.2.1. *Let $\mathbb{P}_i = \{x \vert F_i x \leq b_i\}$, then $\mathbb{P}_1 \subseteq \mathbb{P}_2$ if and only if there exists an element-wise nonnegative matrix $H \geq 0$ satisfying:*

$$HF_1 = F_2, \quad Hb_1 \leq b_2 \tag{3.20}$$

3.3 Problem Formulation

As this chapter is concerned with parameter uncertainty, we only include multiplicative uncertainties in the system matrices.

$$x_{k+1} = A(\omega_k)x_k + B(\omega_k)u_k \tag{3.21}$$

The prediction of states in the control horizon is defined as,

$$x_{k+i|k} = A(\omega_{k+i-1})x_{k+i-1|k} + B(\omega_{k+i-1})u_{k+i-1|k} \tag{3.22}$$

where, $x_{k|k} = x_k$ is the current state measured.

The constraints are imposed on the predicted states and inputs,

$$\begin{cases} Gu_{k+i-1|k} \leq b, \quad \forall i > 0 \\ \text{Prob}(F_s x_{k+i|k} \leq b_s) \geq 1 - \delta_i, \quad \forall i > 0 \end{cases} \tag{3.23}$$

As origin is treated as the steady state, we can further assume $b > 0,\ b_s > 0$. The chance constraints on the states can be converted into hard constraints by requiring $\delta_i = 0$. The probability distribution of (A, B) has finite support contained within a polytope with vertices $(A^{(j)}, B^{(j)})$ for $j = 1, \ldots, r$.

$$(A, B) \in Conv(A^{(j)}, B^{(j)}) \tag{3.24}$$

where $Conv$ stands for convex hull.

It can be written equivalently as:

$$A = A_0 + \sum_{j=1}^{r} \omega_j \Delta A_j, \quad B = B_0 + \sum_{j=1}^{r} \omega_j \Delta B_j \tag{3.25}$$

where $\omega_j \in [0, 1]$ follows a certain distribution.

The MPC cost is expressed as,

$$J_k = \sum_{i=0}^{\infty} \mathbb{E}_k(x_{k+i|k}^T Q x_{k+i|k} + u_{k+i|k}^T R u_{k+i|k}) \tag{3.26}$$

We have already introduced in last chapter that the closed-loop control policy $u_k = f(x_k)$ has a better performance than arbitrary control input. However, a general control policy has a computational load exponential to the prediction horizon. State-affine feedback or disturbance-affine feedback greatly reduces the computation and shows good performance. We employ a dual mode strategy here [34]:

$$\begin{cases} u_{k+i|k} = Kx_{k+i|k} + c_{k+i|k}, \quad 1 \leq i \leq N-1 \\ u_{k+i|k} = Kx_{k+i|k}, \quad i \geq N \end{cases} \tag{3.27}$$

The control policy follows an state-affine feedback. $u = Kx$ is designed to be stabilizing and optimal for the objective function 3.26 in the absence of constraints. K is calculated off-line and will be explained in details next section. Perturbations $c_{k+i|k}$ is calculated on-line to optimize the control performance with the constraints present.

So the chance constraint problem in LPV system is formulated as:

$$\begin{aligned}
&\min J_k \\
&\text{s.t.} \\
&x_{k+1} = A(\omega_k)x_k + B(\omega_k)u_k \\
&\begin{cases} u_{k+i|k} = Kx_{k+i|k} + c_{k+i|k}, & 1 \le i \le N \\ u_{k+i|k} = Kx_{k+i|k}, & i > N \end{cases} \\
&Gu_{k+i-1|k} \le b, \qquad \forall i \ge 1 \\
&\text{Prob}(F_s x_{k+i|k} \le b_s) \ge 1 - \delta_i, \quad \forall i \ge 1
\end{aligned} \tag{3.28}$$

3.4 State Tube MPC

The offline calculation is started with feedback control policy:

$$u_{k+i} = Kx_{k+i} + c_{k+i}, \quad 0 \le i \le N-1 \tag{3.29}$$

where, K is calculated based on finding Quadratic Lyapunov function.

Consider the system now given by:

$$x_{k+1} = \Phi x_k + Bu_k \tag{3.30}$$

where, $\Phi = A + BK, u_k = c_k$.

3.4.1 Stability of Unconstrained LPV System

The stability guarantee of unconstrained LPV system is given here. Consider $u_k = Kx_k$ as the feedback control law for $x_{k+1} = A(\omega_k)x_k + B(\omega_k)u_k$. It is known from Laypunov stability theorem for discrete system that the following condition can guarantee

stability:

$$(A+BK)^T P(A+BK) - P \prec 0, \tag{3.31}$$

$$P \succ 0 \tag{3.32}$$

For the system associated with parameter uncertainty: $\Phi = \Phi_0 + \sum_{j=1}^{r} \omega_j \Delta\Phi_j,\ \Phi_0 = A_0 + B_0 K,\ \Delta\Phi_j = \Delta A_j + \Delta B_j K$. For any vector x, the set of all vectors $y(\omega) = \Phi(\omega)x$ is a polytope and, since the norm is a convex function, the expression

$$\sqrt{x^T \Phi(\omega)^T P \Phi(\omega) x} = ||\Phi(\omega)x||_P \tag{3.33}$$

is thought of as a function of ω, and reaches the maximum on one of the vertices. [9]

So we have a equivalent condition by requirement on the polyhedral vertices of the uncertainty set:

$$(A_j + B_j K)^T P(A_j + B_j K) - P \prec 0, \quad j = 1, 2, \ldots, r \tag{3.34}$$

$$P \succ 0 \tag{3.35}$$

By pre and post multiplying both sides by $Q = P^{-1}$ and by defining $KQ = R$, we have:

$$(QA_j^T + R^T B_j^T) Q^{-1} (A_j Q + B_j R) - Q \prec 0, \quad j = 1, 2, \ldots, r \tag{3.36}$$

$$Q \succ 0 \tag{3.37}$$

It is known to be equivalent to the set of linear matrix inequalities (LMIs), [10]

$$\begin{bmatrix} Q & QA_j^T + R^T B_j^T \\ A_j Q + B_j R & Q \end{bmatrix} \succ 0, \quad j = 1, 2, \ldots, r \tag{3.38}$$

$$Q \succ 0 \tag{3.39}$$

A feasible solution of Q can be calculated by semidefinite programming based on some minimization criteria, like quadratic objective. The feedback parameter is thus obtained: $K = RQ^{-1}$.

3.4.2 State Tube

The state tube cross-sections are constructed as the polyhedral set as:

$$T_k = \{x | Vx \leq \alpha_k \mathbf{1}\}, \quad \forall k \geq 0 \tag{3.40}$$

The tube size is parameterized by the variable α_k. Parameter V is calculated off-line to guarantee recursive feasibility, with a design factor λ.

V is chosen such that the polyhedral set $\{x | Vx \leq \mathbf{1}\}$ defines a λ-contractive set for system $x_{k+1} = \Phi(\omega)x_k$, with $\lambda < 1$. It is equivalent to having positive invariant set $\{x | Vx \leq \mathbf{1}\}$ for system $x_{k+1} = \dfrac{\Phi(\omega)}{\lambda}x_k$.

3.4.3 On-line Optimization

Here the constraints on the states are relaxed by requiring two conditions:

1. states belong to the tubes

2. constraints restrictions on the tubes

The state tubes have fixed structure defined from the polyhedral set. Set size is being optimized on-line to guarantee the two required conditions.

The state tube will be used to refer to the sequence of sets $\{T_k\}$. The MPC controller, at each time instant, should generate a state tube $\{T_k\}$ and a sequence of inputs $\{c_k\}$ such that, for all $k \geq 0$:

$$T_k \subseteq \{x_k | Gu_k \leq b\} \quad \forall k \tag{3.41}$$

$$T_k \subseteq \{x_k | \text{Prob}(F_s x_{k+1} \leq b_s) \geq p\} \quad \forall k \tag{3.42}$$

$$T_k \subseteq \{x_k | Ax_k + Bu_k \in T_{k+1}\}, \quad \forall k \tag{3.43}$$

$$x_0 \in T_0 \tag{3.44}$$

Inequality 3.41 is the hard constraint imposed on the tube in the control horizon, 3.42 defines the chance constraint imposed on the tube, and 3.43 requires the states to be confined in the tubes recursively (replaced by $x_k \in T_k$), 3.44 requires the initial states to be inside the initial tube.

The inequality constraints 3.41, 3.42 and 3.43 are equivalent to 3.45, 3.46 and 3.47:

$$\begin{cases} T_k \subseteq \{x_k | GKx_k + Gc_k \leq b\}, & 0 \leq k \leq N-1 \\ T_k \subseteq \{x_k | GKx_k \leq b\}, & k \geq N \end{cases} \tag{3.45}$$

$$\begin{cases} T_k \subseteq \{x_k | \text{Prob}(F_s\Phi(\omega)x_k + F_sBc_k \leq b_s) \geq p\}, & 0 \leq k \leq N-1 \\ T_k \subseteq \{x_k | \text{Prob}(F_s\Phi(\omega)x_k \leq b_s) \geq p\}, & k \geq N \end{cases} \tag{3.46}$$

$$\begin{cases} T_k \subseteq \{x_k | \Phi(\omega)x_k + Bc_k \in T_{k+1}\}, & 0 \leq k \leq N-1 \\ T_k \subseteq \{x_k | \Phi(\omega)x_k \in T_{k+1}\}, & k \geq N \end{cases} \tag{3.47}$$

Tube Constraints

With Proposition 3.2.1, we can change the constraints into solvable formats.

First, we denote $\Phi^{(j)} = \Phi_0 + \Delta\Phi_j$. Then we would have, $\Phi = \sum_{j=1}^{r} \omega_j \Phi^{(j)},\ \sum_{j=1}^{r} \omega_j = 1,\ \omega_j \geq 0, \forall j$.

For the invariant set constraints 3.47, we can get the recursive constraints:

$$\begin{cases} Vx_k \leq \alpha_k \mathbf{1} \\ V(\Phi x_k + Bc_k) \leq \alpha_{k+1}\mathbf{1} \end{cases} \qquad 0 \leq k \leq N-1 \tag{3.48}$$

$$\begin{cases} Vx_k \leq \alpha_k \mathbf{1} \\ V\Phi x_k \leq \alpha_{k+1}\mathbf{1} \end{cases} \qquad k \geq N \tag{3.49}$$

From $\Phi = \sum_{j=1}^{r} \omega_j \Phi_j$, we know $V\Phi x_k = \sum_{j=1}^{r} \omega_j V\Phi^{(j)} x_k$. As $0 \leq \omega_j < 1$, we can obtain $\max V\Phi x_k = \max_{j=1}^{r} V\Phi^{(j)} x_k$.

So we can formulate the recursive constraints 3.48, 3.49 as:

$$\begin{cases} Vx_k \leq \alpha_k \mathbf{1} \\ V\Phi^{(j)}x_k + VBc_k \leq \alpha_{k+1}\mathbf{1} \end{cases} \qquad 0 \leq k \leq N-1,\ \forall j \tag{3.50}$$

$$\begin{cases} Vx_k \leq \alpha_k \mathbf{1} \\ V\Phi^{(j)}x_k \leq \alpha_{k+1}\mathbf{1} \end{cases} \qquad k \geq N,\ \forall j \tag{3.51}$$

Let nonnegative matrix $H^{(j)}$ satisfy the condition $H^{(j)}V = V\Phi^{(j)}$, we can have the inequality constraint formulated as:

$$\begin{cases} Vx_k \leq \alpha_k \mathbf{1} \\ H^{(j)}Vx_k + VBc_k \leq \alpha_{k+1}\mathbf{1} \end{cases} \qquad 0 \leq k \leq N-1,\ \forall j \tag{3.52}$$

$$\begin{cases} Vx_k \leq \alpha_k \mathbf{1} \\ H^{(j)} V x_k \leq \alpha_{k+1} \mathbf{1} \end{cases} \qquad k \geq N,\ \forall j \tag{3.53}$$

From Proposition 3.2.1, we know that with the condition $H^{(j)}V = V\Phi^{(j)}$ and $Vx_k \leq \alpha_k \mathbf{1}$, if $H^{(j)}\alpha_k \mathbf{1} + VB^{(j)}c_k \leq \alpha_{k+1}\mathbf{1}$ is satisfied, the inequality $H^{(j)}Vx_k + VBc_k \leq \alpha_k \mathbf{1}$ will also be satisfied. So the states requirement can be relaxed into the initial condition and recursive constraint:

$$\begin{cases} Vx_0 \leq \alpha_0 \mathbf{1} \\ H^{(j)}\alpha_k \mathbf{1} + VB^{(j)}c_k \leq \alpha_{k+1}\mathbf{1}, \quad 0 \leq k \leq N-1, \forall j \\ H^{(j)}\alpha_k \mathbf{1} \leq \alpha_{k+1}\mathbf{1}, \quad k \geq N, \forall j \end{cases} \tag{3.54}$$

where,

$$H^{(j)}V = V\phi^{(j)} \tag{3.55}$$

We will give here one choice of selecting $H^{(j)}$ off-line:

$$H^{(j)} = \arg\min_{H^{(j)}} \{\| H^{(j)} \|_1 | H^{(j)}V = V\Phi^{(j)}, H^{(j)} \geq 0,\ \forall j\} \tag{3.56}$$

where one-norm is defined as $\| A \|_1 = \max_{1 \leq i \leq m}(\sum_{j=1}^{n} |A_{ij}|)$.

The purpose of minimizing the one-norm is to relax the constraints applied on line. The user can of course choose other methods to select $H^{(j)}$ to get more optimal result.

Hard Constraints

For the hard constraints 3.45, by letting $HV = GK$ and following similar procedures, we would yield the result:

$$\begin{cases} H\alpha_k \mathbf{1} + Gc_k \leq b, \quad 0 \leq k \leq N-1 \\ H\alpha_k \mathbf{1} \leq b, \quad k \geq N \end{cases} \tag{3.57}$$

where, $HV = GK$.

$$H = \arg\min_{H} \{\| H \|_1 | HV = GK,\ H \geq 0\} \tag{3.58}$$

Soft Constraints

For soft constraints 3.46:

$$\begin{cases} \text{Prob}(F_s\Phi_k x_k + F_s B_k c_k \le b_s) \ge p, & 0 \le k \le N-1 \\ \text{Prob}(F_s\Phi_k x_k \le b_s) \ge p, & k \ge N \end{cases} \tag{3.59}$$

For polyhedral set $\mathcal{P}(V) = \{x|Vx \le \mathbf{1}\}$, and its dual $\mathcal{V}(V_g) = \{x|x = V_g z, \mathbf{1}^T z = 1, z \ge 0\}$. $V_g(i)$ is the i_{th} column of V_g. In the state tubes, by adding the set size, we have $T_k = \{x_k|Vx_k \le \alpha_k \mathbf{1}\}$, its dual form $T_k = \{x_k|x_k = \alpha_k V_g z, \mathbf{1}^T z = 1, z \ge 0\}$.

So the states inside the tube can also be written as: $x_k = \sum_{i=1}^m \lambda_i V_g(i)\alpha_k,\ \sum_{i=1}^m \lambda_i = 1,\ \forall \lambda_i \ge 0$. Thus we know,

$$F_s\Phi_k x_k = F_s\Phi_k \sum_{i=1}^m \lambda_i V_g(i)\alpha_k = \sum_{i=1}^m \lambda_i F_s\Phi_k V_g(i)\alpha_k \le \max_i F_s\Phi_k V_g(i)\alpha_k \quad \sum_{i=1}^m \lambda_i = 1,\ \forall \lambda_i \ge 0 \tag{3.60}$$

where,

$$\max_i F_s\Phi_k V_g(i)\alpha_k = \max(\begin{bmatrix} F_s\Phi_k V_g(1)\alpha_k & \dots & F_s\Phi_k V_g(i)\alpha_k & \dots & F_s\Phi_k V_g(m)\alpha_k \end{bmatrix}, 1) \tag{3.61}$$

where, $\max(A, 1)$ is defined as maximization of matrix A over the 1^{st} dimension (row). For example,

$$\max(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \end{bmatrix}, 1) = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \tag{3.62}$$

The soft constraints 3.59 are then relaxed to, (as a matter of simplification, the derivation for $k \ge N$ is omitted)

$$\text{Prob}[\max_i F_s\Phi_k V_g(i)\alpha_k + F_s B_k c_k \le b_s] \ge p,\ 0 \le k \le N-1 \tag{3.63}$$

It is equivalent to,

$$\text{Prob}[F_s\Phi_k V_g(i)\alpha_k + F_s B_k c_k \le b_s,\ \forall i] \ge p,\ 0 \le k \le N-1 \tag{3.64}$$

The inequality inside the probability constraint becomes:

$$F_s(\Phi_0 + 0.5\sum_{j=1}^{r}\Delta\Phi_j)V_g(i)\alpha_k + F_s(B_0 + 0.5\sum_{j=1}^{r}\Delta B_j)c_k - b_s +$$

$$\left[F_s\frac{\Delta\Phi_1}{2}V_g(i)\alpha_k + F_s\frac{\Delta B_1}{2}c_k \quad \dots \quad F_s\frac{\Delta\Phi_r}{2}V_g(i)\alpha_k + F_s\frac{\Delta B_r}{2}c_k\right]\begin{bmatrix}2\omega_{1,k}-1\\ \vdots \\ 2\omega_{r,k}-1\end{bmatrix} \le \mathbf{0}, \quad i = 1,\dots,m \tag{3.65}$$

For time $0 \le k \le N-1$, let:

$$\xi_k = \begin{bmatrix}2\omega_1 - 1\\ \vdots \\ 2\omega_r - 1\end{bmatrix},\ y^i_{0,k} = F_s(\Phi_0 + 0.5\sum_{j=1}^{r}\Delta\Phi_j)V_g(i)\alpha_k + F_s(B_0 + 0.5\sum_{j=1}^{r}\Delta B_j)c_k - b_s$$

$$\mathbf{y}^i_k = \left[F_s\frac{\Delta\Phi_1}{2}V_g(i)\alpha_k + F_s\frac{\Delta B_1}{2}c_k \quad \dots \quad F_s\frac{\Delta\Phi_r}{2}V_g(i)\alpha_k + F_s\frac{\Delta B_r}{2}c_k\right]^T$$

The probability constraint can be derived as:

$$\text{Prob}[y^i_{0,k} + {\mathbf{y}^i_k}^T\xi_k \le \mathbf{0}, \quad i = 1,\dots,m] \ge p, \quad 0 \le k \le N-1 \tag{3.66}$$

Use q to denote the index of soft constraint, and Q to denote the number of soft constraints (number of rows in b_s).

Let $y^{i,q}_{0,k} = r_q y^i_{0,k}, {\mathbf{y}^{i,q}_k}^T = r_q {\mathbf{y}^i_k}^T$.

We have,

$$\text{Prob}[\max_{i,q}(y^{i,q}_{0,k} + {\mathbf{y}^{i,q}_k}^T\xi_k) \le 0, \quad i = 1,\dots,m, q = 1,\dots,Q] \ge p, \quad 0 \le k \le N-1 \tag{3.67}$$

Apply the RO method here in solving the chance constraints, we have:

$$\begin{cases} \phi_k + \sum_j \gamma_k^{i,q} \leq \delta_k t_k \\ \phi_k \geq u_{0k} + \max_{\xi \in \mathbb{U}_k} \xi^T \mathbf{u}_k \\ \phi_k \geq 0 \\ \gamma_k^{i,q} \geq v_{0k}^{i,q} + \max_{\xi \in \mathbb{U}_k} \xi^T v_k^{i,q} \quad \forall i, q \\ \gamma_k^{i,q} \geq 0 \quad \forall i, q \\ u_{0k} = \omega_{0k} + t_k \\ \mathbf{u}_k = \boldsymbol{\omega}_k \\ v_{0k}^{i,q} = y_{0k}^{i,q} - \omega_{0k} \quad \forall j \\ \mathbf{v}_k^{i,q} = \mathbf{y}_k^{i,q} - \boldsymbol{\omega}_k \quad \forall i, q \end{cases} \qquad 0 \leq k \leq N-1 \tag{3.68}$$

For time $k \geq N$, let:

$$\xi_k = \begin{bmatrix} 2\omega_1 - 1 \\ \vdots \\ 2\omega_r - 1 \end{bmatrix}, \; y_{0,k}^i = F_s(\Phi_0 + 0.5 \sum_{j=1}^{r} \Delta\Phi_j) V_g(i) \alpha_k - b_s$$

$$\mathbf{y}_k^i = \begin{bmatrix} F_s \frac{\Delta\Phi_1}{2} V_g(i) \alpha_k & \dots & F_s \frac{\Delta\Phi_r}{2} V_g(i) \alpha_k \end{bmatrix}^T$$

The RO approximated constraints are also constructed in the same form as in 3.68.

Terminal Constraints

The second mode is defined for the time $k \geq N$, which is infinite time horizon. A finite period is definitely preferred in solving the problem. We want to have a terminal time instant M such that $M \geq N$ and there always exist tubes after time M which satisfy the feasibility constraint of both hard constraint 3.57 and soft constraint 3.71:

$$H \alpha_k \mathbf{1} \leq b, \quad k \geq M \tag{3.69}$$

$$\text{Prob}[H_s(i) \alpha_k \mathbf{1} \leq b_s, \; \forall i] \geq p, \quad k \geq M \tag{3.70}$$

To find the time instant M which satisfies the conditions, we need to give some proofs here.

For $k \geq N$, the soft constraints can also be defined as:

$$\text{Prob}[H_s(i)\alpha_k \leq b_s,\ \forall i] \geq p, \quad k \geq N \tag{3.71}$$

where,

$$H_s(i) = F_s(\Phi_0 + \frac{1}{2}\sum_{j=1}^{r}\Delta\Phi_j)V_g(i) + \frac{F_s}{2}\left[\Delta\Phi_1 V_g(i) \quad \Delta\Phi_2 V_g(i) \quad \dots \quad \Delta\Phi_r V_g(i)\right]\xi_k$$

In [35]: the set $\mathcal{P}(V, \mathbf{1}) = \{x | Vx \leq \mathbf{1}\}$ is a positively invariant set of system $x_{k+1} = \Phi x_k$ if there exists non-negative matrix $H^{(j)}$ that,

$$H^{(j)}V = V\Phi^{(j)}, \quad \forall j \tag{3.72}$$

$$(H^{(j)} - I)\mathbf{1} \leq \mathbf{0}, \quad \forall j \tag{3.73}$$

where the inequality constraints require to be satisfied element-wise.

Recall the definition of joint spectral radius in 3.2.9. Design positive invariant set $\mathcal{P}(V, \mathbf{1})$ for the system $x_{k+1} = \dfrac{\Phi}{\lambda}x_k$.

Thus, we have:

$$\begin{aligned} &\frac{H^{(j)}}{\lambda}V = V\frac{\Phi^(j)}{\lambda}, \quad \forall j \\ &(\frac{H^{(j)}}{\lambda} - I)\mathbf{1} \leq \mathbf{0}, \quad \forall j \end{aligned} \tag{3.74}$$

Thus, we have:

$$H^{(j)}\mathbf{1} \leq \lambda\mathbf{1}, \quad \forall j \tag{3.75}$$

$$(H^{(j)}\mathbf{1})_m \leq \lambda, \quad \forall j \tag{3.76}$$

where m denotes the row number.

As $H^{(j)}$ is nonnegative, we have $|H^{(j)}|_\infty = \max_m H^{(j)}\mathbf{1} \leq \lambda$. So we can give the conclusion that:

For the λ-contractive set $\{x | Vx \leq \mathbf{1}\}$ of system $x_{k+1} = \Phi(\omega)x_k$ and satisfaction of $H^{(j)}V = V\Phi^{(j)}$, we have $|H^{(j)}|_\infty \leq \lambda$ for all j.

Then we can find the condition of the existence of M which satisfies the recursive feasibility. We can show that as long as $M \geq N$, there exists tube size α_k which satisfies the above requirements. Proof is given as the following:

Construct the sequence of $\widetilde{\alpha}_k$ which satisfies the tube conditions:

$$\widetilde{\alpha}_k = \alpha_k, \quad 0 \leq k \leq M \tag{3.77}$$

$$\widetilde{\alpha}_{k+1} = \max_{j,m}[(H^{(j)})_m \alpha_k \mathbf{1}], \quad k \geq M+1 \tag{3.78}$$

Then we can get,

$$\begin{aligned} \widetilde{\alpha}_{k+1} &\leq \max_{j,m} |(H^{(j)})_m|_\infty \widetilde{\alpha}_k \\ &= \max_{j} |H^{(j)}|_\infty \widetilde{\alpha}_k \\ &= \lambda \widetilde{\alpha}_k \end{aligned} \tag{3.79}$$

As we know,

$$H\tilde{\alpha}_M \mathbf{1} \leq b \tag{3.80}$$

$$\text{Prob}[H_s(i)\tilde{\alpha}_M \mathbf{1} \leq b_s, \ \forall i] \geq p \tag{3.81}$$

Assume $b \geq 0$, from 3.79, 3.80, we have:

$$H\tilde{\alpha}_{M+l}\mathbf{1} \leq \lambda^l H\tilde{\alpha}_M \mathbf{1} \leq \lambda^l b \tag{3.82}$$

Assume $b_s \geq 0$,from 3.79, 3.81, we have:

$$H_s(i)\tilde{\alpha}_{M+l}\mathbf{1} \leq \lambda^l H_s(i)\tilde{\alpha}_M \mathbf{1} \leq H_s(i)\tilde{\alpha}_M \mathbf{1}, \quad \forall i \tag{3.83}$$

$$\text{Prob}[H_s(i)\tilde{\alpha}_M \mathbf{1} \leq b_s, \ \forall i] \geq p \tag{3.84}$$

Thus, we can get:

$$\text{Prob}[H_s(i)\tilde{\alpha}_{M+l} \mathbf{1} \leq b_s, \ \forall i] \geq p \tag{3.85}$$

So we know that as long as $M \geq N$, there always exists solution α_k with $k \geq M$ that satisfies the feasibility constraint. It is equivalent to saying as long as N is large enough so that $c_N = 0$, recursive feasibility is guaranteed.

In summary, we have the constraint as:

Tube Constraints:

$$\begin{gathered} Vx_0 \leq \alpha_0 \mathbf{1} \\ H^{(j)}\alpha_k \mathbf{1} + VB^{(j)}c_k \leq \alpha_{k+1}\mathbf{1}, \quad 0 \leq k \leq N-1, \forall j \\ H^{(j)}\alpha_N \mathbf{1} \leq \alpha_{N+1}\mathbf{1}, \quad \forall j \end{gathered} \tag{3.86}$$

Hard Constraints:

$$\begin{aligned} H\alpha_k\mathbf{1} + Gc_k \leq b, \quad 0 \leq k \leq N-1 \\ H\alpha_N\mathbf{1} \leq b \end{aligned} \tag{3.87}$$

Soft Constraints:

$$\begin{cases} \phi_k + \sum_j \gamma_k^{i,q} \leq \delta_k t_k \\ \phi_k \geq u_{0k} + \max_{\xi\in\mathbb{U}_k} \xi^T \mathbf{u}_k \\ \phi_k \geq 0 \\ \gamma_k^{i,q} \geq v_{0k}^{i,q} + \max_{\xi\in\mathbb{U}_k} \xi^T v_k^{i,q} \quad \forall i,q \\ \gamma_k^{i,q} \geq 0 \quad \forall i,q \\ u_{0k} = \omega_{0k} + t_k \\ \mathbf{u}_k = \boldsymbol{\omega}_k \\ v_{0k}^{i,q} = y_{0k}^{i,q} - \omega_{0k} \quad \forall j \\ \mathbf{v}_k^{i,q} = \mathbf{y}_k^{i,q} - \boldsymbol{\omega}_k \quad \forall i,q \end{cases} \qquad \forall k \tag{3.88}$$

where,

$$y_{0,k}^i = F_s(\Phi_0 + 0.5\sum_{j=1}^{r}\Delta\Phi_j)V_g(i)\alpha_k + F_s(B_0 + 0.5\sum_{j=1}^{r}\Delta B_j)c_k - b_s$$

$$\mathbf{y}_k^i = \left[F_s\frac{\Delta\Phi_1}{2}V_g(i)\alpha_k + F_s\frac{\Delta B_1}{2}c_k \quad \dots \quad F_s\frac{\Delta\Phi_r}{2}V_g(i)\alpha_k + F_s\frac{\Delta B_r}{2}c_k\right]^T,\ 0 \leq k \leq N-1$$

$$y_{0,k}^i = F_s(\Phi_0 + 0.5\sum_{j=1}^{r}\Delta\Phi_j)V_g(i)\alpha_k - b_s$$

$$\mathbf{y}_k^i = \left[F_s\frac{\Delta\Phi_1}{2}V_g(i)\alpha_k \quad \dots \quad F_s\frac{\Delta\Phi_r}{2}V_g(i)\alpha_k\right]^T,\ k = N$$

Objective Function

The objective function is formulated as a quadratic function of expected states and inputs:

$$\min_{\alpha_k, c_k}((\mathbb{E}\mathbf{x})^T Q_{\mathbf{x}}(\mathbb{E}\mathbf{x}) + (\mathbb{E}\mathbf{u})^T Q_{\mathbf{u}}(\mathbb{E}\mathbf{u})) \tag{3.89}$$

where,

$$\mathbf{x} = \begin{bmatrix} x_{k+1|k} \\ x_{k+2|k} \\ \vdots \\ x_{k+N|k} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_{k|k} \\ u_{k+1|k} \\ \vdots \\ u_{k+N-1|k} \end{bmatrix} \tag{3.90}$$

Here we also define:

$$\mathbb{E}x_{k+i+1|k} = A_0\mathbb{E}x_{k+i|k} + B_0u_{k+i|k}, \quad i \geq 0 \tag{3.91}$$

$$\mathbb{E}x_{k|k} = x_k \tag{3.92}$$

Thus for the expected augmented states, we have:

$$\mathbb{E}\mathbf{x} = Gx_k + H\mathbf{u}, \quad \mathbb{E}\mathbf{u} = K\mathbb{E}\mathbf{x} + \mathbf{c} \tag{3.93}$$

where,

$$G = \begin{bmatrix} A_0 \\ A_0^2 \\ \vdots \\ A_0^N \end{bmatrix}, H = \begin{bmatrix} B_0 & 0 & \cdots & 0 \\ A_0B_0 & B_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_0^{N-1}B_0 & A_0^{N-2}B_0 & \cdots & B_0 \end{bmatrix},$$

$$K_M = \begin{bmatrix} K & 0 & \cdots & 0 \\ 0 & K & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & K \end{bmatrix}, \mathbf{c} = \begin{bmatrix} c_{k|k} \\ c_{k+1|k} \\ \vdots \\ c_{k+N-1|k} \end{bmatrix}$$

As,

$$\mathbb{E}\mathbf{x} = (I - HK_M)^{-1}Gx_k + (I - HK_M)^{-1}H\mathbf{c} \tag{3.94}$$

$$\mathbb{E}\mathbf{u} = K(I - HK_M)^{-1}Gx_k + [K(I - HK_M)^{-1}H\mathbf{c} + \mathbf{c}]\mathbf{c} \tag{3.95}$$

By letting,

$$Q_{\mathbf{x}x} = (I - HK_M)^{-1}G,\ Q_{\mathbf{xc}} = (I - HK_M)^{-1}H, \tag{3.96}$$

$$Q_{\mathbf{u}x} = K_M(I - HK_M)^{-1}G,\ Q_{\mathbf{u}c} = [K(I - HK_M)^{-1}H\mathbf{c} + \mathbf{c}] \tag{3.97}$$

The objective function is equivalent to,

$$\min_{\alpha_k, c_k} \mathbf{c}^T Q_{q\mathbf{c}}\mathbf{c} + 2Q_l\mathbf{c} + m \tag{3.98}$$

where, $Q_{q\mathbf{c}} = Q_{\mathbf{xc}}^T Q_{\mathbf{x}} Q_{\mathbf{xc}} + Q_{\mathbf{uc}}^T Q_{\mathbf{u}} Q_{\mathbf{uc}}$, $Q_l = x^T(Q_{\mathbf{x}x}^T Q_{\mathbf{x}} Q_{\mathbf{xc}} + Q_{\mathbf{u}x}^T Q_{\mathbf{u}} Q_{\mathbf{uc}})$, $m = x_k^T(Q_{\mathbf{x}x}^T Q_{\mathbf{x}} Q_{\mathbf{x}x} + Q_{\mathbf{u}x}^T Q_{\mathbf{u}} Q_{\mathbf{u}x})x_k$.

3.5 Case Study

We give a case study here utilizing the method of state tube MPC method based on RO approximation:

The state space is given as:

$$x_{t+1} = A(\omega)x_t + B(\omega)u_t$$

$$A(\omega) = A_0 + \sum_{i=1}^{3} \omega_i \Delta_A^{(i)},\ B(\omega) = B_0 + \sum_{i=1}^{3} \omega_i \Delta_B^{(i)}$$

$$A_0 = \begin{bmatrix} 0.3221 & -0.6774 \\ 0.8659 & -1.7650 \end{bmatrix},\ B_0 = \begin{bmatrix} -0.5045 \\ -1.9386 \end{bmatrix}$$

$$\Delta_A^{(1)} = \begin{bmatrix} 0 & 0.05 \\ -0.05 & 0 \end{bmatrix},\ \Delta_A^{(2)} = \begin{bmatrix} 0.01 & 0 \\ 0.05 & -0.01 \end{bmatrix},\ \Delta_A^{(3)} = \begin{bmatrix} -0.01 & -0.05 \\ 0 & 0.01 \end{bmatrix}$$

$$\Delta_B^{(1)} = \begin{bmatrix} 0.0478 \\ 0.0589 \end{bmatrix},\ \Delta_B^{(2)} = \begin{bmatrix} -0.0478 \\ -0.0589 \end{bmatrix},\ \Delta_B^{(3)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ N = 5$$

The matrices defining the constraints 3.23 are:

$$G = \begin{bmatrix} -0.1 \\ 0.1 \end{bmatrix},\ F_s = \begin{bmatrix} 9 & 3 \\ -9 & -3 \\ 0 & 6.67 \end{bmatrix},\ b = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ b_s = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

The matrices defining the objective 3.26 is:

$$Q_{\mathbf{x}} = 10,\ Q_{\mathbf{u}} = 0$$

To use RO based tube method, objective function is reformulated as in 3.98, constraints are reformulated as a composition of tube constraints 3.86, hard constraints 3.87 and soft constraints 3.88.

For simulation results, λ-contractive set $\mathcal{P}(V)$ 3.19 is plotted in Fig 3.1. System trajectory with probability requirement 70% and 90% are drawn in Fig 3.2. Detailed results are recorded in Table 3.2.

From the results, we can tell that: the probability constraint is satisfied within the tolerance. The objective function performs better with less restraint on the probability satisfaction. Different trajectories under different probability requirements are plotted in Fig 3.2a, 3.2b. The constraint violation happens at step 2. The state comes to the original at step 5.

Table 3.2: *Comparison in LPV System Case Study*

Required Probability $P = 0.7$	Required Probability $P = 0.9$
Monte Carlo Simulated Probability $P_{MC} = 0.69$	Monte Carlo Simulated Probability $P_{MC} = 0.8999$
Objective Value $OBJ = 0.0399$	Objective Value $OBJ = 0.0495$
Tube Size $\alpha =$ [0.9972 1.3445 1.6729 1.6123 1.7303]	Tube Size $\alpha =$ [0.9972 1.2695 1.6621 1.5844 1.6905]
Iteration Time $N_{iter} = 5$	Iteration Time $N_{iter} = 3$

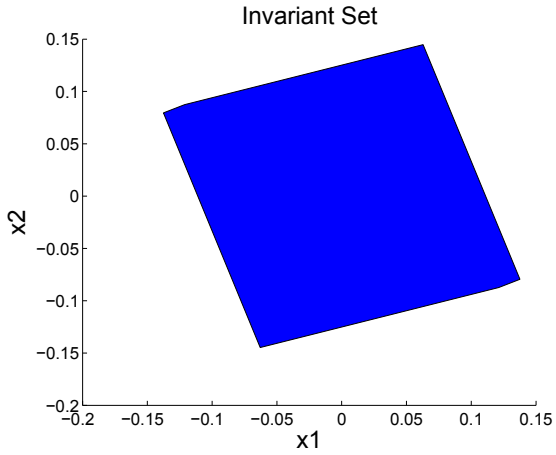
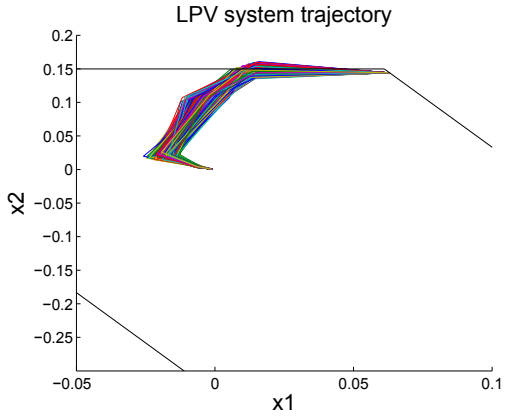
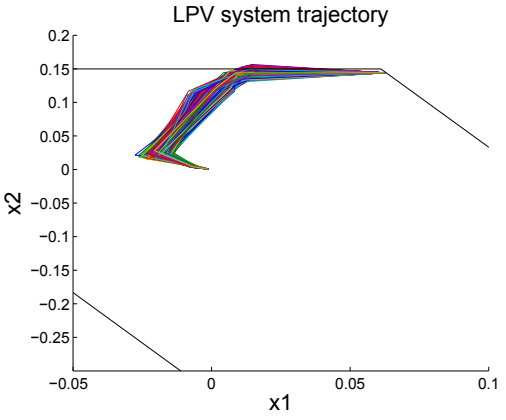


Figure 3.1: *Invariant Polyhedral Set*



(a) *System Trajectory with 70 % Probability Requirement*



(b) *System Trajectory with 90 % Probability Requirement*

Figure 3.2: *System Trajectory with Tube MPC*

3.6 Conclusions

This chapter extended RO based CCMPC to systems with parameter uncertainty. However, RO method cannot be directly applied due to multiplicative uncertainties. Stochastic tubes are introduced to help solving the problem. By imposing constraints on the tubes, multiplicative uncertainties are reduced. The proposed method also guarantees a recursive feasibility and stability. However, this method can be conservative in solutions and are sensitive to initial states. It also requires uncertainty distribution to be bounded. A numerical case study is shown at the end with probability requirement satisfied.

Chapter 4

Application of CCMPC in SAGD Process

4.1 Introduction

Alberta is rich in oil sands deposit. However, oil sands extraction is different from conventional oil development. The two main methods for oil sands extractions are: open mining and steam assisted gravity drainage (SAGD) process. Open mining deals mainly with oil sands in shallow areas. SAGD is designed for oil sands buried deep under ground. SAGD was introduced by Butler in the 1970s [36]. He also brought the concept of steam chamber and developed analytical models [37]. The model described the relationship between chamber shape and time, bitumen drainage rate and chamber height. The concepts and related theories Butler built for the SAGD process greatly stimulated the technology to be applied in industry. It is now one of the premier recovery methods for heavy oil and oil sands [38].

Oil sand is a mixture of sand, clay and other minerals, water and bitumen, which is heavy and extremely viscous. Traditional pumping method cannot extract oil sands from deep underground. However, as the temperature goes up, bitumen viscosity will drop significantly, which makes it possible for the bitumen to be extracted. SAGD is a typical thermal recovery method efficient in extracting oil sands from deep underground. A typical SAGD process refers to a pair of wells draining horizontally. The well pair includes one injection well and one production well, with a distance

of 3 to 5 meters in between each other. The injection well injects hot steam into the horizontal wellbore, heating bitumen and reducing its viscosity. The production well extracts bitumen drained into the lower wellbore. During the process, steam chamber expands both vertically and horizontally. Based on shape of steam chamber and production amount, the process is separated into four periods: start-up, ramp-up, peak and blow-down. Start-up is the period which requires establishment of oil mobility between the two wells. The temperature between the well pair is sufficiently high which makes the oil flowable enough to gather around the production well. Oil mobility is achieved by circulating steam in each of the well. It usually takes three to six months. During the ramp-up period, which is around half a year to two years, the chamber expands quickly. More steam is injected and more bitumen is produced. Then in peak period, production rises to maximum amount and chamber touches the top of pay zone. The peak period lasts for three to six years with maximum injection and production rate. After the peak period, energy usage efficiency significantly drops and production keeps falling down. It takes three to five years for the final period. For a reservoir with SAGD implemented, it takes ten to twenty years for the whole process to complete. The whole plant combines fuel and gas system, steam generator, injection and production system, oil treating equipment and water recycling system. In this thesis, only the underground reservoir part is being studied. For later mentioning of the SAGD process, it is meant to be the underground part.

Many studies and industry practices have shown that it is important to keep the steam chamber well developed so that neither liquid around producer accumulates nor is steam produced. In other words, maintain the liquid-steam interface in between the horizontal injector and producer. It is called steam trap by Butler first, using analogy of the operation in steam-heated radiator [39], [40]. Without using steam trap, various problems can be introduced. If the condensate and bitumen are removed quickly from producer, steam has a tendency to flow directly to producer, leading to steam breakthrough. Instead of transferring heat to the reservoir, steam is greatly wasted by being passed to producer directly. If the condensate and bitumen are not removed in time, it leads to liquid-steam interface higher than injector well.

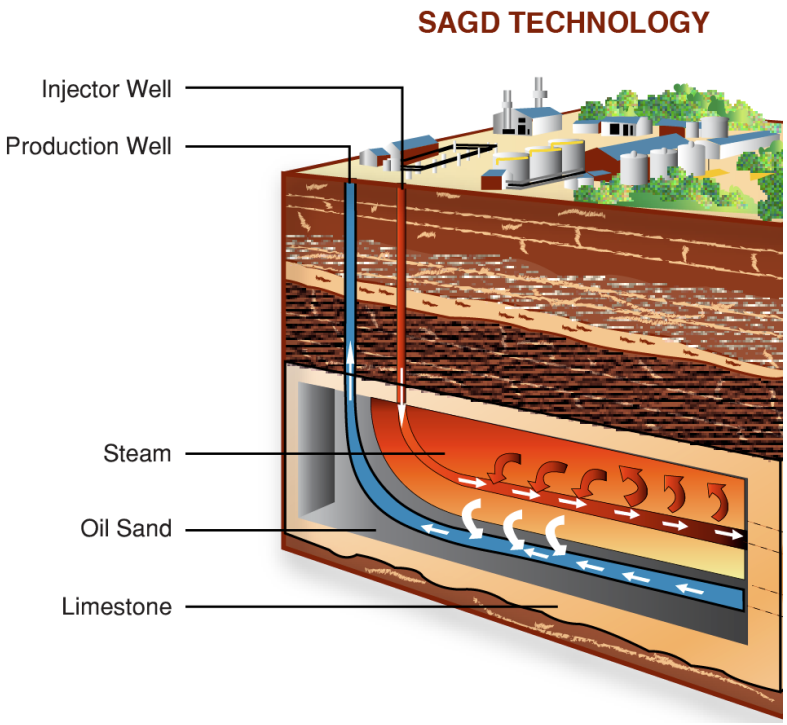


Figure 4.1: *SAGD Process Illustration*

Steam coming out of the injector would loss a lot of heat to the liquid pool before getting to the boundary of chamber. While in reality, interface level is difficult to measure. Thermal measuring methods are shown to be effective in reflecting the interface position [41]. Sub-cool is introduced as the temperature difference between tubing steam saturated temperature and tubing temperature at certain parts of the producer. Sub-cool can reflect the distance between interface and producer. When sub-cool is small and close to zero, the interface is close to the production well. While sub-cool is high, the interface is far away from the production well. Sub-cool values can be different with different geological parameters, measured locations, operating conditions and various factors. In current literatures and applications, set-point of sub-cool is usually given by operator experience. Sub-cool is required to be operated in a safe region, which will not lead to steam breakthrough or energy waste. The range usually runs from 10 to 50F. Conventionally, sub-cool control is implemented using production rate (pump frequency) as control variable. Increasing production rate will

cause the interface to drop down, with smaller sub-cool. Decreasing production rate will cause the interface to rise, with larger sub-cool. Steam injection has an opposite effect, but much more slowly, which is not used in sub-cool control yet. Currently in industry, sub-cool is being controlled manually or by PID controller with pump frequency as the manipulate variable.

Another important factor is the chamber pressure, measured at certain parts of the injector and producer. Inadequate steam chamber pressure can lead to safety issues, like steam chamber collapse or fracture. If there exist thief zones (gas cap), with chamber pressure higher than gas-cap pressure, steam and some of the bitumen will possibly be pushed into the gas cap [42], [43]. Also small chamber pressure encourages fluid inflow from surrounding areas, for example, water from connected aquifers [44]. Apart from operation safety, economic benefits are also affected by the chamber pressures. An industry study in Hangingstone reservoir pointed out the reduction in injection pressure led to a suspension of vertical growth of the steam chamber [45],[46]. While at the same time, low pressure gravity drainage is seen as an economic way of operating the SAGD process, by fully utilizing gravity as the driving mechanism [47]. Other important performance evaluation factors include: water-cut, cumulative steam oil ratio (cSOR), recovery factor (RF), etc. Water-cut denotes the amount of water contained in the total liquid produced. The water contains both steam injected and original water underground. A low water-cut is preferred as to further reduce the energy used to extract the bitumen. $cSOR$ is defined as:

$$cSOR = \frac{\int_0^T q_{steam}(t)dt}{\int_0^T q_{bitumen}(t)dt} \tag{4.1}$$

$cSOR$ measures the average volume of steam required to produce one barrel of bitumen. The less $cSOR$ is, the more energy we save in obtaining bitumen. RF is the recoverable amount of bitumen, it indicates the maximum extent we can extract from the reservoir.

Modeling of the SAGD process started from Roger Butler's analytical model to

predict oil production rate:

$$q = 2L\sqrt{\frac{2kg\alpha\rho_0\phi\delta S_0 H}{m\mu_{os}}} \tag{4.2}$$

Butler model is based on fundamental theory of flow and heat transfer, with the assumption that steam chamber does not touch the pay zone top and geographical parameters are uniform. It is generally acknowledged that the production rate in 4.2 is overestimated. Butler used this model as a tool to evaluate the feasibility of SAGD technique in oil sands extraction. Later improvement is based on modifying the parameters of 4.2, which can give a better estimation. However, the developed analytical equations have not been used in control and optimization for their inaccuracies.

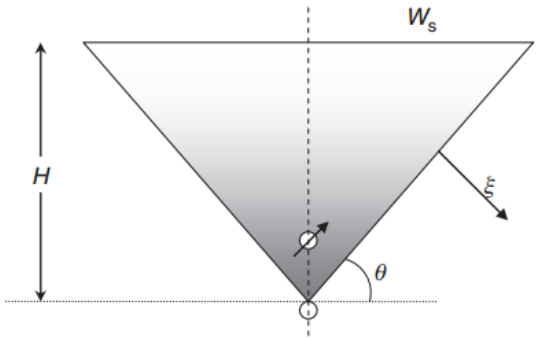


Figure 4.2: *Illustration of Butler Model*

Simulator based modeling is another important research area with fast development these years. Development in geostatistics and increased computation power help obtaining reliable simulation results [48]. Reservoir simulation brings into the benefit of generating field performance under various production schemes. It aids selection of an optimal set of producing conditions for the reservoir.

In evaluating SAGD economics, SOR comes first than oil rate per well pair. It has been stated that capital and operating cost of producer wells only account for $20 - 30\%$ of the total supply cost, while steam supply accounts for more than half. The reduction of steam used can result in less bitumen produced, but the reduction in SOR can greatly compensate the reduction in bitumen production [47]. The process

mechanism remains gravity driven fundamentally. Additional steam capacity to increase the pressure gradient is seen as an unattractive investment [47]. Topics about SAGD optimization in literature are mostly based on heuristic methods and treat the simulator as a black box. Typical methods include DECE (a heuristic optimizer developed by CMG) [49], genetic algorithm [50], trail and error method [47], [51]. The previous methods are computationally expensive due to time-consuming reservoir numerical simulations. Large number of decision variables is also another problem. Experimental design methods help reduce number of choices in variable selection [52], [49]. Surrogate model or proxy model is proposed to deal with the expensive computation effort [52], [53], [54], [55], [56], [57]. However, these surrogate models require complicated model structures,which bring great difficulty in implementing in real-time control and optimization. Simpler and effective models are desired.

Our intention here is to get a simple and reliable model of SAGD process for CCMPC.

4.2 Numerical Simulator

From the modeling point of view, numerical simulator serves as an important tool for modeling as it can provide process performance information in various circumstances, either in the situation of different geological parameters or in the situation of different operating conditions. Another reason is that doing identification experiment on SAGD process can be extremely expensive. Experiment test not only reduces economy benefit but also increases the operation risk.

Our research is based on the reservoir numerical simulator: Petroleum Experts. Petroleum Experts is a state of art software in petroleum engineering. Its distinguished feature REVEAL includes the thermal recovery process of heavy oil. We designed a simple case of one well pair to study its production behavior.

Simulation parameters are given in Table 4.1.

In the simulation work, grid dimension is reduced to 2, with only one grid in the horizontal dimension. The reason is that the simulation time can be significantly large

Table 4.1: *SAGD Process Parameters*

Types of Parameters	Name	Value
Geology Parameters	Grid Number	$1 \times 73 \times 55$
	Total Length	$1329ft \times 1120.7ft \times 165ft$
	Top Depth	340 ft
	Porosity	0.3
	Horizontal Permeability	$2500md$
	Vertical Permeability	$1750md$
Operation Data	Preheat Period	83 days
	Preheat Injection Rate	495.27 STB/day
	Preheat Injection Heat	1127.86 BTU/lb
	Steam Quality	0.9
Physical Properties	S_{wc}	0.3
	S_{owc}	0.2
	S_{ogc}	0.2
	S_{gc}	0.05
Initial Condition	Original Oil in Place	8.30045×10^9 STB
	Initial Pressure	217 psia
	Initial Temperature	60 F

in 3 dimensions, taking around 3 to 5 hours for a simulation of ten years. The 2 dimensional simulation can greatly reduce the simulation time to several minutes. With uniform property along the well dimension, 2D simulation does not show apparent difference from 3D simulation.

It shows the expansion of the steam chamber in Fig 4.3. The chamber grows rapidly in vertical direction at the beginning. When the chamber touches the overburden, it grows horizontally.

As we have introduced before, it is important to have the sub-cool and chamber pressure at certain levels, or the production efficiency can be greatly hindered, even

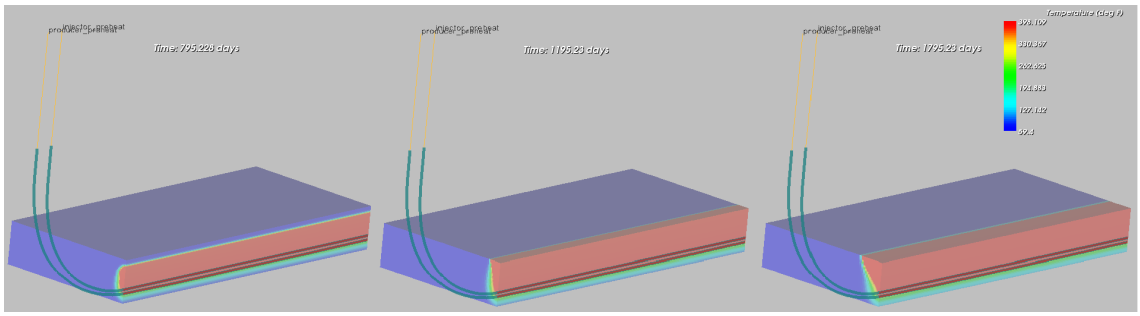
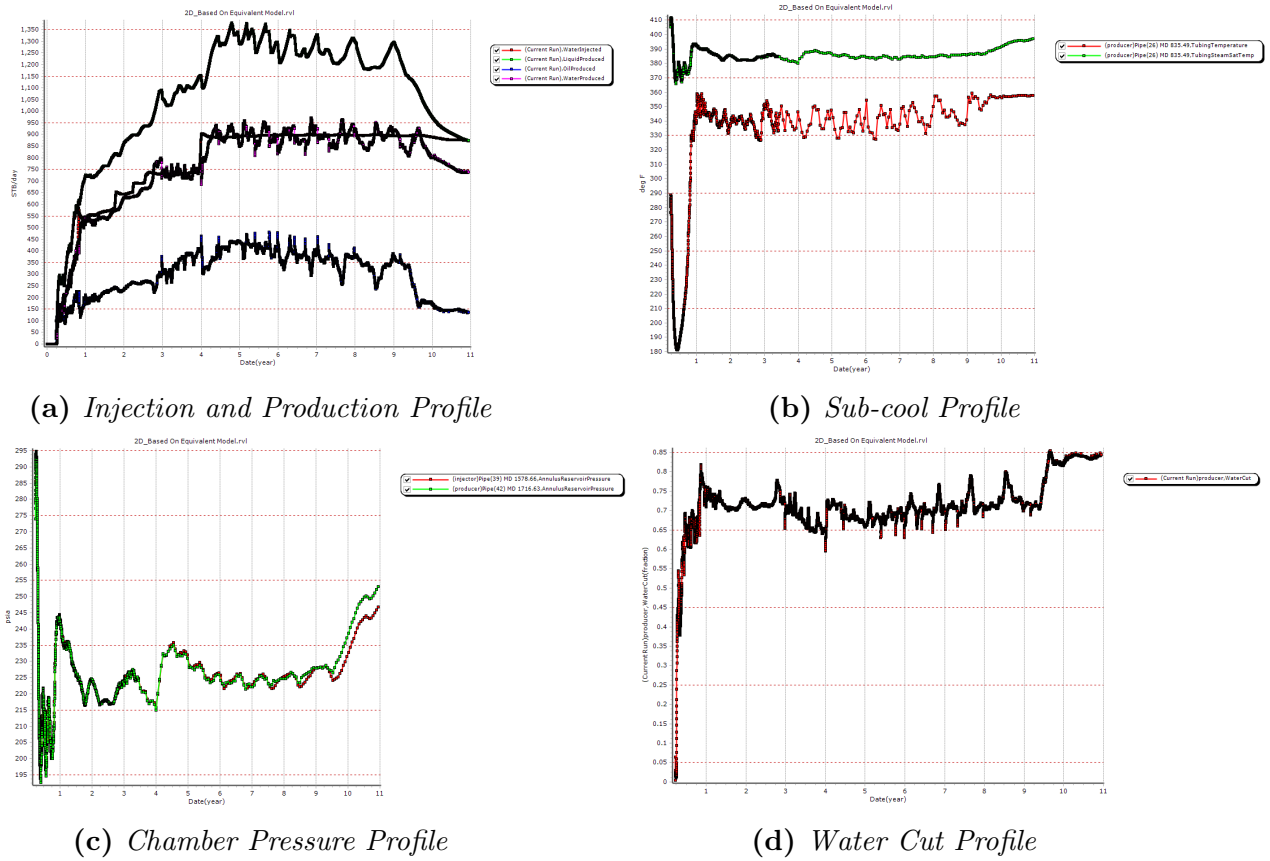


Figure 4.3: *Chamber Expansion in Simulation*

dangerous operating situations can happen. Appropriate operating conditions need to be applied for such requirements. Feedback control is implemented here as a simple method to maintain the appropriate working conditions. The feedback utilizes two PID controllers separately for sub-cool control and pressure control. Production profiles are shown in Fig 4.4a, 4.4b, 4.4c, 4.4d. PID controllers intend to operate sub-cool at 40F and chamber pressure at 217psia. The controller parameters are fixed through the whole operating period, but time-varying property of the system still leads to large variations of the controlled variables.

The operation profile varies a lot in the sub-cool, chamber pressure and water-cut during the first year. The first year includes the start-up period and a rapid chamber expansion time. Communication between the two wells is established and chamber expands to touch larger areas of bitumen. Peak production periods happen between the 4th year and 8th year. In the period between day 1500 and 3500, sub-cool varies from $27.4F$ to $58.9F$ and chamber pressure varies from $221.5Psia$ to $235.9Psia$. It is desirable to maintain these variables around specific levels and reduce their fluctuations for better economy benefit. As we have mentioned before, too large or too small the sub-cool or chamber pressure is will lead to various safety issues or poor economic behaviors. The purpose in this chapter is to obtain models for CCMPC and implement CCMPC on SAGD. It is desirable to obtain reliable predictions, maintain the system feasibility while the system is pushed towards better economic performance.



(a) *Injection and Production Profile*

(b) *Sub-cool Profile*

(c) *Chamber Pressure Profile*

(d) *Water Cut Profile*

Figure 4.4: *Detailed Comparison of Different Methods*

4.3 Economic Evaluation of SAGD Process

The peak period in SAGD process yields the largest amount of bitumen with a chamber already maturely developed. So we focus on the time horizon in between day 1500 and 3500. Before evaluating the economic profit of SAGD process, controller for the peak period is finely tuned to reduce steady state error and dynamic error. In sub-cool control, manipulated variable is production rate, controlled variable is sub-cool. In chamber pressure control, manipulated variable is injection rate and controlled variable is chamber pressure. For sub-cool control, the tuning parameters are $K_p = 50, K_I = 40$, maximum variation for liquid produced and sub-cool are 16.19% and 16.64% respectively. For chamber pressure control, the tuning parameters are $K_p = 50, K_I = 60, K_D = 2$, maximum variation for liquid injected and chamber

Table 4.2: *Comparison in Different Sub-cool Set-point*

Sub-cool Set-point $25F$	Sub-cool Set-point $40F$
Cumulative Bitumen Production $1.60 \times 10^5 STB$	Cumulative Bitumen Production $1.58 \times 10^5 STB$
cSOR 2.149	cSOR 2.177

Table 4.3: *Comparison in Different Chamber Pressure Set-point*

Chamber Pressure Set-point $211Psia$	Chamber Pressure Set-point $217Psia$
Cumulative Bitumen Production $1.5836 \times 10^5 STB$	Cumulative Bitumen Production $1.5828 \times 10^5 STB$
cSOR 2.1684	cSOR 2.2287

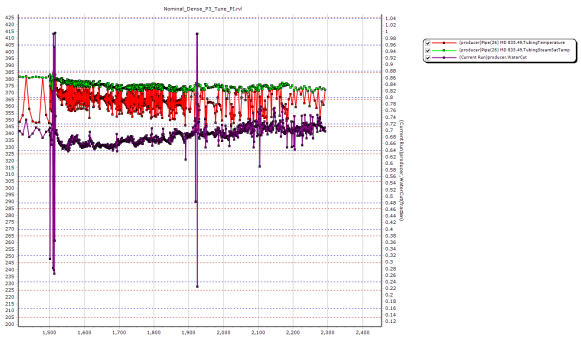
pressure are 18.88% and 2.02%.

With a finely tuned system, we examine economic benefits with different set-points. It has shown that for a certain range and time period, lower sub-cool and lower chamber pressure yield better economic performance. A typical comparison is given in Table 4.2 and Table 4.3. With lower sub-cool and lower chamber pressure, more bitumen is produced with less steam used.

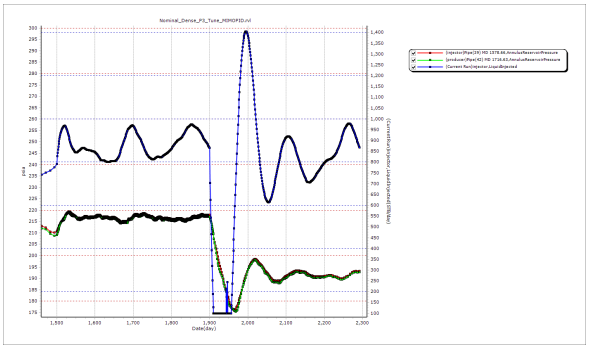
When the system has high sub-cool or high chamber pressure, we want to drive the system to low set-points with satisfaction of constraints. As a case study, we focus on transferring the system from $40F$ to $25F$ in sub-cool, and from $217Psia$ to $211Psia$ in chamber pressure, to yield more economic profit. We try to track the new set-points fast and satisfy probabilistic constraints.

However, merely using PID controllers can have problems in tracking new states. First one is during transition, PID controllers can drive sub-cool into a lower zone. Too low the sub-cool (e.g. 10F) is will lead to steam breakthrough Fig 4.5a. Another problem is the change of injection rate. Time constant for chamber pressure is rather large (up to 100 days). To reduce the response time, the injector is forced to shut down (simulator has a low limit of 100 STB/d). In Fig 4.5b, the set-point changes from 217 Psia to 191.7 Psia, and injector is forced to drop down to the lowest injection rate. However, for most of the cases, it cannot be tolerated to vary the injection rate a lot. Due to these problems, CCMPC is brought in. But there still exists the question

whether we shall put CCMPC in replace of PID controllers or put CCMPC at upper layer to give set-points to PID controllers. To answer this question, we will analyze the SAGD system with simulation tests.



(a) *Steam Breakthrough with Sub-cool Set-point 10F*



(b) *Injector Forced to be Shut Down with Chamber Pressure Drop*

Figure 4.5: *Violations of Operating Conditions*

4.4 Nonlinearity and Time-Varying Property of SAGD Process

In this section, we will analyze briefly about nonlinearity and time-varying property of SAGD process. From control point of view, the independent inputs to the system are injection rate and production rate. Note that the injection rate refers to the flow rate of steam injected, and the production rate refers to the flow rate of the produced liquid, which includes both bitumen, water and small amounts of gas. It is mostly argued that producer pump frequency is directly manipulated in industry. It can be shown that in simulation and in industry data that, flow rate and frequency are highly linear correlated. We select the flow rate for its easy use in the simulator. An equivalent pump frequency can be easily obtained from the flow rate. The outputs are selected as sub-cool and chamber pressure. It can be noticed from the previous sections that sub-cool and chamber pressure are vital to the operation safety and economy benefits. No measurement noises are added to the system in simulation.

Our focus is on the peak period from day 1500 to 2300 in Fig 4.6. Although

bitumen production reaches the maximum, steam injected and liquid production keeps slightly increasing to keep the chamber expanding. A constant injection rate and production rate is given to the process. Fig 4.7a shows the response of sub-cool. For some periods, a steady state of sub-cool can be reached in 20 to 30 days. However, sub-cool keeps increasing due to the unbalance between production rate and bitumen drainage rate. Even if we try to increase the value of production rate, we can not find a steady operating point for a time longer than 30 days because of the time-varying property (chamber expanding, etc). The response of chamber pressure is shown in Fig 4.7b. A slight drop is shown in the response of chamber pressure. In 20 days, the chamber pressure can still drop 4 psia, which can be negligible in a short time horizon. But in years, it can still become a problem. We try to use the trend to modify the injection and production profile to obtain a stable system response. Sigmoid function $f(x) = \dfrac{1}{1+e^{-x}}$ is used here to fit the trend. The response of sub-cool and chamber pressure is shown in Fig 4.8a, 4.8b. The improvement is that the sub-cool and chamber pressure varies around a set-point. But sub-cool shows a behavior of large fluctuation. We reason the fluctuation as a result of expanding chamber and unbalance between the injection rate and production rate. Chamber pressure shows a smaller moving trend with fluctuation (around 1 to 2 psia in 20 days) than the constant input. From this, we can tell that both sub-cool and chamber pressure show a nonlinear and time-varying property, which makes it hard to find a steady operating point.

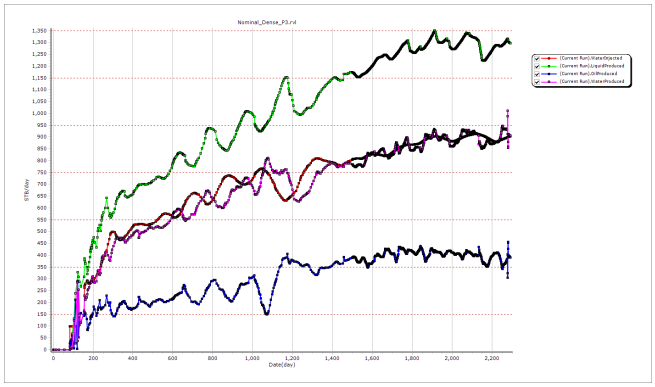
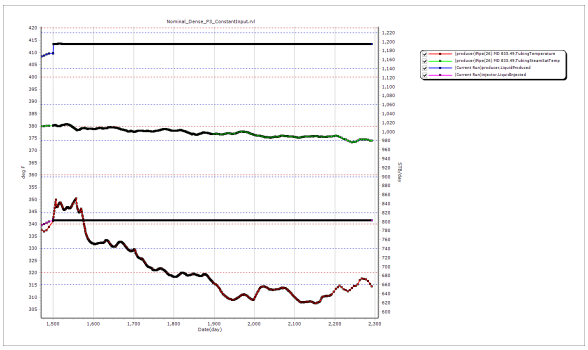
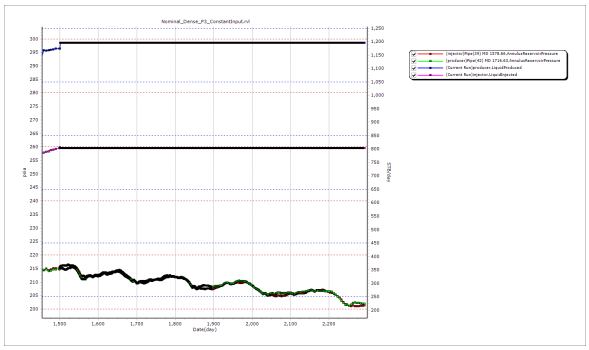


Figure 4.6: *Injection Rate and Production Rate Profile*

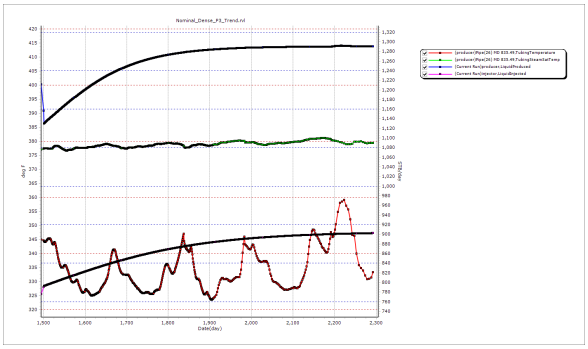


(a) *Sub-cool*

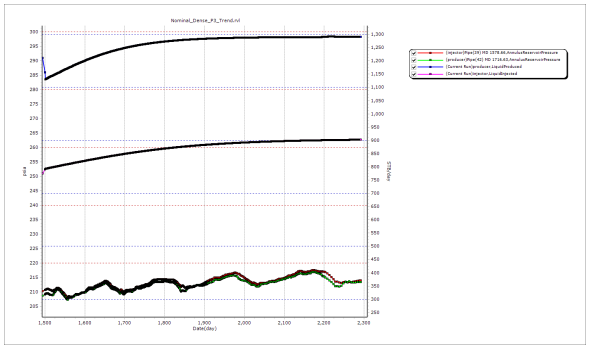


(b) *Chamber Pressure*

Figure 4.7: *Response with Constant Input*



(a) *Sub-cool*



(b) *Chamber Pressure*

Figure 4.8: *Response with Sigmoid Function Input*

By taking a closer look at detailed sub-cool behavior, we also notice that the sub-cool would have different behaviors in different operating regions. The sub-cool tends to be flat at regions when the sub-cool is relatively large, and tends to fluctuate when the sub-cool is in the middle of range, tends to varies a lot when sub-cool is small or close to zero. It can be explained in simple illustrations. In Fig 4.9, we show the cross-section area of the steam chamber and liquid production zone. Liquid steam interface is at the level where we want to control. To show the dynamic behavior of high and low sub-cool, we show the illustration in Fig 4.10. For low sub-cool, we have a low steam-liquid interface in the left figure. For high sub-cool, we have a high steam-liquid interface in the right figure. Knowing that sub-cool is mostly affected by the interface level, we assume that same change in interface level will lead to same

change in sub-cool. With the change of same liquid amount in the liquid zone, the distance change, however, will be larger in the left figure than in the right figure. From our assumption, we will have larger change in sub-cool for the circumstance with low sub-cool.

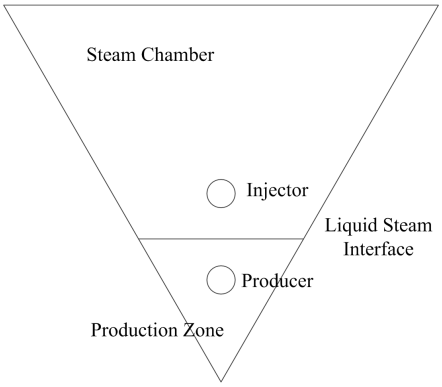


Figure 4.9: *Illustration of Steam Interface and Sub-cool*

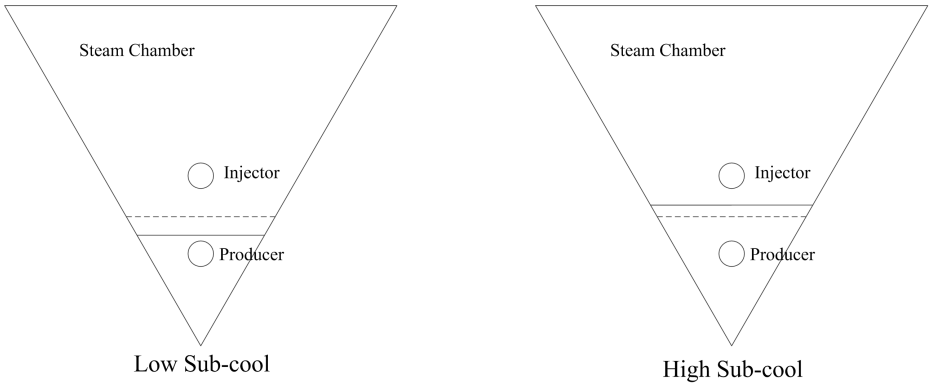


Figure 4.10: *Illustration of High Sub-cool and Low Sub-cool*

The behavior of sub-cool and chamber pressure exhibits a strong nonlinearity and time-varying property. However, we can still observe consistent properties of the SAGD process.

1. Different behaviors for sub-cool in different operating regions;

2. Trend of the production rate and injection rate in different periods.

4.5 System Identification of Closed-loop System

Due to the nonlinearity and time-varying property of the SAGD process, it is hard to do open-loop identification of SAGD process. However the system with a fine tuned controller possesses linear time invariant property and is easy to identify, CCMPC is going to be applied at upper layer, giving set-points to PID controllers. It tries to bring the system from low yield to high yield while satisfying probabilistic constraints.

So our modeling focuses on the whole system with both PID controllers and SAGD process. The reasons of having closed-loop can be summarized as the following:

1. Unstable systems must be operated in closed-loop;
2. Operational constraints may require closed-loop control;
3. Closed-loop controller maintains the system close to the operating point of interest;

The system we want to build is shown in Fig 4.11. A general system identification includes:

1. Selection of measured variables and controlled variables, which include system inputs, measured disturbances, outputs;
2. Preliminary process tests: nominal operation test, step test, staircase test, white noise test. The purpose is to get a general structure of the system model, also estimate parameter of system gain, time constant, frequency band width and time delay;
3. Experiment with persistent excitation;
4. Model structure selection;
5. Obtain model parameters based on certain criteria;
6. Model validation: correlation and autocorrelation analysis, fitness test.

Table 4.4: *Step Test Result*

Reference / Response	Reference Change in Sub-cool	Reference Change in Chamber Pressure
Response in Sub-cool	$K = 1, t_d = 0, \tau = 25$	Small Effect
Response in Chamber Pressure	Small Effect	$K = 1, t_d = 0, \tau = 70$

For the variables selection, system outputs are sub-cool and chamber pressure, system inputs are sub-cool external reference and chamber pressure external reference to PID controllers. As for the simulation, we does not add measurement noise. The system uncertainties are assumed to be uncertainties coming from inaccurate modeling and SAGD process. Preliminary step test has given for an estimate of the parameters shown in Table 4.4.

To identify the closed-loop system, persistently exciting reference signals are given to PID controllers. One persistently exciting signal is pseudorandom binary signal (PRBS). It is a signal that shifts between two levels in a certain fashion [58]. PRBS requires the user to select the levels and bandwidth. In previous section of economic evaluation, we have focused on tracking the new set-point. So the levels we have are, $40F, 25F$ for sub-cool and $217Psia, 211Psia$ for chamber pressure respectively. For bandwidth, we need to roughly guess the time constant of the process. A rough guess of the sub-cool and chamber pressure are 25 days and 70 days as shown in Table 4.4. In Matlab, the bandwidth is expressed as fractions of the Nyquist frequency:

$$0 \leq \omega \leq \frac{kT_s}{\pi\tau} \tag{4.3}$$

where, k usually takes value between 2 and 3, T_s is sampling time, which is 1 day here, τ is time constant estimate.

PRBS is given to the simulator for generating training and validation data. State space model, with its easy implication in control and optimization, is selected as model structure.

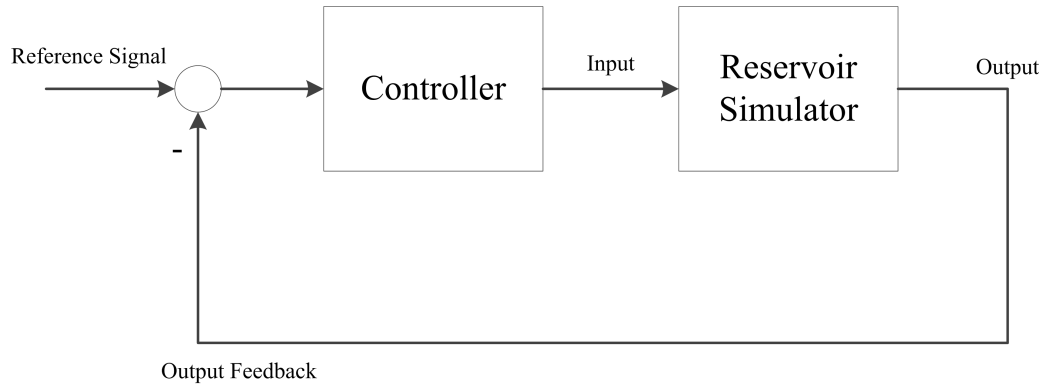


Figure 4.11: *Closed-loop System Diagram*

4.5.1 Subspace Method

A brief introduction is given [59] of using subspace methods for estimating state space models. Only the procedures are given here:

A general state space model is given as:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + \omega(t) \\ y(t) &= Cx(t) + Du(t) + v(t) \end{aligned} \tag{4.4}$$

where, $y(t)$ is a p-dimensional column vector, $u(t)$ is a m-dimensional column vector, and $x(t)$ is n-dimension. $\omega(t)$ and $v(t)$ are Gaussian white noise.

The first step is to estimate the observability matrix O_r. To estimate the observability matrix, we write the system output in augmented vectors.

$$Y_r(t) = O_r x(t) + S_r U_r(t) + V(t) \tag{4.5}$$

where,

$$\begin{aligned} Y_r(t) &= \begin{bmatrix} y(t) \\ y(t+1) \\ \vdots \\ y(t+r-1) \end{bmatrix}, \; U_r(t) = \begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+r-1) \end{bmatrix} \\ V(t) &= \begin{bmatrix} v(t) \\ C\omega(t) + v(t+1) \\ CA\omega(t) + C\omega(t+1) + v(t+2) \\ \vdots \\ CA^{r-1}\omega(t) + CA^{r-2}\omega(t+1) + \cdots + v(t+r-1) \end{bmatrix} \end{aligned} \tag{4.6}$$

By defining,

$$\begin{aligned} Y &= \begin{bmatrix} Y_r(1) & Y_r(2) & \cdots & Y_r(N) \end{bmatrix} \\ X &= \begin{bmatrix} x(1) & x(2) & \cdots & x(N) \end{bmatrix} \\ U &= \begin{bmatrix} U_r(1) & U_r(2) & \cdots & U_r(N) \end{bmatrix} \\ V &= \begin{bmatrix} V(1) & V(2) & \cdots & V(N) \end{bmatrix} \end{aligned} \tag{4.7}$$

we have,

$$Y = O_r X + S_r U + V \tag{4.8}$$

The observability matrix O_r is estimated by getting rid of U and V, we will show the result as:

$$G = O_r = \frac{1}{N} Y \Pi_{U^T}^{\perp} \Phi^T \tag{4.9}$$

where,

$$\Pi_{U^T}^{\perp} = I - U^T (UU^T)^{-1} U$$

$$\Phi = \begin{bmatrix} \varphi_s(1) & \varphi_s(2) & \cdots & \varphi_s(N) \end{bmatrix}, \ \varphi_s(t) = \begin{bmatrix} y(t-1) \\ \vdots \\ y(t-s_1) \\ u(t-1) \\ \vdots \\ u(t-s_2) \end{bmatrix}$$

Notice that here, $\varphi_s(t)$ is chosen to be uncorrelated with $V(t)$. As long as the requirement is satisfied, we can get any arbitrary combination of $\varphi_s(t)$, which will end in different realizations of state space model. Here $\varphi_s(t)$ is a special choice satisfying the condition.

Then we can obtain the matrix C and A:

$$\begin{aligned} C &= O_r(1:p, 1:n) \\ O_r(p+1:pr, 1:n) &= O_r(1:p(r-1), 1:n) A \end{aligned} \tag{4.10}$$

Notice that $1:n$ is defined in the format as in MATLAB, denoting row number and column number. The latter equation is solved in a least squares sense.

After obtaining C and A, we continue to estimate B, D and x_0 from linear regression problem.

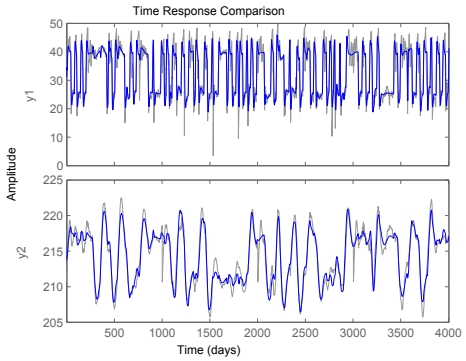
$$\arg \min_{B,D,x_0} \frac{1}{N} \sum_{t=1}^{N} \left|\left| y(t) - C(qI - A)^{-1} B u(t) - D u(t) - C(qI - A)^{-1} x_0 \delta(t) \right|\right|^2 \tag{4.11}$$

Lastly, the noise is estimated:

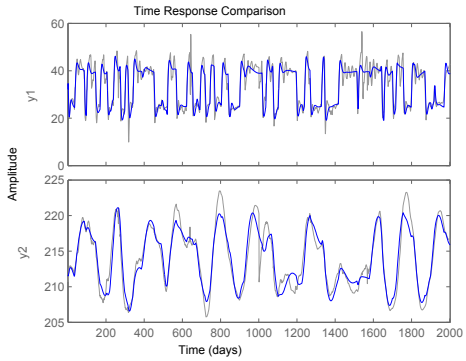
$$\begin{aligned} \hat{X} = L\hat{Y},\ L = R^{-1}U_1^T \\ \omega(t) = \hat{x}(t+1) - \hat{A}\hat{x}(t) - \hat{B}u(t) \\ v(t) = y(t) - \hat{C}\hat{x}(t) - \hat{D}u(t) \end{aligned} \tag{4.12}$$

Further details in given in the reference [59].

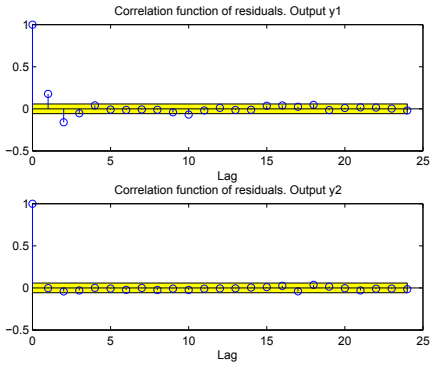
We have the infinite step prediction of training data in Fig 4.12a, infinite step prediction of validation data in Fig 4.12b, auto-correlation and cross-correlation of residuals shown in Fig 4.12c, 4.12d, 4.12e. We have the prediction fitness for sub-cool 60.78% in training and 61.69% in validation, for chamber pressure 72.27% in training and 68.69% in validation. The auto-correlation and cross-correlation test also pass.



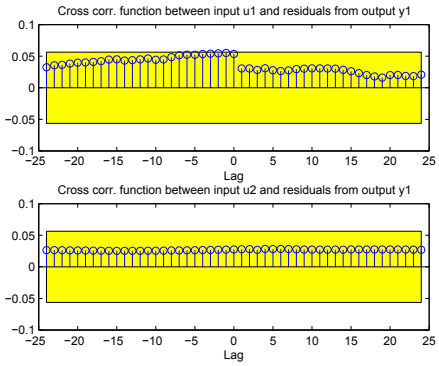
(a) *Training Prediction Result*



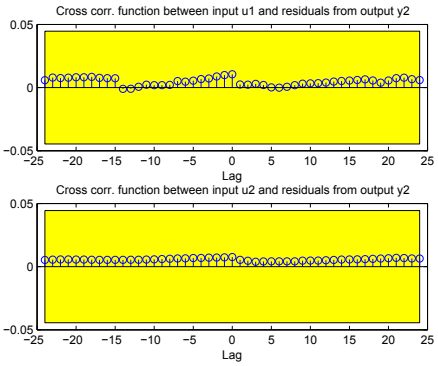
(b) *Validation Prediction Result*



(c) *Validation Autocorrelation Result*



(d) *Sub-cool Validation Error Cross-correlation*



(e) *Chamber Pressure Validation Error Cross-correlation*

Figure 4.12: *Training and Validation Result*

4.6 CCMPC on Linear Proxy Model

From the previous section, we get the linear proxy model:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + W\omega(t) \\ y(t) &= Cx(t) + \omega(t) \end{aligned} \tag{4.13}$$

where, y_1 is sub-cool, y_2 is chamber pressure, u_1 is the external reference of sub-cool, u_2 is the external reference of chamber pressure, $\omega(t) \sim N(0, \sigma^2)$.

The steady states are given by $x_s = (I - A)^{-1}Bu_s$. In the prediction horizon N, we have:

$$\mathbf{x} = Gx(k) + H\mathbf{u} + P\boldsymbol{\xi} \tag{4.14}$$

where,

$$G = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix},\ H = \begin{bmatrix} B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix}$$

$$P = \begin{bmatrix} W & 0 & \cdots & 0 \\ AW & W & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}W & A^{N-2}W & \cdots & W \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_{k+1|k} \\ x_{k+2|k} \\ \vdots \\ x_{k+N|k} \end{bmatrix},\ \mathbf{u} = \begin{bmatrix} u_k \\ u_k + 1 \\ \vdots \\ u_{k+N-1} \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \omega_k \\ \omega_{k+1} \\ \vdots \\ \omega_{k+N-1} \end{bmatrix}$$

Thus, we have:

$$\begin{aligned} x_{k+i|k} &= R_i\mathbf{x},\ \omega_{k+i|k} = r_i\boldsymbol{\xi} \\ y_{k+i|k} &= Cx_{k+i|k} + \omega_{k+i|k} \\ &= CR_iGx(k) + CR_iH\mathbf{u} + (CR_iP + r_i)\boldsymbol{\xi} \end{aligned} \tag{4.15}$$

where,

$$R_i = \begin{bmatrix} 0_{n\times n} & 0_{n\times n} & \cdots & I_n & \cdots & 0_{n\times n} & 0_{n\times n} \end{bmatrix}_{n\times nN}$$

$$r_i = \begin{bmatrix} 0_{q\times q} & 0_{q\times q} & \cdots & I_q & \cdots & 0_{q\times q} & 0_{q\times q} \end{bmatrix}_{q\times qN}$$

In R_i, I_n is $n \times n$ identity matrix in i_{th} block. In r_i, I_q is $q \times q$ identity matrix in i_{th} block.

The soft constraints are defined as:

$$\text{Prob}[y_{k+i|k} \geq \mathbf{s}] \geq 1 - \delta_i,\ i = 1, 2, \cdots, N \tag{4.16}$$

where,

$$\mathbf{s} = \begin{bmatrix} y_1^{lb} \\ y_2^{lb} \end{bmatrix}$$

Equivalently, we can write,

$$\text{Prob}[y_{0i} + y_i \boldsymbol{\xi} \leq \mathbf{0}] \geq 1 - \delta_i, \quad i = 1, \cdots, N \tag{4.17}$$

where,

$$\begin{aligned} y_{0i} &= \mathbf{s} - CR_iGx(k) - CR_iH\mathbf{u} \\ y_i &= -(CR_iP + r_i) \end{aligned} \tag{4.18}$$

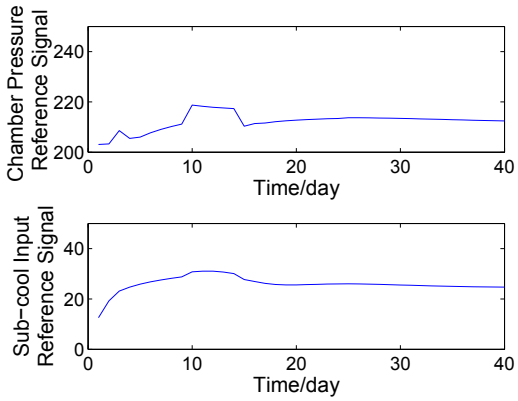
Thus the optimization at each step is solved as:

$$\begin{aligned} &\min_{\mathbf{u}} (E\mathbf{x} - \mathbf{x}_s)' Q_{\mathbf{x}} (E\mathbf{x} - \mathbf{x}_s) + (\mathbf{u} - \mathbf{u}_s)' Q_{\mathbf{u}} (\mathbf{u} - \mathbf{u}_s) \\ &\text{s.t.} E\mathbf{x} = Gx(k) + H\mathbf{u} \\ &\text{Prob}[y_{0i} + y_i \boldsymbol{\xi} \leq \mathbf{0}] \geq 1 - \delta_i, \quad i = 1, \cdots, N \end{aligned} \tag{4.19}$$

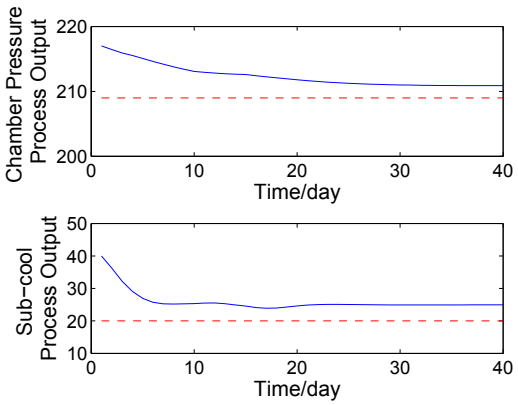
where, $\delta_i = 0.1,\ \forall i$

The result on simulation of the proxy model is shown in Fig 4.13a, 4.13b, 4.13c. Red dashed line shows lower bound. The result is shown with nominal prediction when the realization of noise equals its expectation value. For time response of sub-cool, rising time t_r is 4 days, settling time t_s equals 18 days, and the stabilizing error is within $0.09F$. For the response of chamber pressure, it is a much slower process, but still reaches within $1Psia$ error in 30 days. For consideration of constraints, both sub-cool and chamber pressure keep a safe distance from the boundary with a minimum distance of $3.9F$ and $1.9Psia$. Notice in Fig 4.13a, the reference given to the controller is small at the beginning, to get a faster respond. The set-point is adjusted backwards soon to avoid constraint violation and obtain stabilized output. The result is shown with a probability of satsifaction of 90%. However for arbitrary set-point given, which is shown in Fig 4.13d, the result is much worse: chamber pressure drops below the lower constraint and does not stabilize. For the nominal

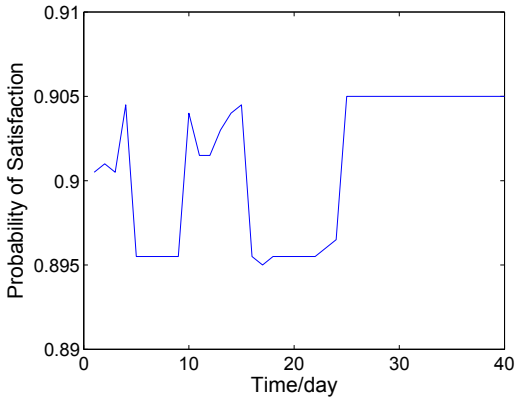
trajectory of sub-cool, it does not violate, the minimum distance from the boundary is $1.54F$. By merely arbitrary set-point, the system cannot be stabilized nor satisfy the constraints.



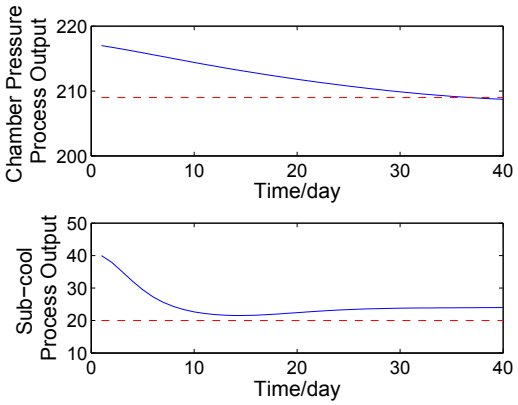
(a) *Chamber Pressure and Sub-cool Reference Signal Using CCMPC*



(b) *Process Output Chamber Pressure and Sub-cool Using CCMPC*



(c) *Probability of Satisfaction Using CCMPC*

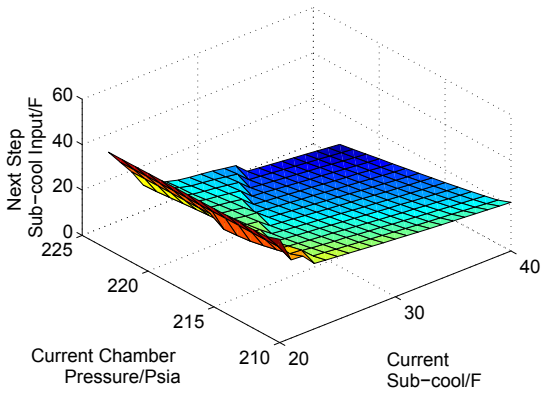


(d) *Chamber Pressure and Sub-cool Response under Original Arbitrary Set-point*

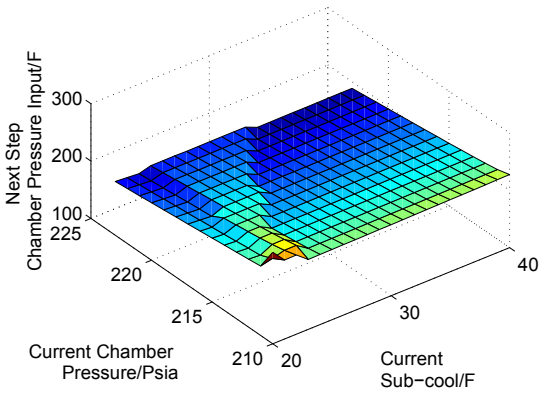
Figure 4.13: *Comparison Between Performance with and without CCMPC*

Record the CCMPC control inputs (reference signals to PID controllers) from selected measured outputs, and it is shown with 3D plots in Fig 4.14a, 4.14b. For detailed analysis of the control policy, the control input is plotted with each individual output in 2D plot in Fig 4.15a, 4.15b, 4.15c, 4.15d. For sub-cool control, the output of sub-cool is the major factor while chamber pressure is the minor factor. Larger

sub-cool will force the sub-cool set-point to be smaller and wise versa. For chamber pressure in most of the cases, larger chamber pressure will force the sub-cool control input to be smaller because the bitumen drainage rate is higher when chamber pressure is larger. Even though it is not monotonic all the time. However, for low sub-cool, the chamber pressure almost does not affect for the sub-cool control. For chamber pressure control, larger chamber pressure will force the chamber pressure control input to be smaller and wise versa. The sub-cool almost has no affect on selecting the chamber pressure control input except for some small ranges.



(a) *Sub-cool Control Input with Measured Outputs of Chamber Pressure and Sub-cool*

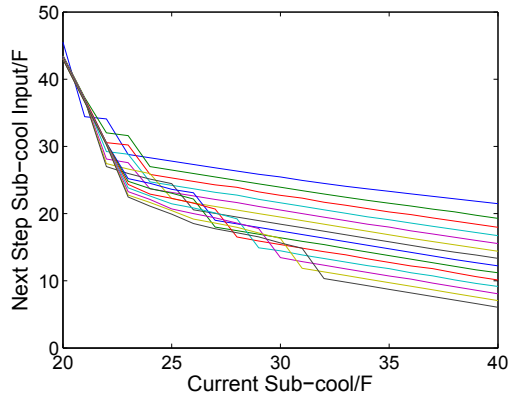


(b) *Chamber Pressure Control Input with Measured Outputs of Chamber Pressure and Sub-cool*

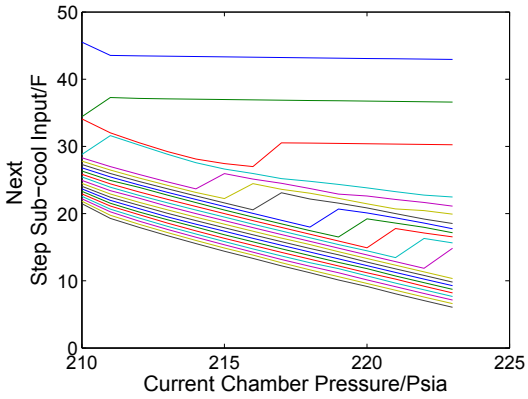
Figure 4.14: *Recorded Control in 3 Dimension*

The controller is also tested on the Petroleum Experts for the performance. A comparison between arbitrary set-point and CCMPC is given in Fig 4.16. The sub-cool control shows a great performance with fast response and keeps good track of required set point. It also keeps a safe distance from violating the constraints. For chamber pressure control, it also keeps away from the constraints and reduces the variations of chamber pressure greatly. In Table 4.5, we also show the comparison in violation of constraints. Arbitrary set-point method has a much larger violation with 37.60%, while CCMPC satisfies probabilistic constraints with 100%.

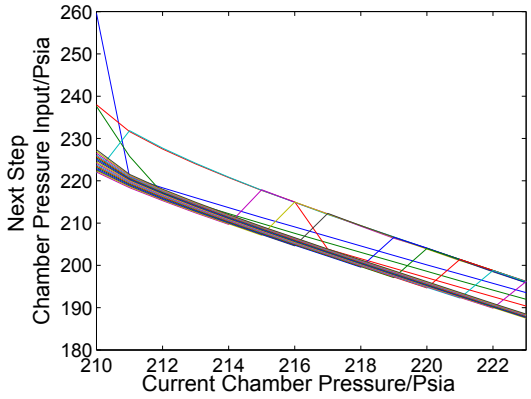
Due to the discrepancy between proxy model and simulator, CCMPC does not push the control performance limit to 90% of probability satisfaction. CCMPC



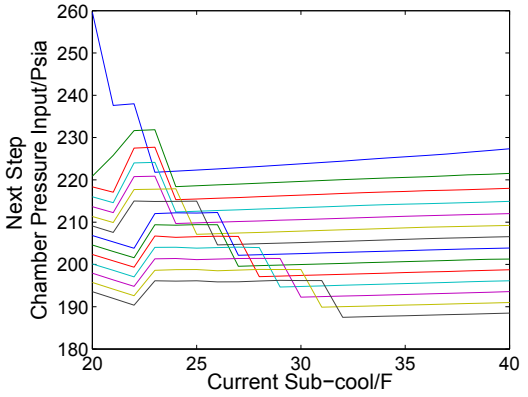
(a) *Sub-cool Input to PID controller with Measured Outputs of Sub-cool*



(b) *Sub-cool Input to PID controller with Measured Outputs of Chamber Pressure*



(c) *Chamber Pressure Input to PID controller with Measured Outputs of Chamber Pressure*



(d) *Chamber Pressure Input to PID controller with Measured Outputs of Sub-cool*

Figure 4.15: *Calculated Control Input with Measured Output*

achieves a conservative result with no violations of constraints. Certainly, safety issues due to the system uncertainties are greatly reduced with the use of CCMPC. If a better proxy model can be developed, CCMPC can achieve better performance in its application.

4.7 Conclusions

In this chapter, SAGD simulation is set up in REVEAL. Various tests are given and SAGD process shows a strong nonlinearity and time-varying property. To compensate

Table 4.5: *Comparison between CCMPC and without CCMPC*

Violation with CCMPC	Violation without CCMPC
0%	37.60%

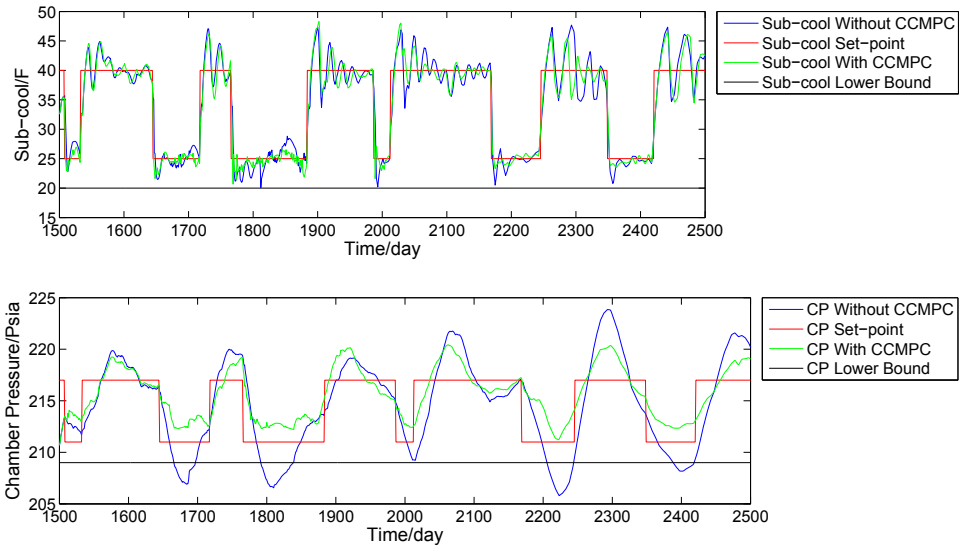


Figure 4.16: *Response Comparison between CCMPC and without CCMPC*

this deficiency, CCMPC is designed at upper level, giving reference signals to PID controllers of the SAGD process. State space model is identified using subspace method. CCMPC is designed on the proxy model (state space model). Off-line control input is pre-recorded and implemented in the simulator. The result shows that with the use of CCMPC, violation of constraints and variation of system output are all greatly reduced.

Chapter 5

Conclusions

5.1 Summary of this thesis

This thesis focuses on chance constrained MPC methods for linear system with additive and multiplicative uncertainty. Conventional methods are restricted by certain distributions. RO method is proposed to deal with arbitrary distribution. RO based CCMPC is further applied to SAGD process. Constraint violations due to uncertainties are greatly reduced.

Motivation, thesis contributions and thesis outline are included in Chapter 1.

Chapter 2 first formulates the chance constrained problem in MPC. Two typical analytical methods are given, including ellipsoidal relaxation and iterative risk allocation. However, both methods have to assume Gaussian distribution of the uncertainty. RO method is proposed based on the idea of robust optimization and uncertainty set. Probability of random variables within the set is calculated based on Monte Carlo method. Arbitrary distributions can be included in this framework. With one layer having guarantee of the probability, another layer is being used to optimize the cost. The optimization performance can be further improved with the inclusion of closed-loop framework.

Chapter 3 proposes how to deal with multiplicative (parametric) uncertainty with RO based CCMPC. Stochastic tubes are introduced to relax the original constraints into recursive ones. The recursive constraints preserve linear chance constraint forms, which can be solved by RO method. Apart from that, a recursive feasibility is also

guaranteed.

Chapter 4 applies RO based CCMPC to SAGD process. Different sets of operation data are generated from reservoir simulator first. Linear state space model is built with sub-space method. The linear state space model is treated as a proxy model for the real process. CCMPC is tested on the reservoir simulator and greatly reduces violations due to uncertainties.

5.2 Directions of future work

This thesis focuses on RO based CCMPC and its applications. The work includes optimization, control and identification of process with uncertainties. To improve the performance of RO method, following aspects can be considered in the future:

1. In the formulation of robust optimization approximation, t is given in $\frac{1}{t}[(\eta + t)]^+$ as upper bound for the indicator function. t controls the difference between upper bound and indicator function. Investigation of how the value t will affect the objective function is needed for further research.

2. Stochastic tubes constructed for LPV system have requirements for both stability and feasibility. It should be pointed out that not all initial states in the feasible region can be included in the stochastic tubes. The use of stochastic tubes requires the initial states to be in a smaller area than the original feasible region. Efforts to expand stochastic tubes to include larger areas can be meaningful to the current work.

3. Gaussian process (GP) dynamical model is one non-parametric method to capture the nonlinearities of data. GP is a stochastic process governing the properties of the functions. Without explicit form of the function, GP can give inference of the prediction distribution. Further investigation into GP dynamic model might help increase model accuracy of SAGD process.

4. In application of SAGD process, CCMPC is implemented to the reservoir simulator based on a fixed proxy model. As we know, SAGD process itself possesses

time-varying properties. Adaptive models built from real-time operation data need to be implemented. Better performance can be achieved with adaptive models.

Bibliography

[1] Alberto Bemporad and Manfred Morari. Robust model predictive control: A survey. In *Robustness in identification and control*, pages 207–226. Springer, 1999.

[2] Zhi Q Zheng and Manfred Morari. Robust stability of constrained model predictive control. In *American Control Conference, 1993*, pages 379–383. IEEE, 1993.

[3] Ali Mesbah. Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems Magazine, Accepted*, 2016.

[4] Daniele Bernardini and Alberto Bemporad. Scenario-based model predictive control of stochastic constrained linear systems. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 6333–6338. IEEE, 2009.

[5] Masahiro Ono and Brian C Williams. Iterative risk allocation: A new approach to robust model predictive control with a joint chance constraint. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 3427–3432. IEEE, 2008.

[6] Michael P. Vitus. *Stochastic control via chance constrained optimization and its application to unmanned aerial vehicles*. PhD thesis, Standford university, 2012.

[7] Yuan Yuan, Zukui Li, and Biao Huang. Robust optimization approximation for joint chance constrained optimization problem. *Journal of Global Optimization*, pages 1–23, 2015.

[8] Nikolas Kantas, JM Maciejowski, and A Lecchini-Visintini. Sequential monte carlo for model predictive control. In *Nonlinear Model Predictive Control*, pages 263–273. Springer, 2009.

[9] James Luedtke and Shabbir Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699, 2008.

[10] Mark Cannon, Basil Kouvaritakis, Saša V Rakovic, and Qifeng Cheng. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Transactions on Automatic Control*, 56(1):194–200, 2011.

[11] Basil Kouvaritakis, Mark Cannon, Saša V Raković, and Qifeng Cheng. Explicit use of probabilistic distributions in linear predictive control. *Automatica*, 46(10):1719–1724, 2010.

[12] Frauke Oldewurtel, Colin N Jones, and Manfred Morari. A tractable approximation of chance constrained stochastic mpc based on affine disturbance feedback. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 4731–4736. IEEE, 2008.

[13] Frauke Oldewurtel, Alessandra Parisio, Colin Jones, Manfred Morari, Dimitrios Gyalistras, Markus Gwerder, Vanessa Stauch, Beat Lehmann, and Katharina Wirth. Energy efficient building climate control using stochastic model predictive control and weather predictions. In *Proceedings of the 2010 American control conference*, number EPFL-CONF-169733, pages 5100–5105. Ieee Service Center, 445 Hoes Lane, Po Box 1331, Piscataway, Nj 08855-1331 Usa, 2010.

[14] CM Lagoa. On the convexity of probabilistically constrained linear programs. In *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, volume 1, pages 516–521. IEEE, 1999.

[15] Farid Alizadeh and Donald Goldfarb. Second-order cone programming. *Mathematical programming*, 95(1):3–51, 2003.

[16] Mark Cannon, Basil Kouvaritakis, and Desmond Ng. Probabilistic tubes in linear stochastic model predictive control. *Systems & Control Letters*, 58(10):747–753, 2009.

[17] Pu Li, Moritz Wendt, and GüNter Wozny. A probabilistically constrained model predictive controller. *Automatica*, 38(7):1171–1176, 2002.

[18] Lars Blackmore, Masahiro Ono, Askar Bektassov, and Brian C Williams. A probabilistic particle-control approximation of chance-constrained stochastic predictive control. *IEEE Transactions on Robotics*, 26(3):502–517, 2010.

[19] Alexander T Schwarm and Michael Nikolaou. Chance-constrained model predictive control. *AIChE Journal*, 45(8):1743–1752, 1999.

[20] Giuseppe C Calafiore, Marco C Campi, et al. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

[21] Irwin Miller. Probability, random variables, and stochastic processes. *Technometrics*, 8(2):378–380, 1966.

[22] Zukui Li, Ran Ding, and Christodoulos A Floudas. A comparative theoretical and computational study on robust counterpart optimization: I. robust linear optimization and robust mixed integer linear optimization. *Industrial & engineering chemistry research*, 50(18):10567–10603, 2011.

[23] Zukui Li, Qiuhua Tang, and Christodoulos A Floudas. A comparative theoretical and computational study on robust counterpart optimization: Ii. probabilistic guarantees on constraint satisfaction. *Industrial & engineering chemistry research*, 51(19):6769–6788, 2012.

[24] Isaac Meilijson and Arthur Nádas. Convex majorization with an application to the length of critical paths. *Journal of Applied Probability*, pages 671–677, 1979.

[25] Aharon Ben-Tal, Alexander Goryashko, Elana Guslitzer, and Arkadi Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

[26] Luigi Chisci, J Anthony Rossiter, and Giovanni Zappa. Systems with persistent disturbances: predictive control with restricted constraints. *Automatica*, 37(7):1019–1028, 2001.

[27] D Limon Marruedo, T Alamo, and EF Camacho. Input-to-state stable mpc for constrained discrete-time nonlinear systems with bounded additive uncertainties. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 4, pages 4619–4624. IEEE, 2002.

[28] Arthur Richards and Jonathan How. Robust stable model predictive control with constraint tightening. In *2006 American Control Conference*, pages 6–pp. IEEE, 2006.

[29] Wilbur Langson, Ioannis Chryssochoos, SV Raković, and David Q Mayne. Robust model predictive control using tubes. *Automatica*, 40(1):125–133, 2004.

[30] David Q Mayne, María M Seron, and SV Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, 2005.

[31] Qifeng Cheng, Mark Cannon, Basil Kouvaritakis, and Martin Evans. Stochastic mpc for systems with both multiplicative and additive disturbances. *IFAC Proceedings Volumes*, 47(3):2291–2296, 2014.

[32] Martin Evans, Mark Cannon, and Basil Kouvaritakis. Robust and stochastic linear mpc for systems subject to multiplicative uncertainty. *IFAC Proceedings Volumes*, 45(17):335–341, 2012.

[33] Giuseppe C Calafiore and Lorenzo Fagiano. Stochastic model predictive control of lpv systems via scenario optimization. *Automatica*, 49(6):1861–1866, 2013.

[34] James Fleming, Mark Cannon, and Basil Kouvaritakis. Stochastic tube mpc for lpv systems with probabilistic set inclusion conditions. In *53rd IEEE Conference on Decision and Control*, pages 4783–4788. IEEE, 2014.

[35] Georges Bitsoris. On the positive invariance of polyhedral sets for discrete-time systems. *Systems & control letters*, 11(3):243–248, 1988.

[36] Roger Butler et al. Sagd comes of age! *Journal of Canadian Petroleum Technology*, 37(07), 1998.

[37] RM Butler, GS McNab, and HY Lo. Theoretical studies on the gravity drainage of heavy oil during in-situ steam heating. *The Canadian journal of chemical engineering*, 59(4):455–460, 1981.

[38] Al-Muatasim Al-Bahlani and Tayfun Babadagli. Sagd laboratory experimental and numerical simulation studies: A review of current status and future issues. *Journal of Petroleum Science and Engineering*, 68(3):135–150, 2009.

[39] Neil Edmunds et al. Investigation of sagd steam trap control in two and three dimensions. *Journal of Canadian Petroleum Technology*, 39(01), 2000.

[40] Ian D Gates and Christopher Leskiw. Impact of steam trap control on performance of steam-assisted gravity drainage. *Journal of Petroleum Science and Engineering*, 75(1):215–222, 2010.

[41] NR Edmunds, JA Kovalsky, SD Gittins, ED Pennacchioli, et al. Review of phase a steam-assisted gravity-drainage test. *SPE Reservoir Engineering*, 9(02):119–124, 1994.

[42] M Pooladi-Darvish, L Mattar, et al. Sagd operations in the presence of overlying gas cap and water layer-effect of shale layers. *Journal of Canadian Petroleum Technology*, 41(06), 2002.

[43] Roger Butler et al. Some recent developments in sagd. *Journal of Canadian Petroleum Technology*, 40(01), 2001.

[44] Chi-Tak Yee, A Stroich, et al. Flue gas injection into a mature sagd steam chamber at the dover project (formerly utf). *Journal of Canadian Petroleum Technology*, 43(01), 2004.

[45] Y Ito, T Hirata, M Ichikawa, et al. The effect of operating pressure on the growth of the steam chamber detected at the hangingstone sagd project. *Journal of Canadian Petroleum Technology*, 43(01), 2004.

[46] Y Ito, G Ipek, et al. Steam fingering phenomenon during sagd process. In *SPE International Thermal Operations and Heavy Oil Symposium*. Society of Petroleum Engineers, 2005.

[47] Neil Edmunds, Harbir Chhina, et al. Economic optimum operating pressure for sagd projects in alberta. *Journal of Canadian Petroleum Technology*, 40(12), 2001.

[48] Dean S Oliver and Yan Chen. Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1):185–221, 2011.

[49] Card Yang, C Card, L Nghiem, et al. Economic optimization and uncertainty assessment of commercial sagd operations. *Journal of Canadian Petroleum Technology*, 48(09):33–40, 2009.

[50] ID Gates, N Chakrabarty, et al. Optimization of steam assisted gravity drainage in mcmurray reservoir. *Journal of Canadian Petroleum Technology*, 45(09), 2006.

[51] W. Bailey J.Gossuin, P. Naccache and B. Couet. Sagd optimization under uncertainty. In *World Heavy Oil Congress*. World Heavy Oil Congress, 2011.

[52] Nestor V Queipo, Javier V Goicochea, and Salvador Pintos. Surrogate modeling-based optimization of sagd processes. *Journal of Petroleum Science and Engineering*, 35(1):83–93, 2002.

[53] Ehsan Amirian, Juliana Y Leung, Stefan Zanon, and Peter Dzurman. Integrated cluster analysis and artificial neural network modeling for steam-assisted gravity

drainage performance prediction in heterogeneous reservoirs. *Expert Systems with Applications*, 42(2):723–740, 2015.

[54] Eugene Fedutenko, Chaodong Yang, Colin Card, Long X Nghiem, et al. Time-dependent neural network based proxy modeling of sagd process. In *SPE Heavy Oil Conference-Canada*. Society of Petroleum Engineers, 2014.

[55] Jose Walter Prada Vanegas, Clayton Vernon Deutsch, Luciane Bonet Cunha, et al. Uncertainty assessment of sagd performance using a proxy model based on butler's theory. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2008.

[56] Eugene Fedutenko, Chaodong Yang, Colin Card, Long X Nghiem, et al. Forecasting sagd process under geological uncertainties using data-driven proxy model. In *SPE Heavy Oil Conference Canada*. Society of Petroleum Engineers, 2012.

[57] Mahyar Matt Mohajer, Carlos Emilio Perez Damas, Alexander Jose Berbin Silva, Andreas Al-Kinani, et al. An integrated framework for sagd real-time optimization. In *SPE Intelligent Energy Conference and Exhibition*. Society of Petroleum Engineers, 2010.

[58] HG Natke. System identification: Torsten söderström and petre stoica, 1992.

[59] Lennart Ljung. *System Identification: Theory for the User*. Prentice Hall, 1999.