Sensitivity Analysis and Robust Design of Pharmaceutical Manufacturing Processes

Von der Fakultät für Maschinenbau
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung der Würde

eines Doktor-Ingenieurs (Dr.-Ing.)

genehmigte Dissertation

von: M.Sc. Xiangzhong Xie
geboren in: Nanchang, Jiangxi, China

eingereicht am: 14.10.2019
mündliche Prüfung am: 10.02.2020

Vorsitzender:

Prof. Dr.-Ing. Stephan Scholl

Gutachter:

Prof. Dr.-Ing. Ulrike Krewer
Prof. Dr. rer.nat. habil. Sebastian Sager

2020

Statement of authorship

I, M.Sc. Xiangzhong Xie, hereby state that this PhD thesis, titled *Sensitivity Analysis and Robust Design of Pharmaceutical Manufacturing Processes*, was written entirely independently by myself and that it was not used in any other examination before. Furthermore, I declare to the best of my knowledge that no other work with any altered version exists, and that all references required for the completion of this work have been indicated and cited as such.

Xiangzhong Xie Braunschweig, February 2020

Acknowledgments

The work for this thesis was conducted during the period from 2015 to 2019 at the Institute of Energy and Process Systems Engineering at the TU Braunschweig. This work would not have the spirit that it has without the invaluable supports and belief in me as an academic researcher and scientific writer, provided by the following scholars:

Prof. Dr.-Ing. Ulrike Krewer:

I would like to thank you for providing me such an opportunity to work in the field of process system engineering and on the topic of pharmaceutical processes, guiding me throughout my PhD period, and encouraging me to have faith in myself. I also appreciate the discussions we had and the thoughtful comments you provided for my papers and thesis, which make them better.

Dr.-Ing. René Schenkendorf:

I would like to extend my sincere thanks to you, who is not just my mentor but also my friend during my PhD time. Within three years of working with you, I was always encouraged and inspired by the discussions we had. Without your experience and passion, I could not publish my work and finish my thesis as efficient as it is.

I also would like to appreciate the valuable discussion and comments, which provide by the following scholars:

Prof. Dr.-Ing. Stephan Scholl:

I appreciate you for being my second supervisor in the *μProps* project and offering me the feedbacks regarding my work.

Prof. Dr. rer.nat. habil. Sebastian Sager:

I also would like to thank you for being my second supervisor in the International Max Planck Research Schools (IMPRS) and having helpful discussions about my work.

Jun.-Prof. Dr.-Ing. Ulrich Römer:

I am also grateful to you for being my second examiner of my oral examination as well as the insightful discussion we had afterwards.

Other scholars:

I'd also like to extend my gratitude to the editors and anonymous reviewers for my papers. Thank you for taking time and effort to read my work and provide useful com-

ments. Thanks also to the scholars who had discussion with me during the conferences and help me to improve myself.

My last four years will also not be splendid and memorable without the collaborations and discussions with my colleagues. Thanks Victor Emenike and Moritz Schulze for being as a team. It is really nice to working with you. Thanks Nan Lin for the practical discussion and helpful contribution in our collaborated work. I'd also like to recognize the effort that I received from other colleagues. Thanks should also go to Ina Schunke, Wilfried Janßen and Uwe Herrmann for the administration work and technical support.

I also had great pleasure of working with the colleagues from PVZ and other PhD students involved in the $\mu Props$ project. It is really nice to be working with you. Thanks Dr.-Ing. Jörg Stieghan and Dr. Gerlinde Benninger for the guidance and support during the $\mu Props$ project.

Last but not least, I would like to express my deepest appreciation to my parents and family. Mom and Dad, my life will not be possible without the selfless love and nurture from you. I would like to thank my wife for her consistent support and my son for the happiness he brings to my life.

Contents

List of Figures

List of Tables

Abstract

The existence of parameter uncertainties limits model-based process design techniques. It also hinders the modernization of pharmaceutical manufacturing processes, which is necessitated for intensified market competition and Quality by Design (QbD) principles. Thus, in this thesis, proper approaches are proposed for efficient and effective sensitivity analysis and robust design of pharmaceutical processes. Moreover, the point estimate method (PEM) and polynomial chaos expansion (PCE) are further implemented for uncertainty propagation and quantification (UQ) in the proposed approaches.

Global sensitivity analysis (GSA) provides quantitative measures on the influence of parameter uncertainties on process outputs over the entire parameter domain. Different global sensitivity analysis techniques are presented in detail. The PCE is utilized for efficient computation of these sensitivity measures. The results from case studies show that GSA is able to quantify the heterogeneity of the information in parameter uncertainties and model structure. The significant impact of parameter dependencies on GSA ranking and the variation in simulation results are also revealed.

Comprehensive frameworks for robust process design are introduced to alleviate the adverse effect of parameter uncertainties on process performance. The first robust design framework is developed based on the generalized PEM. The proposed approach is able to increase the process robustness at a low computational expense. Here, the negative influence of process constraints as well as the necessity of considering parameter dependencies in robust design are also addressed. Then, the framework is extended to parameter uncertainties with non-Gaussian distributions. A novel approach, in which the Gaussian mixture distribution (GMD) concept is combined with PEM, is proposed to handle this issue. The resulting GMD-PEM concept provides a better trade-off between process efficiency and an acceptable limit of constraint violations than other approaches. The second robust design framework proposed in this work is based on the iterative back-off strategy and PCE. It provides designs with the desired robustness, while the associated computational expense depends no longer on the formulation of the optimization problem. The decoupling of optimization and uncertainty quantification provides the possibility of implementing robust process design to more complex pharmaceutical manufacturing processes with large number of uncertain parameters.

In this thesis, the case studies include unit operations for (bio)chemical synthesis, separation (crystallization) and formulation (freeze-drying), which cover the complete production chain of pharmaceutical manufacturing. Results from the case studies reveal the significant impact of parameter uncertainties on process design. Also they show the efficiency and effectiveness of the proposed frameworks regarding process performance and robustness in the context of QbD.

Kurzfassung

Die pharmazeutische Industrie muss sowohl den gestiegenen Wettbewerbsdruck standhalten als auch die von Regulierungsbehörden geforderte QbD-Initiative (Quality by Design) umsetzen. Neben Erfahrungswerten und Expertenwissen können insbesondere modellgestützte Verfahren einen signifikanten Beitrag leisten, um auch zukünftig kostengünstige Arzneimittelwirkstoffe unter QbD herzustellen. Parameterunsicherheiten (PU) erschweren jedoch eine zuverlässige modellgestützte Prozessauslegung. Das Ziel dieser Arbeit ist daher die Erforschung von effizienten Approaches zur Unsicherheiten- und Sensitivitätsanalyse für ein robustes Prozessdesign der pharmazeutische Industrie. Methoden, Point Estimate Method (PEM) und Polynomial Chaos Expansion (PCE), wurde implementiert, um effizient Unsicherheitenquantifizierung (UQ) zu erlauben.

Der globalen Sensitivitätsanalyse (GSA) ist eine systematische Quantifizierung von Parameterschwankungen auf die Simulationsergebnisse. Verschiedene GSA Techniken werden im Detail vorgestellt, Schwierigkeiten bei der effizienten Berechnung unter Parameterkorrelationen aufgezeigt und Lösungsstrategien an Beispielen demonstriert. Die Ergebnisse zeigen sowohl den Mehrwert der GSA im Kontext des robusten Prozessdesigns als auch die Relevanz zur korrekten Berücksichtigung von Parameterkorrelationen bei der Sensitivitätsanalyse im Allgemeinen.

Um den schädlichen Einfluss von PU auf die modellgestützte Prozessauslegung zusätzlich zu minimieren, wurden weitere Konzepte aus der robusten Optimierung und Unsicherheitenbeschreibung untersucht. Zunächst wurde das erste Kozept basierend auf einer generalisierten PEM-Strategie entwickelt. Hier konnte gezeigt werden, dass die Zuverlässigkeit des robuste Prozessdesign bei einem deutlich reduzierte Rechenaufwand gesteigert werden kann. Parameterkorrelationen und Nebenbedinungen haben sich bei der robusten Prozessauslegung als relevante Einflussgrößen bestätigt und wurden entsprechend in den Simulationen berücksichtigt. In einem zweiten Schritt wurde das Konzept auf nicht normalverteilte PU erweitert. Ein neuer Ansatz, der die Gaußsche Mischverteilung mit der PEM kombiniert, wurde hierzu vorgeschlagen und erfolgreich implementiert. Weiterhin wurde eine iterative Back-off-Strategie erforscht, die durch eine vollständige Berücksichtigung der Unsicherheiteninformation zu besseren Resultaten bei der robusten Prozessauslegung führt. Durch die Entkoppelung von Unsicherheitenquantifizierung und Optimierung können wesentlich komplexere pharmazeutische Herstellungsprozesse mit einer hohen Anzahl an unsicheren Parametern implementiert werden.

Die in dieser Arbeit untersuchten Beispiele beinhalten verfahrenstechnische Grundoperationen und decken somit einen Großteil der gesamten Produktionskette der pharmazeutischen Herstellung ab. Die Ergebnisse der verschiedenen Robustifizierungskonzepte zeigen deutlich den Einfluss von PU und -korrelationen auf das modellgestützte Prozessdesign auf. Mithilfe der vorgeschlagenen Approaches können die Unsicherheiten effektiv und effizient bei einer optimalen Balance von Rechenaufwand und der geforderten Zuverlässigkeit ganz im QbD-Sinne berücksichtigt werden.

Preface

This thesis was written based on the following journal publications and conference proceedings, which have been published in scientific journals or presented at conferences, as be listed in what follows.

Journal Publications

1. Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Robust design of chemical processes based on a one–shot sparse polynomial chaos expansion concept. *Computer Aided Chemical Engineering*, 40(1):613–618, 2017

2. Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Toward a comprehensive and efficient robust optimization framework for (bio)chemical processes. *Processes*, 6(10):183, 2018

3. Xiangzhong Xie, Ulrike Krewer, and René Schenkendorf. Robust optimization of dynamical systems with correlated random variables using the point estimate method. *IFAC-PapersOnLine*, 51(2):427–432, 2018

4. Xiangzhong Xie, Rüdiger Ohs, Antje Spieß, Ulrike Krewer, and René Schenkendorf. Moment-independent sensitivity analysis of enzyme-catalyzed reactions with correlated model parameters. *IFAC-PapersOnLine*, 51(2):753–758, 2018

5. René Schenkendorf, Xiangzhong Xie, and Ulrike Krewer. An efficient polynomial chaos expansion strategy for active fault identification of chemical processes. *Computers & Chemical Engineering*, 122:228–237, 2018

6. René Schenkendorf, Xiangzhong Xie, Moritz Rehbein, Stephan Scholl, and Ulrike Krewer. The impact of global sensitivities and design measures in model-based optimal experimental design. *Processes*, 6(4):27, 2018

7. Nan Lin, Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Efficient global sensitivity analysis of 3d multiphysics model for li-ion batteries. *Journal of The Electrochemical Society*, 165(7):A1169–A1183, 2018

8. Victor N. Emenike, Xiangzhong Xie, René Schenkendorf, Antje C. Spiess, and Ulrike Krewer. Robust dynamic optimization of enzyme-catalyzed carboligation: A point estimate-based back-off approach. *Computers & Chemical Engineering*, 121:232–247, 2019

9. Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Efficient sensitivity analysis and interpretation of parameter correlations in chemical engineering. *Reliability Engineering & System Safety*, 187:159–173, 2019

10. Xiangzhong Xie and René Schenkendorf. Robust optimization of a pharmaceutical freeze-drying process under non-gaussian parameter uncertainties. *Chemical Engineering Science*, 207:805–819, 2019

11. Xiangzhong Xie and René Schenkendorf. Stochastic back-off robust optimization for continuous pharmaceutical processes using polynomial chaos expansion. *Computers & Chemical Engineering*, 124:80–92, 2019

12. Xiangzhong Xie and René Schenkendorf. Robust process design in pharmaceutical manufacturing under batch-to-batch variation. *Processes*, 7(8):509, 2019

Conference Proceedings

1. Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Robust design of chemical processes based on a one-shot sparse polynomial chaos expansion concept. In *10th World Congress of Chemical Engineering (WCCE10) & 27th European Symposium on Computer Aided Process Engineering (ESCAPE 27)*, Barcelona, Spain, 1 - 5 October, 2017

2. René Schenkendorf, Xiangzhong Xie, and Ulrike Krewer. An efficient polynomial chaos expansion strategy for active fault identification of chemical processes. In *10th World Congress of Chemical Engineering (WCCE10) & 27th European Symposium on Computer Aided Process Engineering (ESCAPE 27)*, Barcelona, Spain, 1 - 5 October, 2017

3. Xiangzhong Xie, Ulrike Krewer, and René Schenkendorf. Robust optimization of dynamical systems with correlated random variables using the point estimate

method. In *9th Vienna International Conference on Mathematical Modelling (MATHMOD)*, Vienna, Austria, 21 - 23 February, 2018

4. Xiangzhong Xie, Rüdiger Ohs, Antje Spieß, Ulrike Krewer, and René Schenkendorf. Moment-independent sensitivity analysis of enzyme-catalyzed reactions with correlated model parameters. In *9th Vienna International Conference on Mathematical Modelling (MATHMOD)*, Vienna, Austria, 21 - 23 February, 2018

5. Xiangzhong Xie and René Schenkendorf. The effect of correlated kinetic parameters on (bio)chemical reaction networks. In *DECHEMA - Jahrestreffen Reaktionstechnik*, Würzburg, Germany, 7 - 9 May, 2018

6. Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Robustes Prozessdesign in der Pharmatechnik mittels performanter Ersatzfunktionen. In *ProcessNet-Jahrestagung und 33. DECHEMA-Jahrestagung der Biotechnologen*, Aachen, Germany, 10 - 13 September, 2018

7. René Schenkendorf, Xiangzhong Xie, and Ulrike Krewer. Robustifizierung und Informationsmetriken der modellgestützten Versuchsplanung. In *ProcessNet Jahrestagung und 33. DECHEMA Jahrestagung der Biotechnologen*, Aachen, Germany, 10 - 13 September, 2018

Glossary

APIs	Active Pharmaceutical Ingredients
α_k	coefficients for mutivariate polynomial basis
A	set of all possible index for multivariate polynomial basis
b	back-off term
Cov	Covariance function
δ_{mn}	Kronecker delta function
$\mathbf{\Delta}$	correction term in LAR
D	distance function
δ_i	sensitivity indicator for MISA
h_{eq}	inequality constraint function
η	factor for calculating the back-off term
e_c	probability of constraints violation
$\mathcal{F}$	σ-algebra
F	cumulative density function of random variable
f	probability density function of random variable
g	process function
$G_{ij}(\theta_i, \theta_j)$	partial function from function decomposition
$\mathbb{I}_{\boldsymbol{\xi}}$	domian of random variables $\boldsymbol{\xi}$
h_{nq}	inequality constraint function
$\mathbf{L}$	lower triangular matrix from Cholesky decomposition
$\boldsymbol{\mu}_n$	n^{th} order statistical moment
$\mathcal{N}$	Guassian distribution
θ_n	nominal value of random variable θ
N_C	total number of component distributions
n	dimension of the vector
Ω	sample space
ω_j	weight for ith component distribution

Pr	probability measure
$\Psi_k(\boldsymbol{\xi})$	k^{th} multivariate polynomial basis of random variables $\boldsymbol{\xi}$
$\overline{p}$	order of the multivariate polynomial basis
p	time-invariant parameter
$\Phi(\cdot)$	tranformation function
$\Phi_i^{k_i}(\xi_i)$	k_i order univariate polynomial basis of random variable ξ_i
S_u^U	structural sensitivity indices
S_u^C	correlative sensitivity indices
S_u^{cov}	total covariance-based sensitivity indices
S_i^{uc}	first-order sensitivity indices
S_{ij}^{uc}	interaction sensitivity indices
$S_{T_i}^{uc}$	total sensitivity indices
Σ	covariance matrix
$s(\theta_i)$	shift function
θ	random variables
t	system time
u	control input variable
$\overline{\varepsilon}^*$	approximation error in LAR
ε_{eq}	tolerance factor for equality constraint
$\alpha_{\overline{h}_{nq},4}$	kurtosis of the inequality constraint
$\mu_{\overline{h}_{nq}}$	mean of the inequality constraint
$\alpha_{\overline{h}_{nq},3}$	skewness of the inequality constraint
ε_{nq}	tolerance factor for inequality constraint
$\sigma^2_{\overline{h}_{nq}}$	variance of the inequality constraint
V_{ij}	partial variance from variance decomposition
Var	Variance function
ϑ	scaling value for PEM
w_i	scalar weight factor for PEM
x	state variable
ξ	random variable
y	model output
CDF	Cumulative Density Function

CPM	Continuous Manufacturing Process
CQA	Critical Quality Attributes
CoDSA	Covariance Decomposition-based Sensitivity Analysis
EM	Expectation-Maximization
FIM	Fisher Information Matrix
GF	Generator Function
GMD	Gaussian Mxture Distributions
GQ	Gaussian Quadrature
GSA	Global Sensitivity Analysis
LAR	Least Angle Regression
MCs	Monte Carlo simulations
OED	Optimal Experimental Design
PCE	Polynomial Chaos Expansion
PDF	Probaility Density Function
PEM	Point Estimate Method
RO	Robust Optimization
SSI	Sobol' Sensitivity Indices
UQ	Uncertainty Quantification

Chapter 1

Introduction [1]

1.1 Research Motivation

Manufacturing processes in pharmaceutical industries persuade higher efficient and reliability to be prepared for the stringent competition in global market. Mathematical models, as the key element in the industrial digitalization, are implemented to predicting and forecasting the process status and behavior, which provides the potential for further improve the process performance with low investment costs [20, 21, 22]. The respective physio-chemical phenomena in pharmaceutical processes are described with continuity equations and associated model parameters, which construct in turn the mathematical models (first-principle) [23, 24, 25]. However, the model parameters estimated numerically are uncertain as a consequence of measurement noise and model assumptions(simplifications) [26, 6]. As illustrated in Fig. 1.1, on the account of measurement noises and model simplifications, a perfect agreement between model predictions and experiment measurements could not be achieved no matter how the parameter values are manipulated [27]. In other words, finding a unique value for parameters which lead to zero deviations between model predictions and experimental measurements is not possible in case of using real data. Therefore, the measurement data are accounted as random variables instead of deterministic quantities [27]. And the parameter values are estimated also as random variables, as shown in Fig. 1.1 in which the 99% confidence intervals (CI) indicates the region that includes most of the possible parameter values. The model predictions are also not deterministic as the parameters are uncertain, and vary in a certain range, i.e., 99%CI, in which almost all possible model predictions with different parameter values are included. Indeed, the bandwidth of 99% CI of model prediction indicates the credibility of the model, i.e., a more credible model has a narrower

[1] Part of this chapter has been published in (Xie et al., Reliab. Eng. Syst. Safe., 2018 [9]), (Xie et al., Processes, 6(10), 183, 2018 [2]), (Xie et al., Chem. Eng. Sci., 207, 805-819, 2019 [10]) and (Xie et al., Comput. Chem. Eng., 124, 80-92, 2019 [11])

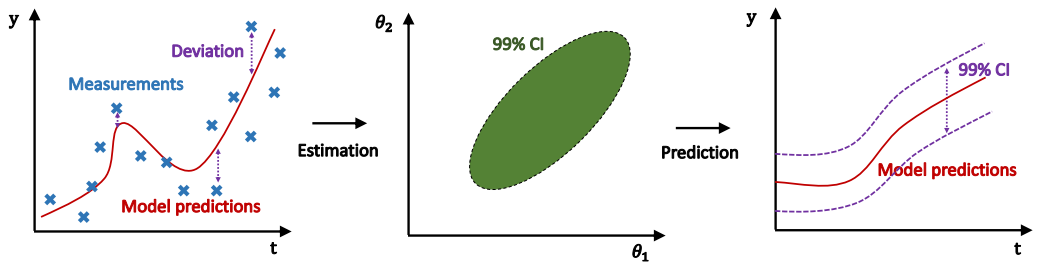


Figure 1.1: Illustration of the imprecision in parameter estimation. Step 1: Comparison between experiment measurements and model predictions. The existing deviation between them could, e.g., come from model simplification and measurement noise. Step 2: The 99% confidence interval (CI) of the estimated parameter values. Step 3: Model predictions with the associated 99% CI.

bandwidth for CI of predictions at different operation conditions. Optimal experimental design (OED) has been proposed and used to increase the model credibility by reducing the uncertainties in the estimated parameter values with informative experiments [27, 6]. However, as also pointed out in [27], the parameter uncertainty could be alleviated but not eliminated, and as such the vagueness in model predictions always exists. Therefore, model-based design of pharmaceutical process with single realization of parameter values and the corresponding individual trajectory of model predictions might result in a design solution with sub-optimal performance and high-risk potential [28, 29, 30].

The effect of parameter uncertainty on the design of pharmaceutical processes is further illustrated in Fig. 1.2. The traditional pattern of pharmaceutical processes is batch-wise and still widely implemented in pharmaceutical industries. The continuous pharmaceutical manufacturing process (CPM), shown in Fig. 1.2, is the pattern of future pharmaceutical process and has been advocated vigorously in the last decade [31]. It is suggested as the technology to modernize the pharmaceutical manufacturing according to the quality-by design concept which will be further explained in the subsequent section. More detailed discussions about how to transfer from traditional batch pattern to CPM and which process system engineering methods could be used to facilitate the transfer are referred to [32, 33]. Regardless of which pattern is used, the pharmaceutical processes consist of three main functional blocks which are synthesis, separation and formulation, as labeled in Fig. 1.2. The synthesis block contains different types of reactors for the synthesis of intermediate compounds and the desired active pharmaceutical ingredients (APIs) via chemical routines. The separation block includes extractor, crys-

tallizer and other type of separators to extract the APIs and remove impurities. These two blocks are referred as upstream and downstream process and normally overlapped to product APIs with desired purity [34]. The formulation block, i.e., downstream process, contains extrusion, granulation, drying and compaction processes to produce final drugs for better preservation and convenient dosage. The design of these three function blocks leads ultimately to the pharmaceutical manufacturing process with high product yield and good drug quality. Research about mathematical modeling and model-based analysis and design of individual function block [35, 36, 37] or even entire production process [25] could be found in the literature. However, the influence of parameter uncertainties is neglected in most of the works, albeit the common existence of them. As explained in Fig. 1.1, the real process behavior could be different from the model predictions with deterministic parameter values. Hence, the performance of the real process with deterministic design could be far away from the desired level which would lead to quality failure in the drug product. And more importantly, the process constraints could also be violated and lead to operation failure, as shown in Fig. 1.2. In short, the performance of process design solutions with deterministic parameter values is not reliable. In order to achieve more reliable model-based results, parameter uncertainties have to be taken into account in the design of pharmaceutical manufacturing processes.

In this work, the generic framework of model-based pharmaceutical manufacturing process development is categorized into three crucial stages, as illustrated in Fig. 1.3. Stage 1 is focused on the plant scale design. In this stage, the overall configuration of the production plant is designed. The mass and energy flow between synthesis, separation, formulation and heat integration compartments are optimized to achieve the desired product quality and yield [38]. In Stage 2, the production plant is divided into individual unit operations. Here, first-principle models are constructed for the unit operations with detailed mechanisms derived from the physio-chemical phenomena. The operation conditions for each unit operation are then designed based on the first-principle model, and the influence of parameter uncertainties is taken into account in Stage 3. In this stage, sensitivity analysis is used to quantify the influence of parameter uncertainties and screen out the ones with appreciable sensitivities. Moreover, robust process design provides the solutions, with which the best trade-offs between performance and robustness for both critical quality attributes and process constraints of the unit operations are achieved. The major focus of this thesis is to introduce and develop the methods needed for stages 3, i.e, sensitivity analysis and robust process design, to alleviate the influence of parameter uncertainties and ensure credible design of pharmaceutical processes.

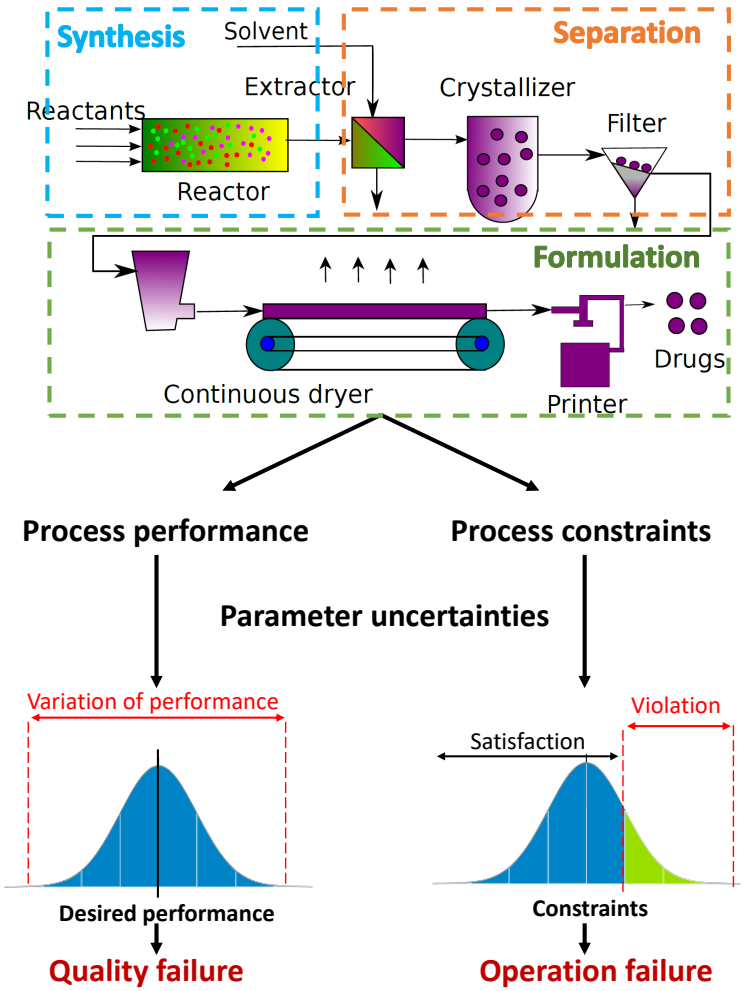


Figure 1.2: Impact of parameter uncertainty on the design of pharmaceutical manufacturing processes.

In the next sections, the history of development of pharmaceutical manufacturing is retrospected and its recent innovation is presented. Afterwards, the major focus of this work, i.e, sensitivity analysis and robust process design, is presented in more detail. Ultimately, the research scope and the structure of the thesis is presented at the end of this chapter.

1.2 Pharmaceutical Manufacturing

Pharmaceutical manufacturing plays a pivotal role in protecting health and saving many lives in the last centuries [39]. It develops and delivers life-saving drugs to patients for rescuing them from suffering of diseases [40]. From the last century until now, the industrialized production processes provide the capacity of pharmaceutical manufacturing to meet the requirements from patients all over the world [40]. However, at the early stage of last century, public health and even their lives were exposed to and threaten by the *medicine quality failures* from unaware contamination or improper manufacturing operations [41, 40]. In order to circumvent such critical situation, the Food and Drugs Administration (FDA) and the European Medicines Agency (EMA) were founded, which play the role of regular agencies and strictly supervise the manufacturing process to deliver qualified pharmaceutical products [39].

The initial focus of the regulatory agencies was on the toxicological analysis of new drug submissions to prevent the recurrence of historical incident in 1932 [40]. Later until 1941, in which the sulfathiazole disaster happened [41], the regulatory agencies realized the manufacturing processes had also to be supervised to deliver drug products with expected qualities. For this reason, procedural-based quality controls, which are also named as Good Manufacturing Practices (GMPs), and the associated guidance documents, which were released in the form of Internal Council on Harmonization (ICH) guidelines, were enacted for the inspection of pharmaceutical manufacturing [39]. Over the decades that followed, the pharmaceutical manufacturers were dedicated to ensure the compliance with GMPs, which hindered the technology development in pharmaceutical manufacturing [39].

With the laggard manufacturing technology and rigorous regulations, the expenditures on drug manufacture amount to more than twice of the research and development investment [42]. Regulatory agencies also sensed the urges on innovation of the current state of pharmaceutical manufacturing, and launched Quality by Design (QbD) [43] and Process Analytical Technology (PAT) initiatives [44] for Current Good Manufacturing Practices (cGMPs) for the 21^{st} Century.

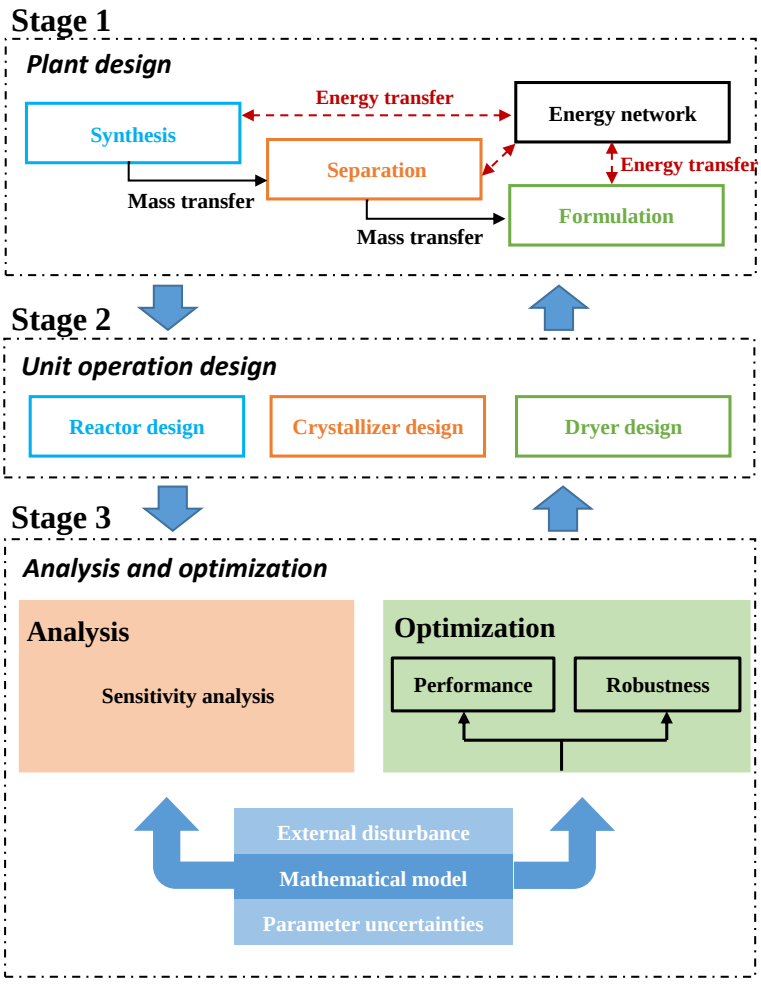


Figure 1.3: Design stages for model-based pharmaceutical manufacturing process development.

Traditionally or even nowadays, in most pharmaceutical manufacturing, the drug quality is still assessed by end-product testing [45]. However, such assessment cannot be relied on to ensure the product quality in the massive production environment [45]. Moreover, it is also not economic and sustainable to discard *"non-compliance"* product based on the unreliable assessment [42]. Thereby, the QbD initiative advocates that the quality should be *built in* rather than *tested in* the drug products [44]. It suggests the pharmaceutical manufacturers to adopt risk- and science- based approaches to pharmaceutical development and manufacturing to ensure a consistent and reliable quality of drug products with less regulatory oversight [42]. The PAT initiative recommends the pharmaceutical company to have in-depth understanding of the manufacturing processes and promotes the implementation of the technologies for designing, analyzing and controlling of manufacturing processes [44, 46].

After the two initiatives, the regulatory agencies introduced further concepts, such as Critical Quality Attributes (CQA), Critical Process Parameters (CPP), Design Space and Risk Assessment, as well as the associated tools to assist the implementation of QbD and PAT in pharmaceutical industries in the form of ICH guidelines [47, 48, 49, 50, 42]. Mathematical models and model-based technologies were encouraged by the regulatory agencies in the ICH guidelines [42]. They also emphasized the importance of incorporating model uncertainties and disturbances in the product and process design, which necessitates the development of stage 3 in Fig. 1.3, for a risk-based product and process design [50].

1.3 Process Analysis and Design with Uncertainties

In this section, the major focus of the thesis, sensitivity analysis and robust process design, are explained in more details.

1.3.1 Sensitivity Analysis

Sensitivity analysis (SA) provides quantitative measures for ranking the importance of parameters on model outputs, with which problem dimension could be reduced significantly by only taking into account the relevant parameters [51]. However, concept and implementation of SA are misinterpreted in many engineer fields, and as a consequence rarely applied, in practice, over there, which is also true for pharmaceutical manufacturing [52].

SA provides a in-depth understanding about the relevance of model parameters and model outputs, which plays an importance role in mathematical modeling [53, 54, 55]. It provides useful information for model reduction [56] and control strategy design [34]. Local and global SA are the two groups normally used for categorizing the SA methods. In local SA, the sensitivity measure is derived from the slope at the nominal value at which the model is linearized. Or it could also be interpreted as the change in model output if a comparatively small step change is added to the parameter at its nominal value. The advantages of local SA are 1) straightforward implementation and interpretation, 2) simple derivation without cumbersome model evaluations. However, information at single parameter point (nominal) and neglect of information from rest parameter space as well as parameter interactions would lead to inevitable loss of information, and as such inaccuracy and unreliable sensitivity measures [57]. On the other hand, global SA (GSA) provides quantitative measures based on the information from the entire parameter space, which ends up with more rigorous definition and capability of describing influence from parameter interactions [51, 55]. In literature, there are already plenty of works, in which local SA is implemented for the manufacturing processes, such as [58, 56, 34]. However, according to the work from [59] and [60], in which the sensitivity measures from local and global SA for chemical processes are compared, GSA is more preferable, especially for highly nonlinear systems, which turns out to be always true for the manufacturing processes. Therefore, GSA should also be the right choice for quantitative analysis of pharmaceutical manufacturing processes. Various measures for GSA were proposed and widely used for analyzing processes based on the assumption of independent parameters [61, 62, 37]. However, correlated parameters, which could be resulted from the inherent dependency or parameter estimation step, appear to be very common in manufacturing processes [59, 63, 64, 65]. And hence it might also be necessary to take them into account at the process analysis stage.

The first purpose of this thesis is to introduce and implement GSA techniques for quantitative analysis of the influence of parameter uncertainties on pharmaceutical manufacturing processes and investigate the impact of parameter correlations. The derived results from GSA can be used to facilitate the subsequent robust process design and also provide the engineers better understanding of the processes.

1.3.2 Robust Process Design

The aim of robust process design is to optimize the process performance and at the same time guarantee the critical constraints, which might be used to ensure process safety and

product quality, under the condition of parameter uncertainties. [66] firstly proposed the concept of robust process design and possible solution strategies, which has been, in the last decades, applied in the design of bio(chemical) processes [69, 70, 67, 68]. The most commonly used approach for robust process design is the so-called scenario-based approach. In this approach, the process is optimized upon on the worst-case scenario, which is derived by solving the robust counterpart [69, 71]. However, this approach suffers from two deficiencies. 1) A bilevel-optimization problem is formulated, which is not directly solvable with the traditional optimization methods. 2) the resulted robust solutions are normally too conservative with excessive scarification on the process performance. Alternatively, the probability-based approach is proposed, in which the uncertainties are described with probability distribution and the robustness from the design is adaptable by changing the threshold for violation probability [30, 3]. It provides adjustable trade-off between process performance and robustness in contrast to the scenario approach, and as such is a more flexible approach for robust process design to prevent excessive loss in process performance.

Uncertainty quantification (UQ) is the key tool for probability-based approach. It propagates the parameter uncertainties through the process model and the variation/probability distribution at model outputs are quantified subsequently. The statistical information of model outputs is then used in the objective function and constraints for process design. Traditional method for uncertainty quantification is the sample-based approaches, i.e, Monte Carlo simulations (MCs), quasi-MCs and Latin Hypercube sampling (LHS; [72]). Enormous amount of samples are drawn from the probability distribution of parameter uncertainties, at which the original computational expensive model is evaluated. The resulting model evaluation could then be used for quantifying the required statistical information of model outputs. This type of method is straightforward for implementation but requires extremely high computational expense and therefore might not be a good candidate to be embed into the robust process design approaches [73]. Alternatively, the Point Estimate Method (PEM, [74, 2]) and Polynomial Chaos Expansion (PCE, [75, 9]), which are more efficient methods for UQ, are presented as the promising solutions for robust process design.

In this thesis, novel and comprehensive frameworks for robust process design are proposed based on the PEM and PCE, respectively. The introduced frameworks are capable to design the processes with uncertain parameters associated with either Gaussian or non-Gaussian distribution at low computational cost. Additionally, the frameworks are implemented on various unit operations which cover the entire process chain of pharmaceutical manufacturing.

1.4 Scope and Structure of the Thesis

In general, the thesis aims to further fertilize the field of pharmaceutical manufacturing with the knowledge from process system engineering. Especially, the focus is to introduce the concept of global sensitivity analysis and robust process design for analysis and design of the pharmaceutical processes in a more efficient and effective way. Moreover, efficient and effective algorithms and solution strategies are introduced based on the PEM and PCE to increase the practicability of the concepts to embed the parameter uncertainty in the product and process development of pharmaceutical manufacturing. The structure of the remaining chapters of the thesis is outlined as follows:

Chapter 2 - Uncertainty Propagation and Quantification

The mathematical basics and implementation details about the Point Estimate Method (PEM) and Polynomial Chaos Expansion (PCE) are provided in this chapter. Moreover, their associated merits and drawbacks are also discussed.

Chapter 3 - Efficient Global Sensitivity Analysis

The basics of two techniques, variance-based and moment-independent GSA are presented firstly under the assumption of parameter independence. However, parameter correlations exist commonly within the uncertainties of model parameter, which are estimated from experimental data. Thereby, the techniques are adapted further to incorporate the effect of parameter correlations. Moreover, efficient computation of all the sensitivity measures are provided based on PCE introduced in Chapter 2. The techniques are demonstrated with a continuous synthesis process that is used to produce API-scaffold. The results manifest the significance of parameter correlations and the relevance of GSA in process analysis and design.

Chapter 4 - Comprehensive and Efficient Framework for Robust Process Design

A highly efficient framework for robust process design is proposed based on PEM. The usage of information obtained from GSA is demonstrated as to reduce the problem dimension. Two moment-based methods are compared in the approximation of robust inequality and equality constraints. Moreover, parameter correlations are also taken into account and their influence on robust process design is presented. The proposed framework is implemented in the chemical and biological synthesis steps of pharmaceutical manufacturing. The results show that the proposed framework for robust process design handles both independent and correlated parameter uncertainties adequately and scales well with the number of uncertain parameters.

Chapter 5 - Robust Process Design with Non-Gaussian Parameter Uncer-

tainties

In the previous chapter, transformation techniques were combined with the PEM to address non-Gaussian and correlated parameter distributions efficiently. However, the transformation step introduces additional nonlinearities and may increase approximation errors. Chapter 5 presents the idea of combining Gaussian mixture distributions (GMD), which decompose the parameter distribution into a finite set of Gaussian distributions, with the PEM to ensures a proper and effective uncertainty quantification for robust process design. The improved approach is applied to a freeze-drying process (lyophilization), which is a part of the formulation block. Results suggest that the novel GMD-PEM algorithm has the potential to outperform conventional approaches regarding credibility and efficiency.

Chapter 6 - Stochastic Back-off Robust Process Design

Limitations of the traditional strategy for robust process design is addressed. For example, even if the efficient algorithms are implemented, performing UQ at each optimization iteration turns the traditional strategy into a computationally expensive process, especially when poor initial guesses are provided for the optimization. In Chapter 6, thus, a novel stochastic back-off strategy is proposed to handle the potential flaws of the traditional strategy for robust process design. Moreover, the statistical information needed for the stochastic back-off strategy is provided by PCE. The proposed approach is demonstrated in the design of a plug-flow crystallizer which is a part of the separation block. The results manifest the efficiency and practicability of the stochastic back-off strategy for robust design of pharmaceutical processes.

Chapter 7 - Conclusions and Future Works

Chapter 7 provides a general summary of the work presented in this thesis, and also discusses the existing challenges and perspectives that could be further investigated in the future work.

Chapter 2

Uncertainty Propagation and Quantification[2]

In this chapter, two methods, which appear throughout this thesis for uncertainty propagation and quantification (UQ), are introduced, i.e., the Point Estimate Method (PEM) and the Polynomial Chaos Expansion (PCE). First, a brief introduction of different UQ methods and the reasons why these two are selected are presented. Afterwards, the mathematical basics and implementation details of these two methods are provided. Additionally, the generalization of these methods to parameter uncertainties with arbitrary probability distributions and dependencies is also presented.

2.1 Preliminaries

Generally, first-principle models are used to describe physio-chemical mechanisms of pharmaceutical processes mathematically. In the field of process system engineering, first-principle models typically consist of nonlinear different algebraic equations (DAEs) equal to:

$$\dot{\mathbf{x}}_{\mathbf{d}}(t) = \mathbf{g}_{\mathbf{d}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}), \tag{2.1}$$

$$\mathbf{0} = \mathbf{g}_{\mathbf{a}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}), \tag{2.2}$$

$$\mathbf{y} = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}), \tag{2.3}$$

$$\mathbf{x}_{\mathbf{d}}(0) = \mathbf{x}_0, \tag{2.4}$$

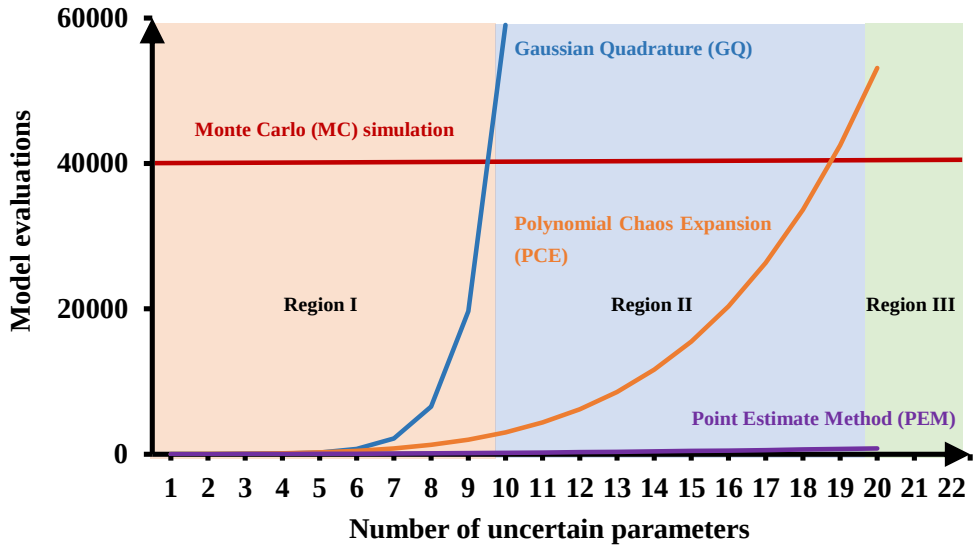
where $t \in [0, t_f]$ denotes the time, $\mathbf{u} \in \mathbb{R}^{n_u}$ the control input vector and $\mathbf{p} \in \mathbb{R}^{n_p}$ the time-invariant parameter vector. $\mathbf{x} = [\mathbf{x_d}, \mathbf{x_a}] \in \mathbb{R}^{n_x}$ is the state vector, while $\mathbf{x}_d \in \mathbb{R}^{n_{x_d}}$ and $\mathbf{x}_a \in \mathbb{R}^{n_{xa}}$ are the differential and algebra states, respectively. $\mathbf{y}$ is the model output vector. $\mathbf{x}_0$ is the vector of the initial conditions for the differential states. Furthermore,

[2]Part of this chapter has been published in (Xie et al., Reliab. Eng. Syst. Safe., 187, 159–173, 2019 [9]), (Xie et al., Processes, 6(10), 183, 2018 [2]), (Xie et al., IFAC-PapersOnline, 52(2), 427-432, 2018 [3])

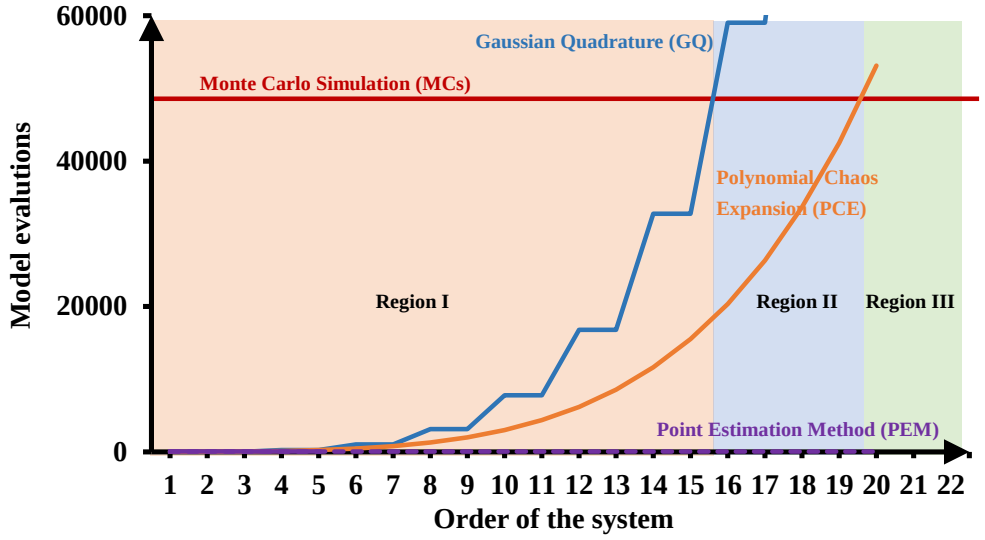
two types of functions $\mathbf{g_d} : \mathbb{R}^{(n_{x_d}+n_{x_a})\times n_u \times n_p} \rightarrow \mathbb{R}^{n_{x_d}}$ and $\mathbf{g_a} : \mathbb{R}^{(n_{x_d}+n_{x_a})\times n_u \times n_p} \rightarrow \mathbb{R}^{n_{x_a}}$ are given, which denote the differential vector field and algebraic expressions of the process model.

Typically, the time-invariant parameters $\mathbf{p}$ and initial conditions $\mathbf{x}_0$ are not known exactly. Measurement and process noise give rise to uncertainties in model parameters, which are estimated through model fitting [76, 63, 27]. In addition, disturbances from the environment and the accuracy of the measurement devices result in uncertain initial conditions. As it is intended to use random variables to describe the uncertainties in the parameters and the initial conditions, a probability space $(\Omega, \mathcal{F}, Pr)$ is defined with the sample space Ω, σ-algebra $\mathcal{F}$, and the probability measure Pr. $\boldsymbol{\theta} = [\mathbf{p}(\omega), \mathbf{x}_0(\omega)]$ is the vector of random variables, which are functions of $\omega \in \Omega$ on the probability space and associated with continuous probability density functions (PDFs) $\mathbf{f}(\boldsymbol{\theta}) = [f_1(\theta_1), \ldots, f_{n_\theta}(\theta_{n_\theta})]$ and correlation matrix Σ. These uncertainties, in turn, lead to variation in the system states $\mathbf{x}$ and outputs $\mathbf{y}$, which could also be characterized with probability distributions. However, quantifying the analytical representation of the probability distribution of system states and outputs is actually challenging due to the complex formulation of PDFs for parameter uncertainties and process models.

The importance of including uncertainty has been realized in many disciplines, e.g., civil engineering. Thereby, various methods were also introduced in the literatures to quantify the probability distribution of model outputs numerically [77]. Sampling-based methods, i.e., (quasi) Monte Carlo simulation (MCs), are the most commonly used approach for UQ. Samples of the uncertain parameters are randomly generated according to their PDFs. The process model is evaluated at each sample and the collection of solutions is used to approximate the statistical information of model outputs. Albeit the MCs method is simple and straightforward for implementation, the approximated statistical information converges slowly with increasing number of samples [78]. On that account, a large number of samples is needed for an accurate approximation, which could lead to excessive computational cost, as shown in Fig. 2.1. Fig. 2.1 provides an illustration of computational demand for different methods for different numbers of uncertain parameters and different system complexity. Gaussian quadrature (GQ), which was developed for solving numerical integration problems [79], is a common approach for computing the statistical moments of model outputs. It utilizes deterministic samples and associated weights, which are derived from orthogonal polynomials, and approximates the statistical moments through weighted superposition of the sample solutions. However, the need of tensor product formulas for multi-dimensional problems render an exponential increase of computational demand on increased number of uncertain param-



(a)



(b)

Figure 2.1: Computational demand (i.e., number of model evaluations) for different methods for uncertainty quantification with increasing (a) number of uncertain parameters and (b) system complexity to achieve similar approximation accuracy.

eters, as shown in Fig. 2.1. Lerner[80] suggested to truncate the expression of GQ at the cost of reduced accuracy on approximating the system interaction and proposed the Point Estimate Method as an alternative for efficient calculation of statistical moments. As can be seen from Fig. 2.1, the increase of computational demand with increasing number of uncertain parameters for PEM is negligible compare to the other methods. But its accuracy is not guaranteed if the model is highly nonlinear. Spectral methods, e.g., polynomial chaos expansion [81, 82], are one of the most widely used methods for UQ. They represents the random model outputs with weighted superposition of orthogonal polynomials of uncertain parameters. PCE has moderate computational demand but still suffers the issue of "curse-of-dimensionality". Nevertheless, a recently proposed algorithm, in which the sparsity effect of PCE is used, provides the possibility of implementing PCE on systems with a large number of uncertain parameters [61, 83]. In the rest of this chapter, mathematical basics and implementation details about PEM and PCE are presented, as they are more likely to be able to quantify the influence of parameter uncertainty for both sensitivity analysis and robust process design with affordable costs.

2.2 Point Estimate Method

The Point Estimate Method (PEM) is a sample-based and an efficient cubature rule for approximating n-dimensional integrals [80, 74, 84]. It is analogous to the concept of the so-called unscented transformation presented by [85], which describes the parameter uncertainty with some deterministic sample points and approximates the statistics of outputs with the corresponding model evaluations, but has different deterministic sample points, associated weights and higher accuracy [80]. The PEM has been successfully applied in the field of sensitivity analysis [86] and optimal experimental design [87, 88, 6] to quantify the influence of measurement imperfections on system identification. A brief introduction to the PEM is given in Section 2.2.1. The concept of extending the PEM to problems with arbitrary and correlated parameter uncertainties is presented in Section 2.2.2.

2.2.1 Basics of the Point Estimate Method

The basic principle of the PEM is illustrated in Figure 2.2. Here, a nonlinear function $\mathbf{k}(\cdot)$ with a two-dimensional parameter $[\xi_1, \xi_2]$ and one model output y is used for demonstration. It is assumed that the two parameters have a bivariate standard Gaus-

sian distribution $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The probability distribution of the parameters does not have to be Gaussian and could be Uniform, Beta distribution or any other parametric distributions; as long as it is symmetric and independent [6]. The reason why the probability distribution of parameters has to be symmetric and independent is provided in [80] and also mentioned later. According to the PEM, nine deterministic sample points, i.e., the cross, circle and star points in Figure 2.2, are generated and used for function evaluations. Afterwards, the integral term is approximated by a weighted superposition of these function evaluations equal to:

$$\int_{\mathbb{I}_{\boldsymbol{\xi}}} \mathbf{k}(\boldsymbol{\xi}) f(\boldsymbol{\xi}) d\boldsymbol{\xi} \approx \sum_{i=1}^{n_p} w_i \mathbf{k}(\boldsymbol{\xi}_i^s), \tag{2.5}$$

where $\boldsymbol{\xi}_i^s$ denotes the i-th sample point; n_ξ and n_p denote the number of random inputs and sample points, which are equal to two and nine in this example; w_i is a scalar weight factor; and $f(\boldsymbol{\xi})$ is the PDF of the uncertain parameters.

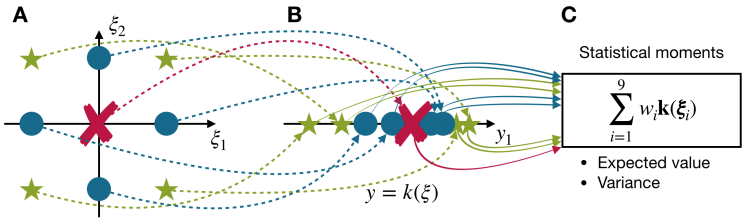


Figure 2.2: Illustration of the point estimation method (PEM) for a nonlinear function $\mathbf{y} = \mathbf{k}(\boldsymbol{\xi})$ that has (**A**) two random inputs; (**B**) one model output y_1; and (**C**) the resulting approximations of statistical moments of y_1.

This figure is originated from [2] and created by René Schenkendorf

The deterministic sample points from PEM are generated by the so-called generator functions (GF) with the scaling value ϑ. The GF generates sample points in $\mathbb{R}^n$ by permutation and the change of sign-combinations of the coordinate values [80]. For instance, the first three GFs (GF[0], GF[$\pm\vartheta$], GF[$\pm\vartheta, \pm\vartheta$]) for sampling in $\mathbb{R}^2$ is given in Eqs. (2.6) to (2.8), which are corresponding to the samples marked with cross, circle and star symbols in Fig. 2.2, respectively. The PEM, as defined in [80], only uses the first three GFs to generate deterministic samples in $\mathbb{R}^{n_\xi}$.

$$GF[0] = \{(0,0)\}, \tag{2.6}$$

$$GF[\pm\vartheta] = \{(\vartheta,0), (-\vartheta,0), (0,\vartheta), (0,-\vartheta)\}, \tag{2.7}$$

$$GF[\pm\vartheta, \pm\vartheta] = \{(\vartheta, \vartheta), (\vartheta, -\vartheta), (-\vartheta, \vartheta), (-\vartheta, -\vartheta)\}, \tag{2.8}$$

This leads to an overall number of $n_p = 2n_\xi^2 + 1$ sample points. Moreover, the weight factors for the evaluations with the samples from the same GF are the same by definition. Therefore, the formulation for the approximation of the integral term could be rewritten in a more specific way as shown in Eq. (2.9).

$$\begin{aligned} &\int_{\mathbb{I}_{\boldsymbol{\xi}}} \mathbf{k}(\boldsymbol{\xi}) f(\boldsymbol{\xi}) d\boldsymbol{\xi} \approx \\ &w_0 \mathbf{k}(GF[0]) + w_1 \sum \mathbf{k}(GF[\pm\vartheta]) + w_2 \sum \mathbf{k}(GF[\pm\vartheta, \pm\vartheta]), \end{aligned} \tag{2.9}$$

Here, in the formulation of PEM approximation, the scaling value ϑ and scalar weight factors w_0, w_1 and w_2 are still not known. As mentioned in [80], they are determined by substituting the function $\mathbf{k}(\boldsymbol{\xi})$ with even-order monomials $\mathbf{k}(\boldsymbol{\xi}) = 1$, $\mathbf{k}(\boldsymbol{\xi}) = \xi_i^2$, $\mathbf{k}(\boldsymbol{\xi}) = \xi_i^4$, $\mathbf{k}(\boldsymbol{\xi}) = \xi_i^2 \xi_{j\neq i}^2$ $(i, j \in \{1, \cdots, n_\xi\})$ as follows.

$$w_0 + 2n_\xi w_1 + 2n_\xi(n_\xi - 1)w_2 = \int_{\mathbb{I}_{\boldsymbol{\xi}}} 1 f(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{2.10}$$

$$2w_1\vartheta^2 + 4(n_\xi - 1)w_2\vartheta^2 = \int_{\mathbb{I}_{\boldsymbol{\xi}}} \xi_i^2 f(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{2.11}$$

$$2w_1\vartheta^4 + 4(n_\xi - 1)w_2\vartheta^4 = \int_{\mathbb{I}_{\boldsymbol{\xi}}} \xi_i^4 f(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{2.12}$$

$$4w_2\vartheta^4 = \int_{\mathbb{I}_{\boldsymbol{\xi}}} \xi_i^2 \xi_{j\neq i}^2 f(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{2.13}$$

In order to solve these equations to calculate the scaling value and scalar weight factors, the probability distribution of uncertainties in parameter $\boldsymbol{\xi}$ has to be specified. According to the previous works [80, 27], the standard Gaussian distribution is used and the calculated values from solving Eqs. (2.10) to (2.13) are $\vartheta = \sqrt{3}, w_0 = 1 + \frac{n_\xi^2 - 7n_\xi}{18}, w_1 = \frac{4 - n_\xi}{18}, w_2 = \frac{1}{36}$. Please note that other types of standard probability distributions, which are symmetric and independent, could also be used. But the values for ϑ, w_0, w_1 and w_2 have to be recalculated as shown in [6]. For odd-order monomials, such as ξ_i^1, ξ_i^3, $\xi_i^1 \xi_j^2$, etc, Eq.(2.9) is always valid as long as the probability distribution of uncertainties in parameter $\boldsymbol{\xi}$ is independent and symmetric [80]. Hence, the PEM formulation is valid for all odd-order monomials and even-order monomials up to 4, the overall precision of PEM is guaranteed for the integral of functions with moderate nonlinearities, i.e., system up to fifth-order [80, 3]. Please note that the system with fifth-order means it can be accurately approximated with the sum of monomials up to order of five. In principle, the PEM can also be adapted to ensure lower or higher precisions, but the proposed setting has the best trade-off between precision and computational cost [74].

2.2.2 Sampling Strategy for Independent/Correlated Random Variables of Arbitrary Distributions

As mentioned above, the proposed PEM with the provided factor values is only applicable in the circumstance of independent standard Gaussian distributions describing the parameter uncertainties. For most practical applications, however, they confront with problems in which the parameter uncertainties have arbitrary and correlated probability distributions. Therefore, the PEM is extended by using Proposition 2.1.

Proposition 2.1. *For two random variables $(\boldsymbol{\theta}, \boldsymbol{\xi})$, where $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\boldsymbol{\theta}$ has an arbitrary distribution, and the function $\Phi(\cdot) = F_{\boldsymbol{\theta}}^{-1}(F_{\boldsymbol{\xi}}(\cdot))$, the following relation for the integral terms of the nonlinear function $k(\boldsymbol{\theta})$ holds [3]:*

$$\int_{I_{\boldsymbol{\theta}}} k(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} = \int_{I_{\boldsymbol{\xi}}} k(\Phi(\boldsymbol{\xi})) f(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{2.14}$$

Based on Proposition 2.1, the integral of function $k(\boldsymbol{\theta})$ is approximated as:

$$\int_{\mathbb{I}_{\boldsymbol{\theta}}} \mathbf{k}(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \approx w_0 \mathbf{k}(\Phi(\boldsymbol{\xi}^1)) + w_1 \sum_{i=2}^{2d+1} \mathbf{k}(\Phi(\boldsymbol{\xi}^i)) + w_2 \sum_{j=2d+2}^{2d^2+1} \mathbf{k}(\Phi(\boldsymbol{\xi}^j)), \tag{2.15}$$

where the samples from the original PEM for $\boldsymbol{\xi}$ are transformed via $\Phi(\cdot) = F_{\boldsymbol{\theta}}^{-1}(F_{\boldsymbol{\xi}}(\cdot))$ to the corresponding points in $\boldsymbol{\theta}$, which can be directly evaluated with function $k(\cdot)$. The joint cumulative density function (CDF) $F_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ in $\Phi(\cdot)$ is typically unknown in practical applications and derived from marginal CDFs $[F_1(\theta_1), \ldots, F_d(\theta_d)]$ and the correlation matrix $\Sigma \in \mathbb{R}^{d \times d}$ for the uncertain parameter $\boldsymbol{\theta}$. Please note that it is actually infeasible to derive an analytical expression for $F_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and $\Phi(\cdot)$ [3]. Thus, Algorithm 1 is introduced to transform the samples from $\boldsymbol{\xi}$ to $\boldsymbol{\theta}$ numerically. The transformed sample points can be directly used for the approximation scheme:

$$\int_{\mathbb{I}_{\boldsymbol{\theta}}} \mathbf{k}(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \approx w_0 \mathbf{k}(\boldsymbol{\theta}^1) + w_1 \sum_{i=2}^{2d+1} \mathbf{k}(\boldsymbol{\theta}^i) + w_2 \sum_{j=2d+2}^{2d^2+1} \mathbf{k}(\boldsymbol{\theta}^j). \tag{2.16}$$

Algorithm 2.1 is derived from the Nataf transformation procedure, which is based on Gaussian-copula [89]. By definition, the Gaussian-copula concept needs only the marginal distributions and the correlation matrix to approximate multivariate distributions. Technically, the Gaussian-copula is used for describing multivariate distributions with linear correlation, and thus might lose accuracy in describing multivariate distributions with non-linear correlations.

Algorithm 2.1 Nataf transformation for deriving generalized PEM samples.

Initialization: Random variables $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, $I \in \mathbb{R}^{d\times d}$; $\boldsymbol{\theta}$ have marginal cumulative density functions $[F_1(\theta_1), \ldots, F_d(\theta_d)]$ and correlation matrix $\Sigma \in \mathbb{R}^{d\times d}$; F is the cumulative density function of standard Gaussian distribution.

1: Sample $\mathbf{U} = [\boldsymbol{\xi}^1, \cdots, \boldsymbol{\xi}^N]$ with size of $N = 2d^2 + 1$ from $\boldsymbol{\xi}$ and dimension d from Generator function $GF[\cdot]$;

2: Cholesky decomposition of $\Sigma = \mathbf{L}\mathbf{L}^T$, where L is a lower triangular matrix;

3: Correlate the sample, $\mathbf{V} = \mathbf{L}\mathbf{U}$;

4: Convert the sample to the corresponding cumulative density $\mathbf{W} = [F(V_1), \cdots, F(V_d)]^T$;

5: Transform into sample of $\boldsymbol{\theta}$, $[\boldsymbol{\theta}^1, \cdots, \boldsymbol{\theta}^N] = [F_1^{-1}(W_1), \cdots, F_d^{-1}(W_d)]^T$.

2.3 Polynomial Chaos Expansion

The Polynomial Chaos Expansion (PCE) was introduced back in the 1930s [81] and was populated for solving stochastic finite element problems in the field of engineering applications [90]. It was firstly constructed for standard normal random variables and then generalized to other type of standard random variables, e.g. Uniform and Beta distribution, [82]. The method is an efficient tool for taking care of stochastic problems and has been further implemented in many other disciplines for accounting the influence of uncertainties (e.g., [75, 83]). The general structure of PCE is given in Section 2.3.1 and methods for computing PCE coefficients for individual problem is presented in Section 2.3.2.

2.3.1 Polynomial Basis

In PCE, any function of finite variance $G(\boldsymbol{\xi}) \in L^2(\Omega, \mathcal{F}, Pr)$, can be represented as [81]:

$$\begin{aligned} G(\boldsymbol{\xi}) =& a_0\Gamma_0 + \sum_{i_1=1}^{n_\xi} a_{i_1}\Gamma_1(\xi_{i_1}) + \sum_{i_1=1}^{n_\xi}\sum_{i_2=1}^{i_1} a_{i_1,i_2}\Gamma_2(\xi_{i_1},\xi_{i_2}) + \cdots \\ &+ \sum_{i_1=1}^{n_\xi}\sum_{i_2=1}^{i_1}\cdots\sum_{i_p=1}^{i_{p-1}} a_{i_1,i_2,\cdots,i_p}\Gamma_p(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_p}) + \cdots, \end{aligned} \tag{2.17}$$

where Γ_p is the p-th order multivariate polynomial for $\boldsymbol{\xi} = \{\xi_i\}_{i=1}^{n_\xi}$. $\boldsymbol{\xi}$ are independent random variables with standard normal distribution. $a_{i_1 \cdots, i_P}$ are the coefficients which quantify the relation between the corresponding polynomial Γ_P and the function $G(\boldsymbol{\xi})$. PCE can also be directly written as:

$$G(\boldsymbol{\xi}) = \sum_{k=0}^{\infty} \alpha_k \Psi_k(\boldsymbol{\xi}), \tag{2.18}$$

where the $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^{\infty}$ and $\{\alpha_k\}_{k=0}^{\infty}$ are multivariate polynomials and corresponding coefficients of infinite degree, respectively. k is the number of the polynomials. The multivariate polynomials $\Psi_k(\boldsymbol{\xi})$ are constructed by the product of the univariate polynomials [82]:

$$\Psi_k(\boldsymbol{\xi}) = \Phi_1^{k_1}(\xi_1)\Phi_2^{k_2}(\xi_2)\cdots\Phi_{n_\xi}^{k_{n_\xi}}(\xi_{n_\xi}), \tag{2.19}$$

where nonnegative integer $\{k_i\}_{i=1}^{n_\xi}$ indicate the individual order of the univariate polynomials $\{\Phi_i^{k_i}(\xi_i)\}_{i=1}^{n_\xi}$. The order of the multivariate polynomials $\Psi_k(\boldsymbol{\xi})$ is the sum according to $\overline{p} = k_1 + k_2 + \cdots + k_{n_\xi}$. The multivariate and univariate polynomials are orthogonal with respect to their corresponding stochastic measure:

$$\int \Phi_i^m(\xi_i)\Phi_i^n(\xi_i)f(\xi_i)d\xi_i = r_m\delta_{mn}, \tag{2.20}$$

$$\int \Psi_m(\boldsymbol{\xi})\Psi_n(\boldsymbol{\xi})f(\boldsymbol{\xi})d\boldsymbol{\xi} = \gamma_m\delta_{mn}, \tag{2.21}$$

$$f(\boldsymbol{\xi}) = f(\xi_1)f(\xi_2)\cdots f(\xi_{n_\xi})d\xi_1 d\xi_2 \cdots d\xi_{n_\xi}, \tag{2.22}$$

where δ_{mn} is the Kronecker delta function which is 1 for identical values of m and n and 0 for the others. r_m and γ_m are the normalized constants of Φ_m and Ψ_m, respectively.

The type of univariate polynomials depends on the probability distributions of the random variables. [81] introduced Hermite polynomials for Gaussian random variables. Although the model can be used for other distributions through isoprobabilistic transformation, the efficiency and accuracy of the model are decreased considerably [77]. Alternatively, the Askey scheme was used to construct orthogonal polynomials for other types of statistical distribution, e.g., Jacobi polynomials for beta distribution, as shown in [82]. Subsequently, PCE has been further adapted to arbitrary distributions by [91, 92] and [93]. The general formulation of PCE works only for independent variables, but PCE can also be constructed for correlated variables [94].

Practically, instead of using a set of an infinite number of polynomials as in Eq. (2.18), a finite number of polynomials is retained to approximate the random variable [95], e.g., polynomials with a total order not exceeding $\overline{p}_{max}$:

$$G(\boldsymbol{\xi}) \approx \sum_{k=0}^{P-1} \alpha_k \Psi_k(\boldsymbol{\xi}), \tag{2.23}$$

where $\overline{P}$ is the dimension of the polynomial basis depending on maximum order $\overline{p}_{max}$ and n_ξ:

$$P = \binom{n_\xi + \overline{p}_{max}}{\overline{p}_{max}} = \frac{(n_\xi + \overline{p}_{max})!}{n_\xi! \overline{p}_{max}!} \tag{2.24}$$

Since the dimension of the input variable (n_ξ) is fixed for certain problems, the optimal option for $\overline{p}_{max}$ is the minimum value needed to guarantee the target accuracy.

2.3.2 PCE Coefficient Calculation

Estimation of the coefficients $\{\alpha_k\}_{k=0}^{P-1}$ plays an essential role in PCE-based calculations as the stochastic properties of PCE are mainly characterized by them. Several methods are available for computing the coefficients, and they are classified into two groups: intrusive methods and non-intrusive methods. Intrusive methods, such as Galerkin projection, have optimal accuracy but require adaptation of the numerical model, which might be challenging to implement for chemical processes [77]. In contrast, non-intrusive methods require only model evaluations for some realizations and can be applied for models of any complexity. The number of required model evaluations, however, might restrict the implementation.

In the vast majority of the cases, heterogeneity in the coefficients of PCE reflects the differences regarding significance of the polynomial basis in representing the original model. For example, polynomial basis, which represent high–order interactions between uncertain parameters, could be close to zero, as they are irrelevant in improving the accuracy of the PCE [96]. Therefore, the number of relevant polynomial basis in PCE could be relatively small compared to the number of the full set. In order to avoid the interference of those insignificant terms and reduce the number of model evaluations, the concept of sparse PCE is introduced by [61]. First, the set of polynomial basis is truncated with certain criteria to roughly remove the redundant terms. One truncated scheme has been already mentioned in the last section, where only the polynomial basis with highest total order of p are retained. Beside that, other truncated schemes, e.g.,

low rank truncated scheme and hyperbolic truncated scheme, could also be implemented ([96]). Moreover, if the sensitivities of the polynomial basis are known in advance, an anisotropic truncated scheme could also be provided for certain problems([61]). Second, the relevant polynomial basis are picked out from the truncated scheme by analyzing the correlation between them and the residual of approximation. Several approaches were described and compared in [61] and [97] for selecting the relevant polynomial basis. And the native and modified Least Angles Regression (LAR) [97] are suggested as a stable and efficient solution. The procedure for estimation of the PCE coefficients, which is developed from Least Angle Regression (LAR) [97, 61], is summarized in Algorithm 2.2.

Here, let us consider a sample set $\mathcal{S}=\{\boldsymbol{\xi}^1,\ldots,\boldsymbol{\xi}^N\}$ generated for random variables $\boldsymbol{\xi}$. $\mathbf{Y}=(G(\boldsymbol{\xi}^1),\ldots,G(\boldsymbol{\xi}^N))$ is the vector of the model evaluations associated with the sample set $\mathcal{S}$. The function $G(\boldsymbol{\xi})$ is approximated by $\sum_{k=0}^{P-1}\alpha_k\Psi_k(\boldsymbol{\xi})$, where P is the number of polynomials of the truncated scheme (Eq. (2.24)). Note that other truncated schemes can be selected based on the information from model structure [61].

The algorithm is initialized by function evaluations $\mathbf{Y}$ and of polynomials $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^{P-1}$ from the truncated scheme, which are evaluated for sample set $\mathcal{S}$. Note that the size of the sample set can be enlarged when the accuracy of the PCE is not satisfied. For each iteration, the size of the multivariate polynomials is enlarged while the maximum order of the polynomials increases from 1 to p_{max}, where p_{max} is the maximum order of the multivariate polynomials allowed for approximation. The most correlated polynomial is selected and moved to the active set at step 6. The coefficients for the polynomials of the active set are adapted in an optimal direction in step 9, using the correction term Δ [97]. All the possible active sets are validated, and the optimal one is selected by using cross validation in steps 11 and 12. Finally, the coefficients for the optimal PCE (step 18) are estimated through ordinary least squares regression. Note that the size of the active set, i.e., the number of multivariate polynomials used for the final estimation, is typically much smaller than the size of the full PCE model. This leads to a considerable reduction in the number of model evaluations. In principle, the resulting sparsity of the PCE model can be quantified by

$$Sparsity=\frac{\mathcal{A}(\text{Active set})}{\mathcal{A}(\text{Full set})}, \tag{2.25}$$

where $\mathcal{A}$ means the cardinality of the set.

Algorithm 2.2 Computing adaptive sparse PCE with LAR.

Initialization

Select the truncated scheme

Compute $\mathbf{Y}$ and $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^{P-1}$ for $\mathcal{S}$

Estimation of polynomial coefficients

for $p \leftarrow 1, p_{max}$ **do**

  Set $\mathbf{a} = \{\alpha_0, \ldots, \alpha_{P-1}\} = \mathbf{0}$, $\mathbf{R} = \mathbf{Y}$, active set = {}, basic set = $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^{P-1}$, $m = 0$

  while $m \leq \min(N, P)$ **do**

    $k^* = \underset{\Psi_k(\boldsymbol{\xi}) \in \text{basic set}}{\operatorname{argmax}} \left| Corr(\mathbf{R}, \Psi_k(\boldsymbol{\xi}) \right|$

    Move basis polynomial $\Psi_{k^*}(\boldsymbol{\xi})$ from the basic set to the active set

    Calculate the correction term $\boldsymbol{\Delta}$ [97]

    Update the polynomial coefficients, $\mathbf{a} = \mathbf{a} + \boldsymbol{\Delta}$

    Update the residual, $\mathbf{R} = \mathbf{Y} - \sum_{k=0}^{P-1} \alpha_k \Psi_k(\boldsymbol{\xi})$

    Recalculate the coefficients for the active set with ordinary least squares regression

    Get mean approximation error $\overline{\varepsilon}_m^{(p)}$ via the cross-validation procedure

    $m = m + 1$

  end while

  Store $\overline{\varepsilon}_p^* = \min(\overline{\varepsilon}_1^{(p)}, \ldots, \overline{\varepsilon}_m^{(p)})$ and the corresponding optimal active set

  Stop if either $\overline{\varepsilon}^* = \min(\overline{\varepsilon}_1^*, \ldots, \overline{\varepsilon}_p^*)$ satisfies the target accuracy or increases for the last two iterations $\overline{\varepsilon}_{p-2}^* \leq \overline{\varepsilon}_{p-1}^* \leq \overline{\varepsilon}_p^*$

end for

Estimate the relevant polynomial coefficients for the last optimal active set via ordinary least squares regression

2.4 Chapter Summary

In this chapter, uncertainty propagation and quantification as well as the associated techniques are discussed and compared. Especially, the PEM and PCE, which possess better compromises between their efficiency and accuracy, are presented in more detail.

The PEM, which computes the statistical moments with the deterministic samples, has the best efficiency among all the techniques. Albeit it is defined and its hyper-parameters are determined based on standard Gaussian distribution, the Nataf transformation procedure is proposed to generalize the method to parameter uncertainties with non-Gaussian distributions and mutual correlations. Unlike the PEM, PCE projects the original model to the space that constitutes with orthogonal polynomials and approximates the relations between uncertain parameters and model outputs with the polynomials. In case of independent parameter uncertainties, the statistical moments of model outputs can

be determined with PCE coefficients straightforwardly. Moreover, the PCE model could also be considered as an efficient substitution of the computationally expensive model for approximating the complete probability distribution of model outputs. The PCE could be extended to any types of probability distribution with generalized or customized polynomial basis and be computed efficiently with the LAR algorithm.

Part 1 – Sensitivity Analysis of Pharmaceutical Processes

Chapter 3

Efficient Global Sensitivity Analysis [3]

In this chapter, the global sensitivity analysis (GSA) techniques are introduced for quantitative analysis of the influence of parameter uncertainties on pharmaceutical manufacturing processes and investigate the impact of parameter dependencies.

The chapter is organized as follows. Section 3.1 provides a brief introduction of GSA, motivation and solution strategy of this chapter. Section 3.2 presents the method for describing and sampling independent/correlated random variables. Section 3.3 describes the mathematical formulation and properties of GSA techniques. Section 3.4 presents the computational framework for these sensitivity measures with PCE. In Section 3.5, the application to a continuous synthesis process, which is used for produce API-scaffold, is illustrated and discussed. Finally, the chapter summary is given in Section 3.6.

3.1 Global Sensitivity Analysis

GSA quantifies variations in model output on the entire domain of the parameter space and comprehensively analyzes the interactions among parameters. Different techniques for GSA are available, such as derivative-based methods [98, 99], non-parametric methods [100, 101], variance-based methods [102, 103], and moment-independent methods [62]. A detailed review of those methods is provided by [51] and [55]. Our focus is on GSA techniques with the following features [51]: 1) the technique has a quantitative measure of parameter sensitivities, 2) it is global for the entire parameter space, 3) it is independent of the model structure, and 4) it is available for independent and correlated model parameters which might be of vital importance. In this chapter, two highly promising methods for GSA, which fulfill the features, are introduced and compared: i) Sobol' sensitivity indices (SSI) representing the classical variance-based approach [104],

[3] Part of this chapter has been published in (Xie et al., Reliab. Eng. Syst. Safe., 187, 159–173, 2019 [9])

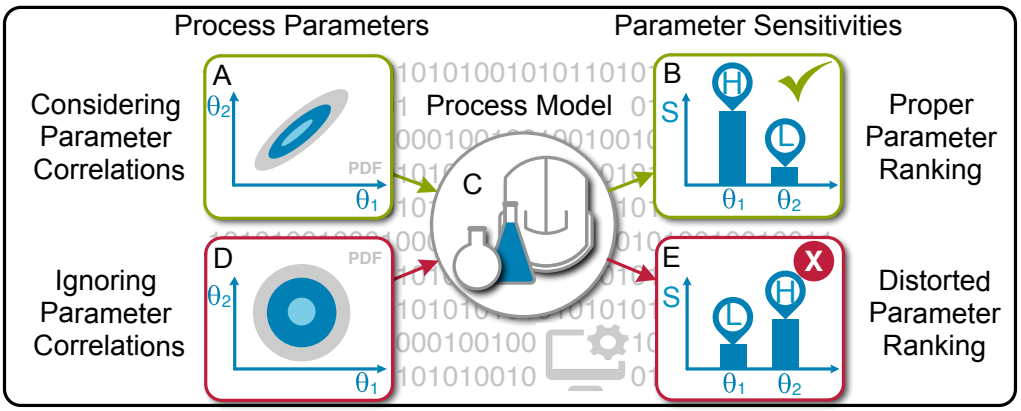


Figure 3.1: Exemplary illustration of the parameter correlation effect on sensitivity analysis: A correlated bivariate probability density function of two process parameters (A) results into a proper parameter sensitivity ranking (B). Parameter θ_1 has a high sensitivity (H), and the sensitivity of θ_2 is low (L). For the same process model (C) but ignoring parameter correlations (D) the sensitivity ranking is misleading (E).

This figure is originated from [9] and created by René Schenkendorf

and ii) the moment-independent sensitivity analysis (MISA) analyzing the entire probability distribution of model outputs [62].

The SSI and MISA are commonly used for sensitivity analysis of problems with independent model parameters [61, 62, 37]. However, problems with correlated model parameters, which arise from inherent parameter dependences or parameter identification procedures, are the standard in chemical processes and other industrial applications [59, 63, 64, 65]. The parameter space is restricted by the parameter dependencies and is different from the one without correlations, which is illustrated in Fig. 3.1. This could lead to a different impact onto the analyzed process model. Thus, the sensitivity analysis and robust design under the hypothesis of independent parameters might lead to unreliable results if the parameter dependencies actually exist. Despite the comprehensive investigation and application with SSI, the method is defined under the assumption of independent parameters [104, 105]. Therefore, explorations on innovative methods upon variance-based sensitivity analysis for systems with correlated random variables were presented in several research works[106, 105, 107, 108, 109]. The covariance decomposition-based sensitivity analysis (CoDSA) proposed by [105] and further explained by [110] is a promising tool for dealing with problems in the presence of pa-

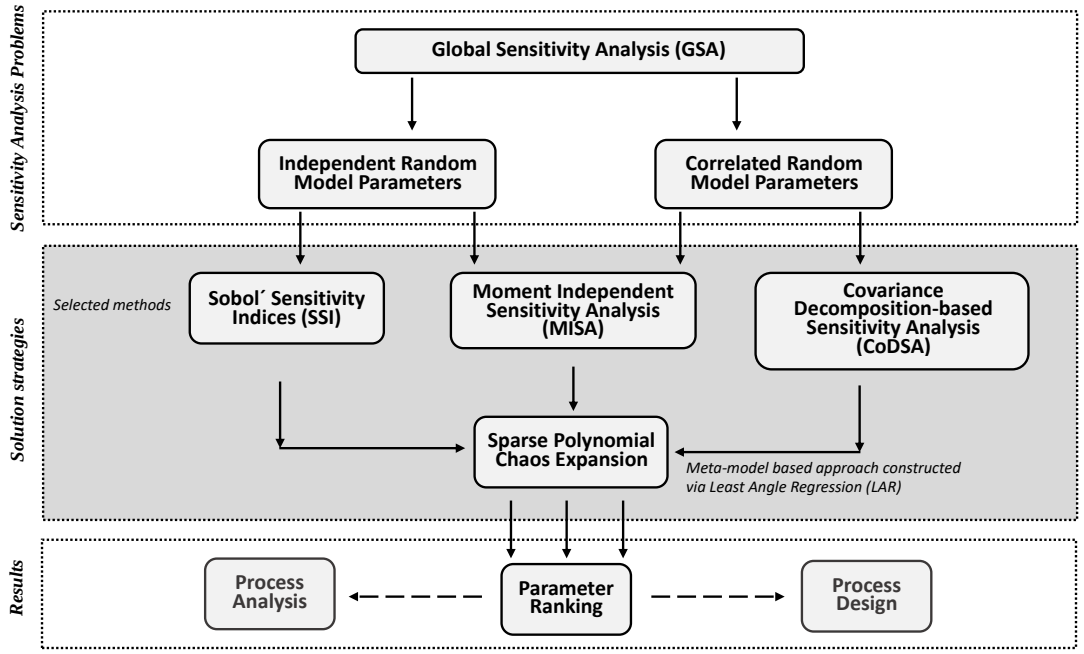


Figure 3.2: Framework for global sensitivity analysis, upper rectangle gives problems, middle rectangle presents the methods to solve the sensitivity analysis problems, lower rectangle presents the results and their potential applications

rameter dependencies. In contrast to the SSI, MISA is not based on the assumption of independent parameters, and thus can be directly extended to problems with correlated parameters [62].

The SSI, CoDSA and MISA are commonly computed by using Monte Carlo simulation [111, 105, 62]. However, Monte Carlo simulations require a large number of deterministic simulations and become computationally expensive especially for large-scale problems. In order to confront the problem, a surrogate model is used to substitute the computationally expensive model. To this end, a polynomial chaos expansion (PCE) surrogate model is implemented [75]. The PCE model is estimated by using LAR (see Chapter 2; [97]). The sensitivity measures are computed through the coefficients of the PCE model or rapid PCE model evaluations [110].

In the current chapter, two questions are answered: 1) *How does the performance of the (co)variance-based SA compare with that of the MISA in analysis of the pharmaceutical processes?* 2) *How do the correlations among parameters impact the parameter sensitivities?* To this end, a deep insight into the different implemented methods for

GSA; i.e., the SSI, CoDSA and MISA approach is firstly provided. Second, the performance of the proposed methods is analyzed with an application to a continuous-flow reactor in the absence and presence of parameter correlations by following the framework shown in Fig. 3.2.

3.2 Independent and Correlated Random Model Parameters

Before starting with the methods for sensitivity analysis, some definitions for multivariate distribution and the difference between independent and correlated model parameters are briefly explained.

The uncertainties are assigned to the model parameters and described with a specific type of probability distribution, for example, Gaussian and uniform distribution. Note that the model parameters could also be design variables, which depends on the type of the problems. The model response Y is represented as the function G:

$$Y = G(\boldsymbol{\theta}) = G(\theta_1, \theta_2, \cdots, \theta_n), \tag{3.1}$$

where $\theta_1, \theta_2, \cdots, \theta_n$ are n random model parameters. The probability of the random parameters on their entire domain are given by an n-dimensional joint probability distribution ($f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$), while the probability of θ_i without reference to the value of the other parameters and with reference to the values of the other parameters is given by the marginal distribution ($f_{\theta_i}(\theta_i)$) and the conditional distribution ($f_{\theta_i|\theta_{\sim i}=\theta_{\sim i}}(\theta_i)$), respectively. Samples for model parameters could be drawn from one of these distributions based on the purpose. For example, a sample group with a constant value for θ_i is generated from the conditional distribution $f_{\theta_{\sim i}|\theta_i=\theta_i}(\theta_{\sim i})$. Note that choosing the right distribution is extremely important because the conditional and marginal distributions are completely different in the case of correlated model parameters.

For independent model parameters, the conditional distribution for θ_i is not affected by the values of other parameters and is equivalent to its marginal distribution. As such, the joint distribution is simplified to the product of the marginal distributions of each parameters:

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = f_{\theta_1}(\theta_1) f_{\theta_2}(\theta 2) \cdots f_{\theta_n}(\theta_n) \tag{3.2}$$

Note that the marginal distributions for each parameters are always given in advance for process analysis purpose. An example of independent parameters is shown in Fig.3.1 (D) which illustrates independent bivariate normal distribution.

For the correlated model parameters, the equivalence between the marginal and conditional distribution and the simplification of the joint distribution in Eq. (3.2) does no longer hold. In order to obtain the joint distribution for correlation model parameters, copula formalism [112] derived from Sklar's theorem provides the link between the joint and marginal distributions for all types of distributions whether they are independent or not [113]. The copula formalism is written as follows [113]:

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = c(F_{\theta_1}(\theta_1), \cdots, F_{\theta_n}(\theta_n)) \cdot \prod_{i=1}^{n} f_{\theta_i}(\theta_i), \tag{3.3}$$

where $F_{\theta_i}(\theta_i)$ is the marginal cumulative density function (CDF) for random parameter θ_i. Here, $c(F_{\theta_1}(\theta_1), \cdots, F_{\theta_n}(\theta_n))$ is the copula density which can be derived by transforming the copula function $C(F_{\theta_1}(\theta_1), \cdots, F_{\theta_n}(\theta_n))$ [113]:

$$c(F_{\theta_1}(\theta_1), \cdots, F_{\theta_n}(\theta_n)) = \frac{\partial^n C(F_{\theta_1}(\theta_1), \cdots, F_{\theta_n}(\theta_n))}{\partial F_{\theta_1}(\theta_1) \cdots \partial F_{\theta_n}(\theta_n)} \tag{3.4}$$

One frequently used copula function is the Gaussian copula which is formulated as [112]:

$$C(F_{\theta_1}(\theta_1), \cdots, F_{\theta_n}(\theta_n); \boldsymbol{\rho}) = F_n(F^{-1}(F_{\theta_1}(\theta_1)), \cdots, F^{-1}(F_{\theta_n}(\theta_n)); \boldsymbol{\rho}) \tag{3.5}$$

Here, $F_n(\cdot)$ is the multivariate cumulative Gaussian distribution function with correlation matrix $\boldsymbol{\rho}$, and $F^{-1}(\cdot)$ is the inverse standard Gaussian distribution function. The correlation matrix consists of Pearson's correlation coefficients (ρ_{ij}), which quantifies the correlation between θ_i and θ_j and is defined as:

$$\rho_{ij} = \frac{cov(\theta_i, \theta_j)}{\sigma_i \sigma_j}, \tag{3.6}$$

where cov is the covariance function, and σ_i is the standard deviation of θ_i. The values of the correlation coefficient ρ are within the range -1 to 1, where 0 and 1(-1) represent independent and completely positive (negative) correlated variables, respectively. The Gaussian copula is available only for linear dependence which is shown in Fig. 3.1 (A). For variables with non-linear correlations, other copulas are required. Additional descriptions of other copulas are beyond the scope of this thesis and refer to [112]. For industrial applications, since only the linear dependence, i.e., correlation matrix, is commonly used for describing the correlations between parameters, the Gaussian copula is a proper choice in this case and as such is also the focus in this work.

Samples of correlated and independent bivariate distributions are illustrated in Fig. 3.1 (A) and (D), respectively. Although the variables have the same marginal distributions, but the shapes of the sample spaces are completely different. The correlation between two parameters restricts the sampling space, which completely changes the information obtained from the sensitivity analysis.

3.3 Methods for Global Sensitivity Analysis in the Absence and Presence of Correlation among the Model Parameters

In this section, the mathematical formulation and definition of the sensitivity measures for the methods used for Global sensitivity analysis, i.e., the SSI, CoDSA and MISA, are briefly reviewed.

3.3.1 Sobol' sensitivity indices for independent model parameters

The Sobol' sensitivity indices (SSI) derived from the decomposition of the model output variance [104] is widely used for global sensitivity analysis of problems with independent model parameters. The following explanation of SSI is based on the definition in [104, 51].

For function $Y = G(\boldsymbol{\theta})$ of finite variance, it can be uniquely decomposed as follows [51]:

$$G(\boldsymbol{\theta}) = G_0 + \sum_{i=1}^{n} G_i(\theta_i) + \sum_{1\le i<j\le n} G_{ij}(\theta_i,\theta_j) + \cdots + G_{12\cdots n}(\theta_1,\cdots,\theta_n), \tag{3.7}$$

where the partial functions are defined as:

$$G_0 = E(Y), \tag{3.8}$$

$$G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}) = E_{\boldsymbol{\theta}_{\sim\mathbf{u}}}(Y|\boldsymbol{\theta}_{\mathbf{u}}) - \sum_{\substack{\mathbf{w}\subset\mathbf{u}\\ \mathbf{w}\neq\emptyset}} G_{\mathbf{w}} - G_0, \tag{3.9}$$

where $E(\cdot)$ denotes the expectation operation. The uniqueness of the function decomposition is ensured based upon the following properties of the partial functions:

$$\int G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}) f_{\theta_i}(\theta_i) d\theta_i = 0 \qquad i \in \mathbf{u}, \tag{3.10}$$

$$\int G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}) G_{\mathbf{v}}(\boldsymbol{\theta}_{\mathbf{v}}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} = 0 \qquad \mathbf{u} \neq \mathbf{v} \tag{3.11}$$

$\mathbf{u}$ and $\mathbf{v}$ are different subsets of the full index set $\{1, \cdots, n\}$. Note that $G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}) = G_0$ if $\mathbf{u}$ is an empty set. $\boldsymbol{\theta}_{\mathbf{u}}$ includes the variables marked with the numbers in subset $\mathbf{u}$.

Based on the orthogonal property of the component function described in Eq. (3.11), the decomposition of the variance of function $G(\boldsymbol{\theta})$ can be deduced from Eq. (3.7) and is written as:

$$Var(Y) = \sum_{i=1}^{n} V_i + \sum_{1 \le i < j \le n} V_{ij} + \cdots + V_{12\ldots n}, \tag{3.12}$$

where

$$V_{\mathbf{u}} = V_{\boldsymbol{\theta}_{\mathbf{u}}}(E_{\boldsymbol{\theta}_{\sim \mathbf{u}}}(Y|\boldsymbol{\theta}_{\mathbf{u}})) - \sum_{\substack{\mathbf{w} \subset \mathbf{u} \\ \mathbf{w} \neq \emptyset}} V_{\mathbf{w}}, \tag{3.13}$$

and $\sim \mathbf{u}$ denoting the complementary subset of $\mathbf{u}$. $Var(Y)$ is the variance of function $G(\boldsymbol{\theta})$. $V_i, V_{ij}, \cdots$, and $V_{123\ldots n}$ are partial variances which describe the effect of individual parameters or parameter interactions on $Var(Y)$.

Based on the normalization of Eq. (3.12), [104] introduced the SSI as follows:

$$S_i^{uc} = \frac{V_i}{Var(Y)}, \tag{3.14}$$

$$S_{ij}^{uc} = \frac{V_{ij}}{Var(Y)}, \tag{3.15}$$

$\ldots$

$$S_{12\ldots,n}^{uc} = \frac{V_{12\ldots n}}{Var(Y)} \tag{3.16}$$

S_i^{uc} are called the first-order sensitivity indices and $S_{ij}^{uc}, \cdots, S_{12\ldots n}^{uc}$ are called interaction sensitivity indices. First-order sensitivity indices give the percentage of the total variance of Y due to the uncertainty in each input variable, whereas the interaction sensitivity indices measure the interaction among input variables. Moreover, the total sensitivity indices $S_{T_i}^{uc}$ are introduced to describe the contribution of single variables θ_i and their interactions with other model parameters, which is defined as [51]:

$$S_{T_i}^{uc} = \frac{V_i + V_{i,j} + \cdots + V_{1\ldots i\ldots n}}{Var(Y)} \tag{3.17}$$

Practically, instead of calculating $2^n - 1$ sensitivity indices, which are the total number of the first-order and interaction terms and increase dramatically with larger n, only the total sensitivity indices and the first-order sensitivity indices (with a number of $2n$) are calculated to save computational cost.

The SSI performs satisfactorily and reflects the model structure for functions with independent model parameters. However, a problem arises in implementing the SSI for functions with correlated model parameters as discussed in [114]. Therefore, other decomposition concept and new sensitivity indices are required for problems with correlated model parameters [105].

3.3.2 Covariance Decomposition-based Sensitivity Analysis (CoDSA) for Correlated Model Parameters

As be described in Sec. 3.2, the entire sampling space for correlated parameters is different from that for independent parameters. This might result in the shift of the mean and the variance, or more generally the entire distribution, of the model response as shown in Fig. 3.3. Therefore, it is necessary to calculate the sensitivities in the presence of parameter correlations.

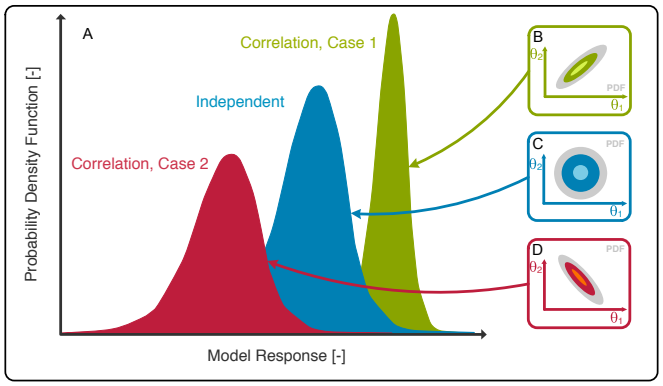


Figure 3.3: Exemplary illustration of probability distributions of the model response (A) with independent (C) or correlated random parameters; i.e., Case 1 (B) and Case 2 (D) are two different situations decided by the structure of the model and dependency among the model parameters.

One idea for dealing with correlated random parameters is using isoprobabilistic transformation concepts, such as the Rosenblatt transformation [115, 65] and the Nataf transformation [89]. By utilizing those transformations, the original correlated variables are converted to new independent variables. For instance, this approach has been applied

in [107] and [108]. Here, some new sensitivity measures are given regarding parameter dependencies. Alternatively, as proposed by [105], new sensitivity indices can be derived, which are based on the covariance decomposition of function $G(\boldsymbol{\theta})$. As there is no need for an isoprobabilistic transformation step, the covariance decomposition approach is of primary interest in this chapter.

3.3.2.1 Covariance Decomposition

Let us consider the second-order function $\mathrm{G}(\boldsymbol{\theta}^{cor})$ again, but with correlated model parameters. As pointed out by [116], the same functional decomposition of $\mathrm{G}(\cdot)$ is still available, as the additional correlations among the model parameters do not affect the structure of the model. Therefore, similar functional relationships between the model parameters and the outputs can be derived as [116]:

$$\begin{aligned} Y = G(\boldsymbol{\theta}^{cor}) =& G_0 + \sum_{i=1}^{n} G_i(\theta_i^{cor}) + \sum_{1\le i<j\le n} G_{ij}(\theta_i^{cor}, \theta_j^{cor}) + \cdots + \\ & G_{123\cdots n}(\theta_1^{cor}, \cdots, \theta_n^{cor}), \end{aligned} \tag{3.18}$$

$$G_0 = E(Y), \tag{3.19}$$

$$G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}) = E_{\boldsymbol{\theta}_{\sim\mathbf{u}}}(Y|\boldsymbol{\theta}_{\mathbf{u}}) - \sum_{\mathbf{w}\subset\mathbf{u}} G_{\mathbf{w}} - \sum_{\substack{\mathbf{u}\not\supseteq\mathbf{w}\subseteq\{1,2,\ldots,n\} \\ \mathbf{u}\cap\mathbf{w}\neq\emptyset}} E_{\boldsymbol{\theta}_{\sim\mathbf{u}}}[G_{\mathbf{w}}], \tag{3.20}$$

where $E_{\boldsymbol{\theta}_{\sim\mathbf{u}}}(\cdot)$ is the operator which calculates the (conditioned) mean value of the output with the marginal distribution of $\boldsymbol{\theta}_{\sim\mathbf{u}}$. After the decomposed functions are constructed, the correlations among the parameters should be considered for the sensitivity analysis. Note that $G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor})$ is not mutually orthogonal for the correlated model parameters, and thus, the variance decomposition of model output Y in Eq. (3.12) is no longer defined. Therefore, the covariance decomposition of $Var(Y)$ was introduced by [105]:

$$\begin{aligned} Var(Y) &= Cov[(G_0 + \sum_{\substack{\mathbf{u}\subset\{1,2,\cdots,n\} \\ \mathbf{u}\neq\emptyset}} G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor})), G(\boldsymbol{\theta}^{cor})], \\ &= \sum_{\substack{\mathbf{u}\subset\{1,2,\cdots,n\} \\ \mathbf{u}\neq\emptyset}} Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G(\boldsymbol{\theta}^{cor})], \end{aligned} \tag{3.21}$$

where $Var(Y)$ is exactly partitioned by $Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G(\boldsymbol{\theta}^{cor})]$ which describes the total contribution of component function $G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor})$ to the total variance of output $Var(Y)$. Further decomposition of $Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G(\boldsymbol{\theta}^{cor})]$ can be used to separate the total contribution into two parts similar to:

$$\begin{aligned} &Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G(\boldsymbol{\theta}^{cor})] \\ &= Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G_0 + \sum_{\substack{\mathbf{u} \subset \{1,2,\cdots,n\} \\ \mathbf{u} \neq \emptyset}} G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor})], \\ &= Var[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor})] + \sum_{\mathbf{v} \neq \mathbf{u}} Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G_{\mathbf{v}}(\boldsymbol{\theta}_{\mathbf{v}}^{cor})] \end{aligned} \tag{3.22}$$

$Var[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor})]$ denotes the contribution of the component function to $Var(Y)$, which depends on the function itself (structure) and the marginal pdf $f_{\boldsymbol{\theta}_{\mathbf{u}}}(\boldsymbol{\theta}_{\mathbf{u}})$ only. $\sum_{\mathbf{v} \neq \mathbf{u}} Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G_{\mathbf{v}}(\boldsymbol{\theta}_{\mathbf{v}}^{cor})]$ denotes the contribution due to the correlation among the model parameters.

Obviously, the covariance decomposition has a completely different structure compared to the variance decomposition for the independent case. The major difference between them is that the component contribution of the covariance decomposition (both covariance parts of Eqs. (3.21) and (3.22)) could be negative, which can never happen in variance-based decomposition. Due to the participation of correlations, new sensitivity indices are required for the sensitivity analysis.

3.3.2.2 Covariance Decomposition-based Sensitivity Indices

[116] introduced three new sensitivity indices for CoDSA, which are obtained by renormalizing Eqs. (3.21) and (3.22) with $Var(Y)$ and formulated as follows:

$$S_u^{cov} = Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G(\boldsymbol{\theta}^{cor})]/Var(Y), \tag{3.23}$$

$$S_u^{U} = Var[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor})]/Var(Y), \tag{3.24}$$

$$S_u^{C} = \sum_{\mathbf{v} \neq \mathbf{u}} Cov[G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^{cor}), G_{\mathbf{v}}(\boldsymbol{\theta}_{\mathbf{v}}^{cor})]/Var(Y) \tag{3.25}$$

The new indices are called total covariance-based sensitivity indices S_u^{cov}, structural sensitivity indices S_u^U, and correlative sensitivity indices S_u^C [116]. The structural and correlative sensitivity indices indicate, in turn, the sensitivity of model structure and correlations among the model parameters on model output Y. S_u^{cov} is the sum of the two sensitivity indices as shown in Eq. (3.26) and presents the total effect on model

output Y:

$$S_u^{cov} = S_u^U + S_u^C \tag{3.26}$$

Note that the sum of S_u^{cov} for all possible subset $\mathbf{u}$ is equal to 1 no matter whether the individual value of S_u^{cov} is positive, negative or larger than 1. The three sensitivity indices could be the sensitivities for individual parameters (first-order) or interactions among two or more variables depending on subset $\mathbf{u}$. The total sensitivity indices can be directly extended for CoDSA similar to Eq. (3.17), and thus, there are the total covariance-based total sensitivity indices $S_{T_i}^{cov}$, structural total sensitivity indices $S_{T_i}^U$, and correlative total sensitivity indices $S_{T_i}^C$. The structure of all the indices used for CoDSA is illustrated in Fig. 3.4, where the structural sensitivity indices are listed on the left hand side and the correlative sensitivity are listed on the right hand side.

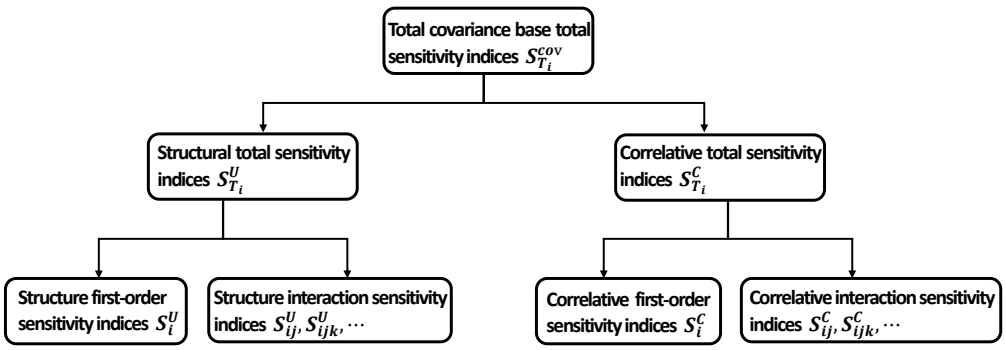


Figure 3.4: Topology of the covariance-based sensitivity indices.

As mentioned previously, the assumption of mutual orthogonality of the component functions in Eq. (3.18) fails for correlated parameters. Thus, the decomposition in Eq. (3.18) is not unique and depends on the applied method, respectively. This may affect the sensitivity results derived from the covariance decomposition. Alternatively, [117] proposed a general and unique decomposition for functions with correlated parameters based on the relaxed vanishing condition, where the component functions are hierarchically orthogonal. Such kind of decomposition is more rigorous but requires sophisticated construction methods. The interested reader is referred to [117, 118] and references therein.

3.3.3 Moment-independent Sensitivity Analysis (MISA)

In the previous subsections, the SSI and CoDSA which are defined based on the (co)variance of the model output are presented. However, using a single statistical moment for SA may lead to an apparent loss of information from other statistical moments as skewness and kurtosis [62]. Therefore, MISA, which is derived directly from the entire distribution of the model output, is introduced and used for sensitivity studies [62]. MISA compares the difference between probability distribution $f_Y(y)$ of output Y and conditional probability distributions $f_{Y|\theta_i}(y)$ of output Y to calculate the parameter sensitivities [62], which is mathematically expressed as:

$$s(\theta_i) = \int |f_Y(y) - f_{Y|\theta_i}(y)| dy \tag{3.27}$$

$s(\theta_i)$ is also called the shift function, and the average of the shift function on the entire distribution of θ_i is then given by:

$$E_{\theta_i}[s(\theta_i)] = \int f_{\theta_i}(\theta_i) \left[\int |f_Y(y) - f_{Y|\theta_i}(y)| dy \right] d\theta_i, \tag{3.28}$$

where $f_{\theta_i}(\theta_i)$ is the marginal density of model parameter θ_i. Based on Eq. (3.28), [62] proposed a new indicator δ_i for global sensitivity analysis, which is defined as follows:

$$\delta_i = \frac{1}{2} E_{\theta_i}[s(\theta_i)] \tag{3.29}$$

The indicator can also be directly extended to a group of parameters equal to:

$$E_{\boldsymbol{\theta_u}}[s(\boldsymbol{\theta_u})] = \int f_{\boldsymbol{\theta_u}}(\boldsymbol{\theta_u}) \left[\int |f_Y(y) - f_{Y|\boldsymbol{\theta_u}}(y)| dy \right] d\boldsymbol{\theta_u}, \tag{3.30}$$

$$\delta_{\mathbf{u}} = \frac{1}{2} E_{\boldsymbol{\theta_u}}[s(\boldsymbol{\theta_u})], \tag{3.31}$$

in which $\mathbf{u}$ is a vector that includes the index of the parameters part of the group, and $f_{\boldsymbol{\theta_u}}(\boldsymbol{\theta_u})$ is the marginal distribution of the parameter group. [62] also concluded 5 properties for the indicator: 1) *The δ_i (or $\delta_{\mathbf{u}}$) varies in the range [0,1]* and 2) *it indicates independence between θ_i (group $\boldsymbol{\theta_u}$) and output Y if $\delta_i (or \delta_{\mathbf{u}}) = 0$.* 3) *The indicator of all model parameters ($\delta_{1,2,\ldots,n}$) equals unity, which is straightforward to understand as the uncertainty of the output vanishes once the uncertainties of all model parameters are*

eliminated. 4) *δ_{ij} for parameters θ_i and θ_j is within the boundary as $\delta_i \leq \delta_{ij} \leq \delta_i + \delta_{i|j}$, in which $\delta_{i|j}$ is the conditional δ.* 5) *The left equal holds if the output is independent of θ_j.* As can be seen, the boundaries for the parameter sensitivity indicator δ is provided within the last two properties. These properties make the sensitivity indicator more representative and comparable. Further descriptions and proofs please refer to [62]. The method is also available for problems with correlated model parameters, as the assumption of independent parameters is not required for its definition.

In this section, three different methods for SA are reviewed. However, the proposed concepts require the calculation of statistical values for the model output, which results in high-order numerical integration problems. Therefore, highly efficient methods for uncertainty quantification are required to lower the computational burden of applying SA to complex chemical processes.

3.4 Estimation of Sensitivity Measures using Polynomial Chaos Expansion

Numerical techniques are the standard in determining parameter sensitivities because analytical solutions are not feasible, especially for highly nonlinear pharmaceutical processes [75, 51]. Methods presented in Chapter 2 can be used for conducting GSA. MCs, its variations and GQ have been used in literatures [51], but their practicability are hindered by their computational burden. PEM, as an efficient solution strategy, has been used to calculate SSI in [27] but is not suitable for conducting MISA. Therefore, PCE is used, as PCE is tailored for GSA [75]. The basics of PCE has been already introduced in Chapter 2. And the computation of all aforementioned GSA techniques with PCE is present in what follows.

3.4.1 Computation of Sensitivity Indices using PCE

Suppose there is a function $Y = G(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the n_θ-dimensional random model parameter vector with given distributions. With the proposed procedures in Chapter 2, the function $Y = G(\boldsymbol{\theta})$ is represented by the PCE truncated at order $\overline{p}_{max}$:

$$Y = G(\boldsymbol{\theta}) \approx \sum_{k=0}^{P-1} \alpha_k \Psi_k(\boldsymbol{\theta}) \tag{3.32}$$

For the sake of simplicity, indices $\mathbf{k} = \{k_1, k_2, \ldots, k_{n_\theta}\}$, where k_i is the order of univariate polynomials for individual input variable i, are defined. Here, the polynomial basis are

assumed to be built upon random variable $\boldsymbol{\theta}$. Otherwise the PCE expression should be a function of random variables used for the construct of polynomial basis, e.g., $\Psi_k(\boldsymbol{\xi})$, as explained in Chapter 2. A is a set of all possible indices $\mathbf{k}$ truncated at order $\overline{p}_{max}$:

$$A = \{\mathbf{k} \in \mathbb{N}^{n_\theta} | k_1 + k_2 + \cdots + k_{n_\xi} \leq \overline{p}_{max}\} \tag{3.33}$$

Therefore, the PCE can be written compactly as:

$$Y = G(\boldsymbol{\theta}) \approx \sum_{\mathbf{k} \in A} \alpha_{\mathbf{k}} \Psi_{\mathbf{k}}(\boldsymbol{\theta}) \tag{3.34}$$

3.4.1.1 Computation of the SSI using PCE

To calculate the SSI, function $G(\boldsymbol{X})$ is decomposed as shown in Eq. (3.7). Due to the orthogonality of the basis polynomials, the component functions $G_{\mathbf{u}}$ in the Sobol' decomposition in Eq. (3.7) are approximated by:

$$G_{\mathbf{u}}(\boldsymbol{\theta}_{\boldsymbol{u}}) \approx \sum_{\mathbf{k} \in A_{\mathbf{u}}} \alpha_{\mathbf{k}} \Psi_{\mathbf{k}}(\boldsymbol{\theta}), \tag{3.35}$$

where $A_{\mathbf{u}} \subset A$ and is defined as:

$$A_{\mathbf{u}} = \{\mathbf{k} \in A | k_i \neq 0 \text{ if and only if i} \in \mathbf{u}\} \tag{3.36}$$

From the definition, it is clear that the PCE in Eq. (3.35) is only a function of the variables included in $\boldsymbol{\theta}_{\boldsymbol{u}}$ because the order of the univariate polynomials of the other variables is zero. The statistical property of function $G(\boldsymbol{\theta})$, especially the mean and the variance, can be directly obtained from the coefficients of their PCE coefficients as:

$$E_{Y,A} = \alpha_0, \tag{3.37}$$

$$V_{Y,A} = \sum_{\mathbf{k} \in \mathbf{A}, \mathbf{k} \neq \mathbf{0}} \alpha_{\mathbf{k}}^2 \tag{3.38}$$

The SSI (Eq. (3.14)) can easily be derived from the representation above:

$$S_{\mathbf{u}}^{uc} = \frac{1}{V_{Y,A}} \sum_{\mathbf{k} \in A_{\mathbf{u}}} \alpha_{\mathbf{k}}^2, \tag{3.39}$$

where $S_{\mathbf{u}}^{uc}$ could be 1) a first-order sensitivity if $\mathbf{u}$ contains only one element or 2) an interaction sensitivity if it contains more than one. The total sensitivity indices are then

given by:

$$S_{T_i}^{uc} = \frac{1}{V_{Y,A}} \sum_{\mathbf{k} \in \overline{A}_i} \alpha_{\mathbf{k}}^2, \tag{3.40}$$

where $\overline{A}_i \subset A$ and includes all the basis polynomials related to variable ξ_i:

$$\overline{A}_i = \{\mathbf{k} \in A | k_i \neq 0\} \tag{3.41}$$

3.4.1.2 Computation of CoDSA using PCE

The construction of the sensitivity indices for function $Y = G(\boldsymbol{\theta})$ with correlated inputs is presented in Sec. 3.3.2.2, where three new sensitivity indices S_u^{cov}, S_u^U and S_u^C are introduced. As be explained in Sec. 3.3.2.2, a unique decomposition where the component functions are hierarchically orthogonal is not easy to construct. Therefore, the function decomposition approach, i.e., the PCE model, to compute the sensitivity indices, is used. The exclusive use of PCE models also makes the computation more comparable for the three GSA techniques. The PCE approximation, as shown in Eqs. (3.34) and (3.35), is constructed as if the random variables are independent. However, the statistical moments in covariance decomposition cannot be derived directly from the coefficients of PCE because the orthogonality between the basis polynomials of PCE does not exist for the correlated variables. For this reason, Monte Carlo simulation are used to estimate the mean, variance and covariance of function $G(\boldsymbol{\theta})$ and component function $G_{\mathbf{u}}(\boldsymbol{\theta}_{\boldsymbol{u}})$ with samples drawn from correlated distributions:

$$E(Y) = \frac{1}{N} \sum_{i=1}^{N} \sum_{\mathbf{k} \in A} \alpha_{\mathbf{k}} \Psi_{\mathbf{k}}(\boldsymbol{\theta}^i), \tag{3.42}$$

$$Var(Y) = \frac{1}{N-1} \sum_{i=1}^{N} (\sum_{\mathbf{k} \in A} \alpha_{\mathbf{k}} \Psi_{\mathbf{k}}(\boldsymbol{\theta}^i) - E(Y))^2, \tag{3.43}$$

$$E(G_{\mathbf{u}}(\boldsymbol{\theta}_{\boldsymbol{u}})) = \frac{1}{N} \sum_{i=1}^{N} \sum_{\mathbf{k} \in A_{\mathbf{u}}} \alpha_{\mathbf{k}} \Psi_{\mathbf{k}}(\boldsymbol{\theta}_{\boldsymbol{u}}^i) \tag{3.44}$$

Here, $\boldsymbol{\theta}^i$ is one sample vector for the random model parameter vector, and n samples are drawn from the given distribution. Eqs. (3.45) and (3.46) represent an evaluation of the functions approximated by PCE with the sample $\boldsymbol{\theta}^i$:

$$G(\boldsymbol{\theta}^i) = \sum_{\mathbf{k} \in A} \alpha_{\mathbf{k}} \Psi_{\mathbf{k}}(\boldsymbol{\theta}^i), \tag{3.45}$$

$$G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^i) = \sum_{\mathbf{k}\in A_{\mathbf{u}}} \alpha_{\mathbf{k}}\Psi_{\mathbf{k}}(\boldsymbol{\theta}^i) \tag{3.46}$$

The covariance-based sensitivity indices computed from the Monte Carlo estimation are formulated as:

$$S_u^{cov} = \frac{\sum_{i=1}^N (G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^i) - E(G_{\mathbf{u}}(\boldsymbol{\theta}_{\boldsymbol{u}})))(G(\boldsymbol{\theta}^i) - E(Y))}{(N-1)Var(Y)}, \tag{3.47}$$

$$S_u^{U} = \frac{\sum_{i=1}^N (G_{\mathbf{u}}(\boldsymbol{\theta}_{\mathbf{u}}^i) - E(G_{\mathbf{u}}(\boldsymbol{\theta}_{\boldsymbol{u}})))^2}{(N-1)Var(Y)}, \tag{3.48}$$

$$S_u^{C} = S_u^{cov} - S_u^{U} \tag{3.49}$$

We should note that the computational cost is low as all the evaluations are conducted on the PCE model. The computational effort for the Monte Carlo simulation on the PCE model is negligible compared to that for estimation of the coefficients for the PCE model.

3.4.2 Computation of MISA using PCE

According to the definition of MISA in Sec. 3.3.3, the density functions $f_Y(y)$ and $f_{Y|\theta_i}(y)$ for the model output are required for calculating the shift function and the indicator, and therefore, the kernel density estimator (KDE) is used to estimate the distributions [119]. The KDE is a non-parametric method for estimating the probability density function of random variables with an arbitrary probability distribution [120] according to:

$$\hat{f}(h) = \frac{1}{n\Delta}\sum_{i=1}^{n} K(\frac{h-h_i}{\Delta}), \tag{3.50}$$

where n is the number of samples, K is the kernel function, and Δ is the bandwidth. K is normally a non-negative function, e.g., a standard normal density function. h_i is the constraint evaluation at sample i. Moreover, the double-loop Monte Carlo method illustrated in Fig. 3.5 is used to calculate the indicator δ [121].

The double-loop method starts with two groups of samples generated from the distribution of the random model parameters. The unconditioned density function $f_Y(y)$ is estimated from the model evaluations associated with the samples in A. For loop 2, the conditioned density function $f_{Y|\theta_i}(y)$ is estimated by n_i samples, and the shift function $s(\theta_i)$ is calculated at a given value θ_i^j. Note that the samples are different for the independent and correlated model parameters. For the independent case, the samples are

generated directly by replacing the i-th column of sample B with the (i,j)-th element θ_{ij} in sample A. However, this does not work for the correlated case as the conditioned density $f_{\theta_{\sim i}|\theta_i=\theta_i^j}(\theta_{\sim i})$ is different for different values of θ_i. Therefore, the samples should be updated for each iteration according to the conditioned density with parameter θ_i specified to the value θ_i^j in sample A. In loop 1, loop 2 is repeated for different values of θ_i from sample A, and the indicator δ_i is calculated with the values for the shift function. Note that the calculation formula for shift function $s(\theta_i)$ and δ_i is the Monte Carlo estimation of Eqs. (3.27) and (3.29). The total number (n_t) of model evaluations for the double-loop Monte Carlo method is $(n \times n_i + 1) \times n_o$, where n, n_o and n_i are the number of model parameters, the size of samples A and B, respectively. The number n_t could be prohibitively high as the KDE requires a large sample size to ensure the accuracy of the estimated density function. Thus, the PCE model derived above is used for the model evaluations to ensure low computational costs.

3.5 Case Study: a Continuous Synthesis of an API–Scaffold

3.5.1 Problem Statement

In this study, a model of a continuous-flow reactor processing the synthesis of aminopyrimidine as an API–scaffold [63] is considered. The mechanism of the reactions is described as follows:

$$\mathrm{A} + \mathrm{B} \xrightarrow{\mathrm{r_1}} \mathrm{C} \tag{3.51}$$

$$\mathrm{A} + \mathrm{B} \xrightarrow{\mathrm{r_2}} \mathrm{D} \tag{3.52}$$

$$\mathrm{C} + \mathrm{B} \xrightarrow{\mathrm{r_3}} \mathrm{E} \tag{3.53}$$

$$\mathrm{D} + \mathrm{B} \xrightarrow{\mathrm{r_4}} \mathrm{E} \tag{3.54}$$

Eq. (3.51) to Eq. (3.54) describe the nucleophilic aromatic substitution reactions ($\mathrm{S_NAr}$) of 2,4-dichloropyrimidine (A) and morpholine (B) in ethanol which produce the desired product 2-substituted aminopyrimidine (D), the less-desired product 4-substituted aminopyrimidine (C) and side product 2,4-substituted aminopyrimidine(E) [63]. A complete description of the reactions and their potential application in the pharmaceutical field can be found in [63]. The governing equations of the continuous-flow reactor are formulated as follows:

$$\frac{dC_A}{dt} = -k_1 C_A C_B - k_2 C_A C_B \tag{3.55}$$

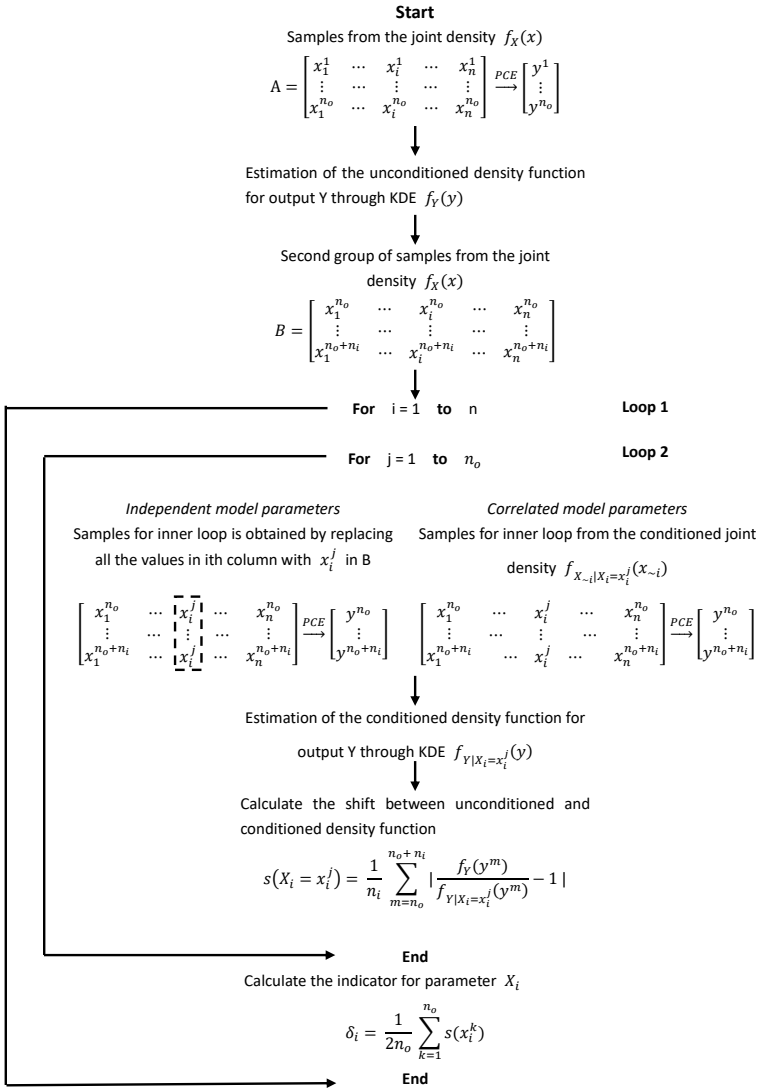


Figure 3.5: The framework of the double-loop method for the calculation of indicator δ [121].

$$\frac{dC_B}{dt} = -k_1 C_A C_B - k_2 C_A C_B - k_3 C_B C_C - k_4 C_B C_D \tag{3.56}$$
$$\frac{dC_C}{dt} = k_1 C_A C_B - k_3 C_B C_C \tag{3.57}$$
$$\frac{dC_D}{dt} = k_2 C_A C_B - k_4 C_B C_D \tag{3.58}$$
$$\frac{dC_E}{dt} = k_3 C_B C_C + k_4 C_B C_D, \tag{3.59}$$

where

$$k_i = A_i exp(-\frac{E_{Ai}}{RT}), \quad i \in \{1, 2, 3, 4\} \tag{3.60}$$

in which C_j with $j \in \{A, B, C, D, E\}$ are the concentrations, k_i are the kinetic constants, A_i and E_{Ai} are pre-exponential factors and activation energies, and r_i are the reaction rates determined by the corresponding kinetic constant and concentration of reactants. According to [63] an isothermal reactor is assumed, i.e., the temperature T is constant along the reactor. Residence time t is defined as the position in the tubular reactor divided by the flow rate of the substance based on the assumption that plug flow is inside the reactor. The values for the initial conditions and parameters are listed in Table 3.1, where t_{end} is the final residence time which is decided by the volume of the reactor and volumetric flow rate of the components inside the reactor. The kinetic parameters A_i and E_{Ai} estimated from the experiments are not accurate and their uncertainties are characterized by normal distributions as shown in Table 3.1. For the first part of the sensitivity analysis, the eight parameters are assumed to be independent. However, the data provided by [63] reveal strong correlations among the parameters measured by the correlation matrix in Table 3.2. This correlation matrix is used in the second part of the sensitivity analysis. The model output that is of interest here is the final concentration of product D (C_{Df}). Therefore, the surrogate model is constructed for the eight kinetic parameters and C_{Df} before starting the sensitivity analysis. The PCE model is constructed and estimated in UQLAB© [122], software for UQ and surrogate modeling. Sensitivity analysis performed later are conducted in MATLAB®, and the function *ode15s* is used to solve the dynamic equations of the reactor.

3.5.2 Construction of the PCE Model

The PCE model for C_{Df} (model output) and eight kinetic parameters (inputs) is constructed based on the probability distribution assigned to the parameters. Table 3.3 lists the information for the PCE model. According to Table 3.3, the coefficients of the

Table 3.1: Parameters and uncertainties for the continuous-flow reactor [63].

Parameters		Nominal Value	Uncertainty
C_{A0}(M)		0.150	–
C_{B0}(M)		0.375	–
C_{C0}(M)		0	–
C_{D0}(M)		0	–
C_{E0}(M)		0	–
T(K)		373.15	–
t_{end}(s)		1200	–
R(J/mol·k)		8.314	–
$\log_{10}(A_1)(\mathrm{M^{-1}s^{-1}})$	(θ_1)	3.4	$\mathcal{N}(3.4\ ,\ 0.1)$
E_{A1}(kJ/mol)	(θ_2)	27.0	$\mathcal{N}(27.0\ ,\ 0.6)$
$\log_{10}(A_2)(\mathrm{M^{-1}s^{-1}})$	(θ_3)	3.5	$\mathcal{N}(3.5\ ,\ 0.1)$
E_{A2}(kJ/mol)	(θ_4)	32.1	$\mathcal{N}(32.1\ ,\ 0.6)$
$\log_{10}(A_3)(\mathrm{M^{-1}s^{-1}})$	(θ_5)	4.9	$\mathcal{N}(4.9\ ,\ 0.2)$
E_{A3}(kJ/mol)	(θ_6)	60.0	$\mathcal{N}(60.0\ ,\ 1.6)$
$\log_{10}(A_4)(\mathrm{M^{-1}s^{-1}})$	(θ_7)	3.0	$\mathcal{N}(3.0\ ,\ 0.2)$
E_{A4}(kJ/mol)	(θ_8)	45.0	$\mathcal{N}(45.0\ ,\ 1.7)$

Table 3.2: Correlation coefficients for the eight parameters from [63].

	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8
θ_1	1.000	0.997	0.976	0.968	-0.002	-0.003	0.000	0.000
θ_2	0.997	1.000	0.976	0.973	-0.003	-0.003	0.000	0.000
θ_3	0.976	0.976	1.000	0.997	-0.006	-0.006	0.000	0.000
θ_4	0.968	0.973	0.997	1.000	-0.007	-0.007	0.000	0.000
θ_5	-0.002	-0.003	-0.006	-0.007	1.000	1.000	-0.008	-0.008
θ_6	-0.003	-0.003	-0.006	-0.007	1.000	1.000	-0.008	-0.008
θ_7	0.000	0.000	0.000	0.000	-0.008	-0.008	1.000	1.000
θ_8	0.000	0.000	0.000	0.000	-0.008	-0.008	1.000	1.000

Table 3.3: PCE model settings and characteristics.

Number of random inputs	8
Polynomial basis	Hermite
Maximum order of polynomials	3
Number of model evaluations	200
Estimation error	0.001
Sparsity	40%

PCE model are estimated with 200 random evaluations of the original model by using Algorithm 2.2, and only 40% of the full basis which have maximum order of 3 are activated here. Please note that the maximum order of the polynomials is determined by the desired estimation error and the complexity of the reactor model. The accuracy of the PCE model is indicated by the estimation error in Table 3.3, and further analyzed by comparing with the results from direct Monte Carlo simulations in Fig. 3.6. The probability density function estimated by the PCE model with 200 model evaluations is as good as the one from Monte Carlo simulations with 10,000 model evaluations but much better than the one from Monte Carlo simulations with 200 model evaluations; i.e., using the original process model given in Eqs. (3.55) to (3.59). Based on the PCE model, different sensitivity measures can be then calculate in the absence and presence of correlations as in the following.

3.5.3 Sensitivity Measures in the Absence of Correlation

This section presents the results for the first part of the sensitivity analysis where the correlations among the parameters are neglected. The SSI and MISA for the independent parameters on C_{Df} are calculated by using the sparse PCE model.

The SSI is obtained directly from the coefficients of the preceding PCE model by using Eqs. (3.39) and (3.40). The first-order and total sensitivities of the eight kinetic parameters for the final product concentration C_{Df} are summarized in Table 3.4. Evident differences exist among the magnitude of the sensitivities for different parameters. The first four parameters, $\theta_1, \ldots, \theta_4$, have the strongest impact on the final product concentration and its variance $\mathrm{Var}(C_{Df})$, while the parameters $\theta_5, \ldots, \theta_8$ are less relevant. Moreover, the small deviation between the first-order, S_i^{uc}, and total sensitivities, $S_{T_i}^{uc}$, indicate that the interaction among the parameters is low, i.e, the sum of first-order sensitivities is close to 1.

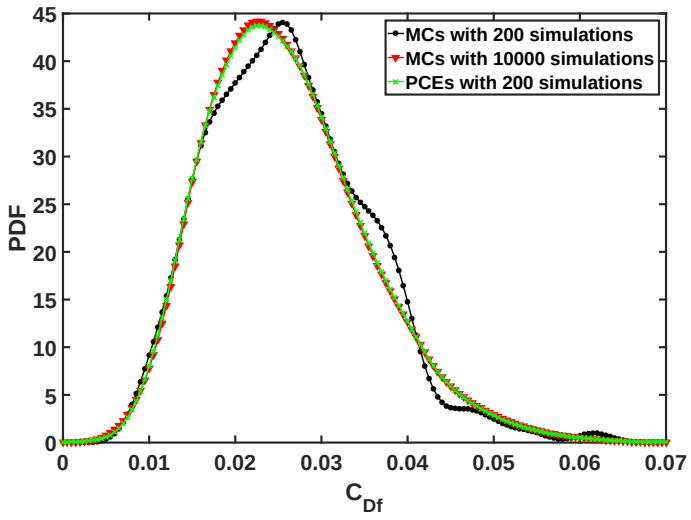
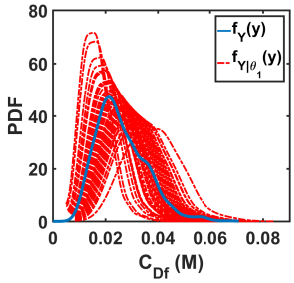


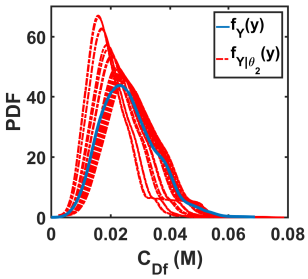
Figure 3.6: Probability density function of component D (C_{Df}) evaluated with 200 and 10,000 Monte Carlo simulations and with PCE from 200 simulations. KDE is used to estimate the probability density functions with the model evaluations.

Unlike the SSI calculated directly from the coefficients of the PCE model, MISA is computed by the method described in Sec. 3.4.2. The samples in the outer loop (n_o) and the inner loop (n_i) are set to 1000. Thus, a total number of 8×10^6 evaluations of the PCE model are required for calculating the indicators. Fig. 3.7 shows the comparison between the unconditioned and conditioned distributions of C_{Df}, where the effect of eliminating the uncertainty of one parameter can be directly observed in the corresponding sub-figures. The quantitative measures for independent parameters, i.e., indicator $\delta_i^{uc} (i = 1 \dots 8)$, are illustrated in Fig. 3.8(a).

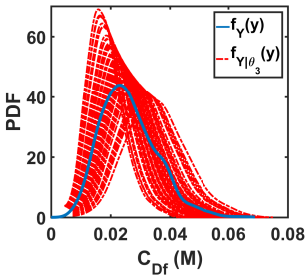
The sensitivity measures from the SSI and MISA reveal the influence of parameter uncertainties on the variation of C_{Df}. According to the results, it can be observed that the trends of the measures from the SSI and MISA are analogous. The kinetic parameters of reactions 1 and 2, i.e., θ_1, θ_2, θ_3 and θ_4, have higher influence than the others. This makes sense from a physical point of view as reactions 1 and 2 are faster than the other reactions and have a direct or indirect relation with product D. it can also be observed that θ_1 and θ_3 have similar importance, which is also true for θ_2 and θ_4. The reason is that reactions 1 and 2 are parallel and have similar competitiveness. In contrast, the kinetic parameters of reactions 3 and 4, i.e.,θ_5, θ_6, θ_7 and θ_8, have lower



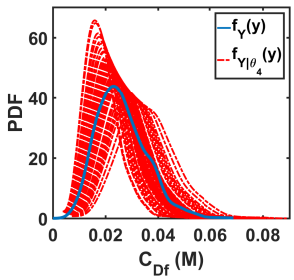
(a) Density distributions considering θ_1



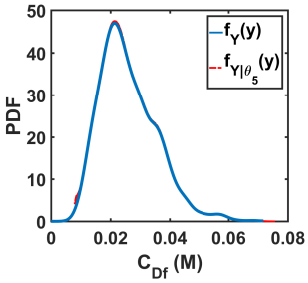
(b) Density distributions considering θ_2



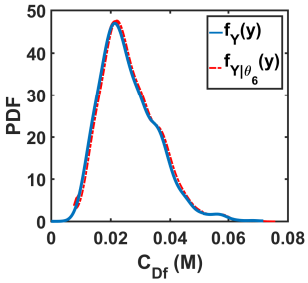
(c) Density distributions considering θ_3



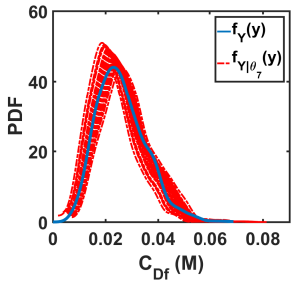
(d) Density distributions considering θ_4



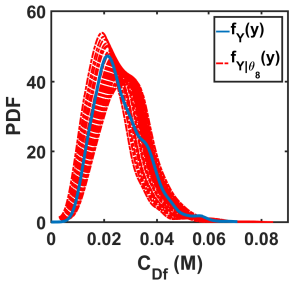
(e) Density distributions considering θ_5



(f) Density distributions considering θ_6



(g) Density distributions considering θ_7



(h) Density distributions considering θ_8

Figure 3.7: Comparison of unconditioned distributions (blue line) and conditioned distributions (red lines) of concentrations of component D representing different parameter realizations in the absence of correlations.

Table 3.4: First-order and total Sobol' sensitivity indices of the kinetic parameters for the final product concentration C_D.

Parameters	S_i^{uc}	$S_{T_i}^{uc}$
θ_1	0.2547	0.2684
θ_2	0.1790	0.1896
θ_3	0.2535	0.2673
θ_4	0.1786	0.1893
θ_5	2.3011×10^{-5}	4.3189×10^{-5}
θ_6	3.7996×10^{-5}	7.7143×10^{-5}
θ_7	0.0400	0.0514
θ_8	0.0577	0.0710
Sum	0.9637	–
$\mathrm{Var}(C_{Df})$	8.5989×10^{-5}	

influence, because reactions 3 and 4 are comparatively slow according to the values of the associated kinetic parameters and concentration of corresponding reactants. However, θ_7 and θ_8 are more important than θ_5 and θ_6 because they have a direct impact on product D. As can be seen, results from both methods are consistent with the structure of the model when the correlations among the parameters are neglected. In the following, the effect of the correlations on the results of the sensitivity analysis is presented.

3.5.4 Sensitivity Measures in the Presence of Correlations

We investigate the effect of parameter correlations given in Table 3.2 on the results of the sensitivity analysis. According to the correlation matrix, the parameters are divided into three subgroups: 1) $\theta_1, \theta_2, \theta_3$ and θ_4, 2) θ_5 and θ_6, and 3) θ_7 and θ_8. The parameters from the same subgroup have a strong correlation, while the parameters from different subgroups have a weak or even no correlation. The samples used in the following calculations are generated with Gaussian copula as Eq.(3.3).

Due to the presence of correlations among the parameters, the SSI is not well defined and, therefore, cannot be used in this situation. CoDSA is available to observe the effect of the parameters, as well as their correlations. The method presented in Sec. 3.4.1.2 is used to compute CoDSA with a sample size of 10,000. The calculated sensitivity indices are listed in Table 3.5, where S_i^U,S_i^C and S_i^{cov} are the first-order sensitivity indices and $S_{T_i}^U$,$S_{T_i}^C$ and $S_{T_i}^{cov}$ are the total sensitivity indices for the corresponding parameters, respectively. When comparing the first-order and total sensitivity indices, it can be

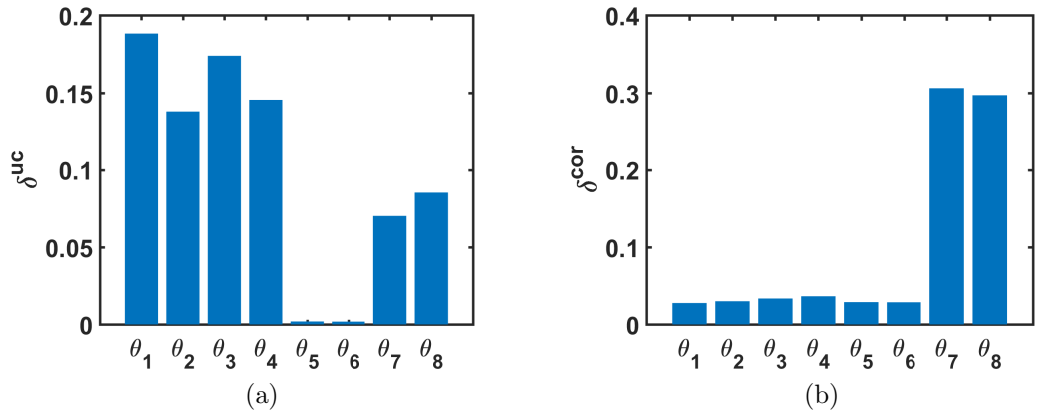


Figure 3.8: MISA sensitivity values of moment-independent measure δ^{uc} for the eight kinetic parameters when ignoring parameter correlations (a), and when parameter correlations are considered properly (b).

observed that the interaction term is not relevant to describe the parameter influence on the model output. Moreover, S_i^U has the same trend as the SSI result but with different magnitudes. The large magnitudes of S_i^U mean the parameter uncertainties have a stronger influence on the model output if they are independent. This can also be observed from S_i^C which represents the effect of correlations and is negative in this case. It turns out the correlations reduce the importance of the parameters and the variance of the model output. The total impact of the parameters on the model output is indicated by the total covariance-based sensitivity indices S_i^{cov}. Please note that the existence of negative values for S_i^{cov} is due to the covariance function formulation and the importance of the parameter is quantified by the absolute value of S_i^{cov}.

In contrast to the variance-based SA methods, MISA is well posed in the presence of correlations among parameters, as its formulation is not based on the assumption of independent parameters. A similar structure for computation as for the independent case is used. Here, however, the correlation matrix of the joint density distribution of the parameters is added. In Fig. 3.8(b), it shows that the indicators δ_i^{cor} which are obtained for the given parameter correlations. Here, the most sensible parameters for C_{Df} are the kinetic parameters θ_7 and θ_8 for reaction 4, which is different from the case with independent parameters. A detailed analysis of the related probability distributions, see Fig. 3.9, explains the new parameter ranking. Here, in Figs. 3.9(g) and 3.9(h), the probability distribution of C_{Df} shifts dramatically if θ_7 and θ_8 are given. θ_1, θ_2, θ_3

Table 3.5: Covariance-based sensitivity indices estimated with sparse PCE.

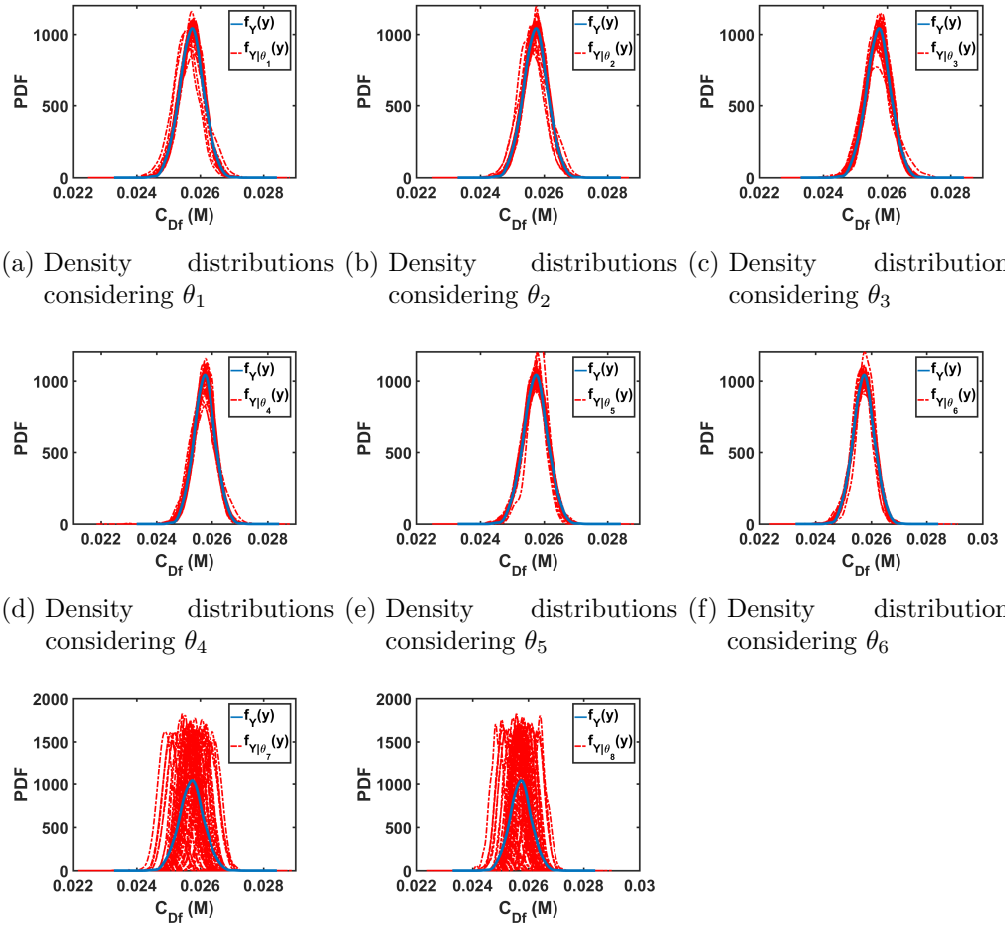
Parameters	S_i^U	S_i^C	S_i^{cov}	$S_{T_i}^U$	$S_{T_i}^C$	$S_{T_i}^{cov}$
θ_1	131.80	-132.98	-1.18	136.23	-137.47	-1.24
θ_2	92.72	-91.55	1.17	99.35	-98.15	1.19
θ_3	131.91	-129.88	2.03	141.63	-139.52	2.11
θ_4	93.09	-94.76	-1.67	96.30	-97.96	-1.66
θ_5	0.01	-0.01	-0.00	0.03	-0.03	0.00
θ_6	0.02	-0.02	0.00	0.02	-0.02	0.00
θ_7	20.68	-24.16	-3.48	20.73	-24.23	-3.50
θ_8	29.87	-25.57	4.30	26.27	-22.26	4.01
Sum	–	–	1.17	–	–	–
Var(Y)			1.65×10^{-7}			

and θ_4 still have a non-ignorable impact on the model response but are impaired by the correlation between the parameters. θ_5 and θ_6 have the weakest impact.

In this case study, it can be concluded that the outcome of the parameter importance ranking is severely affected by the sensitivity measure used; i.e., when applying the (co)variance-based or the moment-independent approach the parameter sensitivities are qualitatively similar. In contrast, the consideration of parameter correlations is crucial for the parameter importance ranking. Thus, the impact of parameter correlations are further discussed in the next section.

3.5.5 Comparison of the Results in the Absence and Presence of Correlations

To compare the resulting variation of C_{Df} in the absence and presence of correlations, the corresponding probability distributions are illustrated in Fig. 3.10. To further demonstrate the effect of parameter correlation which are different from one and less dominating, two additional scenarios are shown in Fig. 3.10. As expected, the correlation has a considerable impact on the resulting distributions. Here, the spread of the shown distribution increases for lower correlation values but is still different compared to the nominal case, i.e., assuming no correlation. It turns out that the uncertainty of the model output, C_{Df}, which is estimated under the assumption of independent parameters for this study is magnified and, therefore, may render a model-based robust design strategy too conservative.



(a) Density distributions considering θ_1 (b) Density distributions considering θ_2 (c) Density distributions considering θ_3

(d) Density distributions considering θ_4 (e) Density distributions considering θ_5 (f) Density distributions considering θ_6

(g) Density distributions considering θ_7 (h) Density distributions considering θ_8

Figure 3.9: Comparison of unconditioned distributions (blue line) and conditioned distributions (red lines) of concentrations of component D considering different parameters in the presence of correlations.

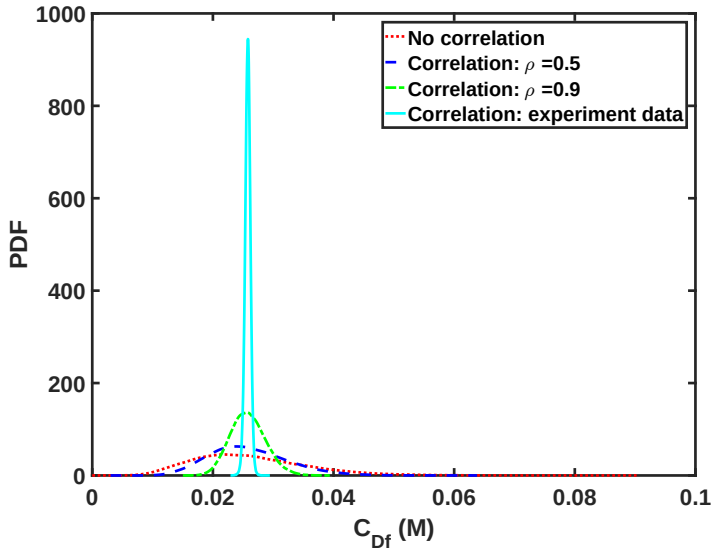


Figure 3.10: Comparison of the resulting probability density functions of C_{Df} in the absence and presence of parameter correlations. The correlation coefficients for $\theta_1, \theta_2, \theta_3, \theta_4, \theta_7, \theta_8$ are assigned with 0.5 and 0.9 assuming fictious, equal correlations for two cases. The result based on a correlation matrix derived from experimental data is illustrated in addition.

In what follows, the importance ranking for the different sensitivity measures neglecting parameter correlation, as well as including parameter correlations based on experimental data are compared; see Table 3.6. For the sake of completeness, results for fictitious correlation coefficients can be found in the Appendix A. The ranking from the first (third) and second (fourth) rows are transposed, while the rankings from first (second) and third (fourth) rows are analogous. To get a more quantitative comparison of the ranking, the Savage score correlation coefficient (SSCC, [123]), which provides a top-down ranking comparison, is calculated for comparing the four sensitivity measures, see Table 3.7. The SSCC has a value range from -1 to 1, where 1 and -1 indicate identical and transposed rankings, respectively. More information regarding the calculation of SSCC is referred to [123]. As it can be seen from Table 3.7, the values of $SSCC(S^{uc}, \delta^{uc})$ and $SSCC(S^{cov}, \delta^{cor})$ are high and close to 1. However, the values of $SSCC(S^{uc}, S^{cov})$ and $SSCC(\delta^{uc}, \delta^{cor})$ are low and negative as the most relevant variables for them are different. According to this, it can be seen that the discrepancy in the ranking of the

Table 3.6: Parameter ranking from (co)variance-based and moment-independent sensitivity analysis in the presence and absence of correlations.

	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8
S_i^{uc}	1	3	2	4	8	7	6	5
S_i^{cor}	5	6	3	4	8	7	2	1
δ_i^{uc}	1	4	2	3	8	7	6	5
δ_i^{cor}	6	5	4	3	8	7	1	2

Table 3.7: Savage score correlation coefficients on importance ranking from variance-based and moment-independent sensitivity in the absence and presence of correlations.

SSCC	S^{uc}	S^{cov}	δ^{uc}	δ^{cor}
S^{uc}	1			
S^{cov}	-0.06	1		
δ^{uc}	0.98	-0.03	1	
δ^{cor}	-0.21	0.78	-0.17	1

most relevant parameters emerges due to the existence of parameter correlations and is less affected by the particular method used for the global sensitivity analysis.

3.6 Chapter Summary

In this chapter, different methods for global sensitivity analysis in the absence and presence of parameter dependencies are presented and compared critically for the continuous synthesis of an active pharmaceutical ingredient scaffold. Sparse polynomial chaos expansion (PCE) was introduced for calculating these sensitivity measures efficiently. Gaussian copulas were utilized to sample from joint and conditional distributions, representing independent and correlated model parameters. In the case study, a continuous-flow reactor model was implemented and analyzed. PCE surrogate model was generated for this reactor model. Here, it can be observed that the presented least angle regression (LAR) algorithm improves the efficiency in PCE modeling. It can also be observed that the PCE model can approximate relevant statistics of simulation results at low computational cost. After performing the sensitivity analysis on the PCE model, the obtained results were compared between (co)variance-based methods (SSI and CoDSA) and the moment-independent method (MISA). Similar parameter sensitiv-

ity ranking for the analyzed metrics are observed, which indicates consistent sensitivity analysis results between these two methods for the situation that parameter uncertainties are Gaussian distributed. Moreover, by comparing the results for independent and correlated parameters, the significant differences between parameter sensitivity ranking were also observed. For independent parameters, the kinetics of reactions 1 and 2 influenced the simulation outcome the most. For the correlated parameters, however, their impact was reduced significantly, and the kinetics of reaction 4 dominated. Moreover, it is noticed that the variance and the width of the distribution of the model output were decreased once parameter correlations were considered.

In summary, global sensitivities provide useful information for analysis and design in the field of pharmaceutical manufacturing and can be derived at acceptable computational cost even for complex problems when using PCE in combination with the LAR algorithm. MISA might be preferable because it is available for problems with independent and correlated model parameters. In addition, MISA is also more rigorous and precise than the (co)variance-based method as MISA considers the entire distribution instead of a single moment of the model output. However, independent of the metrics used for global parameter sensitivity, parameter correlations should always be considered, i.e., utilizing the full information of the parameter (co)variance matrix for process analysis and design. In the next part of the thesis, robust design of pharmaceutical processes under uncertainty reflecting the importance of parameter correlation is also presented.

Part 2 – Robust Design of Pharmaceutical Processes

Chapter 4

Comprehensive and Efficient Framework for Robust Process Design [4]

This chapter introduces a holistic and efficient framework for probability-based robust process design. The PEM introduced in Chapter 2 is used to quantify the statistical moments of model outputs needed in the robust objective function and constraints. Robust formulations of both inequality and equality constraints are investigated. The significance of parameter uncertainties is quantified with the GSA technique introduced in Chapter 3 and used for reducing the computational burden of the robust design. The influence of parameter dependencies, which was shown to be significant on the results of sensitivity analysis, is also investigated for robust process design.

This chapter is organized as follows. Section 4.1 contains the general formulation of the probability-based robust optimization problem. Section 4.2 provides details about robust inequality and equality constraints and approximation methods. The final structure of the proposed framework for robust optimization is given in Section 4.3. To demonstrate the performance of the proposed framework, two case studies are thoroughly discussed in Section 4.4: including a classic jacket tubular reactor and a fed-batch bioreactor for penicillin fermentation. Chapter summary can be found in Section 4.5.

4.1 General Problem Formulation

Model-based design is frequently applied to optimize the performance while satisfying relevant constraints of production processes [22, 27]. However, external disturbances and parameter uncertainties might affect the performance of the processes, which then would deviate from the expected and simulated process characteristics or even result in operation failures [70]. The reliability of the designed processes under various conditions

[4] Part of this chapter has been published in (Xie et al., Processes, 6(10), 183, 2018 [2])

and disturbances is called robustness. Optimization problems that account for process performance and robustness is named as robust process design. Probability-based robust process design, as has been pointed out already in Chapter 1, outperforms the scenario-based approach and could provide solutions with adjustable conservativeness.

The general problem formulation of the probability-based robust process design is given as follow:

Problem 1. Probability-based robust optimization (RO) problem

$$\min_{\mathbf{x}(\cdot),\mathbf{u}(\cdot)} \; E[M(\mathbf{x}_{t_f})] + \alpha Var[M(\mathbf{x}_{t_f})], \tag{4.1a}$$

subject to:

$$\dot{\mathbf{x}}_{\mathbf{d}}(t) = \mathbf{g}_{\mathbf{d}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}), \tag{4.1b}$$

$$\mathbf{0} = \mathbf{g}_{\mathbf{a}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}), \tag{4.1c}$$

$$\mathbf{x}_{\mathbf{d}}(0) = \mathbf{x}_0, \tag{4.1d}$$

$$Pr[\mathbf{h}_{\mathbf{nq}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \geq 0] \leq \varepsilon_{nq}, \tag{4.1e}$$

$$Pr[\mathbf{h}_{\mathbf{eq}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \neq 0] \leq \varepsilon_{eq}, \tag{4.1f}$$

$$\mathbf{u}_{min} \leq \mathbf{u} \leq \mathbf{u}_{max}. \tag{4.1g}$$

First-principle models showed in Chapter 2 are used to describe physicochemical mechanisms of pharmaceutical processes mathematically, as in Equations (4.1b) and (4.1c). $E[\cdot]$ and $Var[\cdot]$ denote the mean and the variance of the cost function $M(\mathbf{x}_{t_f})$, respectively, $Pr[\cdot]$ denotes the probability measure, α denotes a scalar weight factor, ε_{nq} and ε_{eq} are tolerance factors, [$\mathbf{u}_{min}$,$\mathbf{u}_{max}$] are the upper and lower boundaries for the control input vector and $\mathbf{x}_{t_f}$ is the state vector at final time t_f. In detail, $M(\mathbf{x}_{t_f})$ denotes a Mayer objective term that is used for nominal optimal control problems. Please note that certain reformulations can be made to consider optimal Lagrange control problems, as well. The two functions $\mathbf{h_{nq}} : \mathbb{R}^{(n_{x_d}+n_{x_a})\times n_u \times n_p} \rightarrow \mathbb{R}^{n_{nq}}$ and $\mathbf{h_{eq}} : \mathbb{R}^{(n_{x_d}+n_{x_a})\times n_u \times n_p} \rightarrow \mathbb{R}^{n_{eq}}$ are used to represent the inequality and equality constraints, which come from process restrictions, such as temperature limitations. The model equations are also considered as equality constraints as discussed in Section 4.2.

In Problem 1, Equation (4.1a) gives the robust form of the objective function $M(\mathbf{x}_{t_f})$, where $E[M(\mathbf{x}_{t_f})]$ and $Var[M(\mathbf{x}_{t_f})]$ represent the expected performance and the robustness of the objective function, respectively. The trade-off between the performance and the robustness is adjusted by the weight factor α. Equations (4.1e) and (4.1f) give the

robust form of the inequality and equality constraints, respectively. They ensure that the probability of all constraint violations is less than or equal to a certain tolerance factor that can be adjusted according to given specifications and safety rules. However, to solve Problem 1 practically, the following two aspects have to be addressed. First, the estimation of the probabilities of both constraint violations cannot be solved in closed form, and standard numerical methods might be computationally demanding. Second, the robust equality constraints in Equation (4.1f), which were rarely considered in previous robust process design studies [124, 30], are infeasible and render robust optimization insolvable. Therefore, in this chapter, the moment methods introduced by [125] for structural reliability analysis are implemented to approximate the robustified constraints. The statistical moments needed for the approximation are computed with the PEM, as it has the highest efficiency compared to other methods and could be embedded in the optimization with affordable computational demand. Moreover, the robust equality constraint is relaxed to ensure the feasibility of the robust optimization problem.

The dependencies of parameter uncertainties, which are referred to as parameter dependencies in the following context, commonly exist in practical applications [63, 110, 65], but are generally not taken into account in robust process design. Chapter 3 shows the significant impact of parameter dependencies on parameter sensitivities and the resulting probability distributions of the model output. A similar observation is also presented in [3, 4]. Therefore, PEM presented in Chapter 2 is adapted to include parameter dependencies for robust process design. The effect of parameter dependencies on the robust process design is investigated and critically compared with the reference case where parameter dependencies are neglected in the selected case studies. Additionally, the global sensitivity analysis technique is also utilized to obtain a better understanding of the process under study and provide information for simplifying and constructing the robust optimization problem systematically.

4.2 Moment Method for Approximating Robust Inequality and Equality Constraints

In this section, details of inequality and equality constraints are discussed. In Section 4.2.1, the constraints are categorized into two special types, i.e., *hard* and *soft* constraints, and discuss the effects of parameter uncertainties on the constraints. In Sec-

tions 4.2.2 and 4.2.3, a robust formulation of *soft* inequality and *soft* equality constraints and methods for approximating the robustified expressions are presented.

4.2.1 Categorization of the Constraints

There are two types of robust inequality and equality constraints: *hard* and *soft* constraints [126]. *Hard* constraints must be satisfied regardless of uncertainties in the RO. *Hard* constraints ensure that optimized results satisfy physical laws. For instance, in Problem 1, equality constraints Equations (4.1b) and (4.1c), i.e., the governing equations, are *hard* constraints as they describe the underlying (bio)chemical processes and have to be consistently satisfied when assuming deterministic simulation results. *Soft* constraints, in turn, do not have to be exactly satisfied under uncertainties. *Soft* constraints (e.g., Equations (4.1e) and (4.1f)) are typically imposed by the designer to restrict the design space and to satisfy additional process specifications. Therefore, *soft* constraints can be satisfied only in a probabilistic manner and might occasionally be violated, i.e., an acceptable violation probability has to be defined for RO. Please note that the performance of the objective function may decrease if a very low violation probability is required. *Soft* constraints are considerably affected by parameter uncertainties and are investigated in the following section.

4.2.2 Robust Formulation of Soft Inequality Constraints

Soft inequality constraints do not have to be strictly satisfied, but in a probabilistic manner. Inequality constraints $\mathbf{h_{nq}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \leq 0$ formulated on the probability space are also named chance constraints [127] and read as:

$$\mathbf{Pr}[\mathbf{h_{nq}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \leq 0] \geq 1 - \varepsilon_{nq}, \tag{4.2}$$

where the probability of constraint satisfaction must be higher or equal to $1-\varepsilon_{nq}$. Please note that Equation (4.2) can also be equivalently transformed into Equation (4.3) when the probability of a constraint violation is used:

$$\mathbf{Pr}[\mathbf{h_{nq}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \geq 0] \leq \varepsilon_{nq}. \tag{4.3}$$

The probability of constraint violations is frequently estimated by MC simulations. A large number of samples are drawn from given parameter distributions, and the samples, where the constraints are violated, are counted. MC simulations are straightforward in implementation but require a considerable number of CPU-intensive model evaluations.

The computational burden might be prohibitive, especially for the iterative nature of the RO. Moment-based approximation of failure probabilities has been widely applied in the field of reliability analysis [125], and thus, this method is used as an alternative concept to approximate the chance constraints in this work. In addition, it takes the advantage of the proposed PEM for estimating the needed statistical moments.

The basic idea of the moment-based approximation method is to transform the probability distribution of the constraint functions into some specific distributions, e.g., the standard normal distribution $\xi \sim \mathcal{N}(0,1)$ and to obtain the failure probability based on the probability. Here, the one-dimensional constraint function $-h_{nq}(\mathbf{x}(t),\mathbf{u}(t),\mathbf{p})$ with a negative sign is abbreviated as $\overline{h}_{nq}$ and used in the following. The isoprobabilistic transform given in Proposition 2.1 is applied to express the relation between the standard normal distribution and the probability distribution of $\overline{h}_{nq}$ as:

$$\xi = F_{\xi}^{-1}(F_{\overline{h}_{nq}}(\overline{h}_{nq})), \tag{4.4}$$

where F_{ξ}^{-1} indicates the inverse CDF of the standard normal distribution and $F_{\overline{h}_{nq}}$ indicates the CDF of $\overline{h}_{nq}$. Based on this transformation, the failure probability of the constraint function $\overline{h}_{nq}$ is equivalent to the probability of $\xi \leq F_{\xi}^{-1}(F_{\overline{h}_{nq}}(0))$ as shown in Equation (4.5). As the CDF of ξ is known analytically, the failure probability of the constraint function can be determined if $F_{\xi}^{-1}(F_{\overline{h}_{nq}}(0))$ is given. However, the transformation function $F_{\xi}^{-1}(F_{\overline{h}_{nq}}(\cdot))$ is typically not available as the CDF of $\overline{h}_{nq}$ is unknown in practice. Thus, transformation rules that are based only on the statistical moments of $\overline{h}_{nq}$ is a better option [125]:

$$\begin{aligned}\mathbf{Pr}[h_{nq}(\mathbf{x}(t),\mathbf{u}(t),\mathbf{p}) \geq 0] &= \mathbf{Pr}[\overline{h}_{nq} \leq 0], \\ &= \mathbf{Pr}[\xi \leq F_{\xi}^{-1}(F_{\overline{h}_{nq}}(0))].\end{aligned} \tag{4.5}$$

Two representative moment-based approximation methods [125], i.e., the second moment method and the fourth moment method, are used to estimate the failure probability with the first four statistical moments of the probability distribution of the constraint function $\overline{h}_{nq}$, which are the mean ($\mu_{\overline{h}_{nq}}$), variance ($\sigma^2_{\overline{h}_{nq}}$), skewness ($\alpha_{\overline{h}_{nq},3}$) and kurtosis ($\alpha_{\overline{h}_{nq},4}$). The second moment method approximates the transformation function with the first two moments as in Equation (4.6), while the fourth moment method utilizes all four moments and has a more complex structure; see Equation (4.7) [125]. The approximations are incorporated in Equation (4.5) to calculate the failure probability of the

constraints:

$$F_{\xi}^{-1}(F_{\overline{h}_{nq}}(0)) = -\frac{\mu_{\overline{h}_{nq}}}{\sigma_{\overline{h}_{nq}}}, \tag{4.6}$$

$$F_{\xi}^{-1}(F_{\overline{h}_{nq}}(0)) = -\frac{3(\alpha_{\overline{h}_{nq},4}-1)(\frac{\mu_{\overline{h}_{nq}}}{\sigma_{\overline{h}_{nq}}}) + \alpha_{\overline{h}_{nq},3}((\frac{\mu_{\overline{h}_{nq}}}{\sigma_{\overline{h}_{nq}}})^2-1)}{\sqrt{(9\alpha_{\overline{h}_{nq},4} - 5\alpha_{\overline{h}_{nq},3}^2 - 9)(\alpha_{\overline{h}_{nq},4}-1)}}. \tag{4.7}$$

The accuracy of the moment-based approximation methods is determined by two factors. The first factor is the intrinsic approximation error, which results from the approximated transformation function (Equation (4.4)) using a limited number of statistical moments. By definition, the fourth moment method has a lower intrinsic approximation error because this method is more rigorously defined with higher order statistical moments. The second factor is the estimation error of the statistical moments, especially the higher order moments, e.g., skewness and kurtosis. The PEM introduced in Section 2.2 is used to calculate the needed statistical moments with considerably lower computational costs in comparison to MC simulations. However, the precision of the estimated statistical moments deteriorates with higher order statistical moments, because the PEM might fail for highly nonlinear problems of higher order terms. Thus, especially the fourth moment method may suffer from the estimation error. According to these two sources of approximation errors, it is difficult to determine which approximation method, i.e., the second or fourth moment method, is superior for robust process design. Therefore, both concepts are further analyzed and their benefits for efficient and credible robustification strategies are investigated in the case studies.

4.2.3 Robust Formulation of Soft Equality Constraints

Similar to the inequality constraints, *soft* equality constraints are considered in a probabilistic manner for the RO problem and are given as:

$$\mathbf{Pr}[\mathbf{h_{eq}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \neq 0] \leq \varepsilon_{eq} \tag{4.8}$$

However, Equation (4.8) is not directly solvable for most applications as the constraint function h_{eq} has a continuous probability distribution. In other words, the probability of a single point is equal to zero when the random space is continuous [128]. Thus, it

can find that:

$$\mathbf{Pr}[\mathbf{h_{eq}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \neq 0] = 1, \tag{4.9}$$

which contradicts Equation (4.8) if $\varepsilon_{eq} \lneq 1$. Note that the aim is to satisfy the equality constraint with high probability, and thus, $\varepsilon_{eq} \ll 1$. Figure 4.1a shows an example of the equality constraint in the random parameter space. Here, the samples are drawn from their distributions, and the curve shows the locations where the samples satisfy the constraints.

To solve the RO problem, the robust equality constraints must be relaxed as shown in Figure 4.1b. This idea is analogous to the relaxed margin used in support vector machines (SVMs), which have been applied extensively in machine learning [129]. The restriction from the constraints is eased by admitting that samples can lie within a certain range around the constraints. Based on the relaxation, the robust equality constraints in Equation (4.8) are substituted by:

$$\mathbf{Pr}[-\delta_{eq} \leq \mathbf{h_{eq}}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \leq \delta_{eq}] \geq 1 - \varepsilon_{eq}, \tag{4.10}$$

where δ_{eq} indicates the relaxation factor and determines the range of relaxed equality constraints. As can be seen in Figure 4.1b and Equation (4.10), there is a region rather than a single curve where the constraint is satisfied. Thus, probability is nonzero, and the RO problem becomes solvable. The robust equality constraints in Equation (4.10) have nearly the same structure as the robust inequality constraints in Equation (4.2). Therefore, the methods described in Section 4.2.2 can be used to solve Equation (4.10) in RO problems immediately.

As mentioned previously, there is a trade-off between the performance of the objective function and the satisfaction probability of *soft* inequality and *soft* equality constraints. The relevant factors, $\varepsilon_{nq}, \varepsilon_{eq}$ and δ_{eq}, have to be adapted properly. More details about how to select these factors are presented with the given case studies in Section 4.4.

4.3 Robust Optimization with the PEM

The final structure to solve the RO problem defined in Problem 1 is summarized in what follows. Note that $F(\cdot)$ in Equations (4.18) and (4.19) indicates the CDF of a standard Gaussian distribution. The PEM is used to estimate the relevant statistical moments to include the effect of parameter uncertainties. Equations (4.13)–(4.17) are the evaluations of the dynamic system and the constraint functions for all deterministic sample points

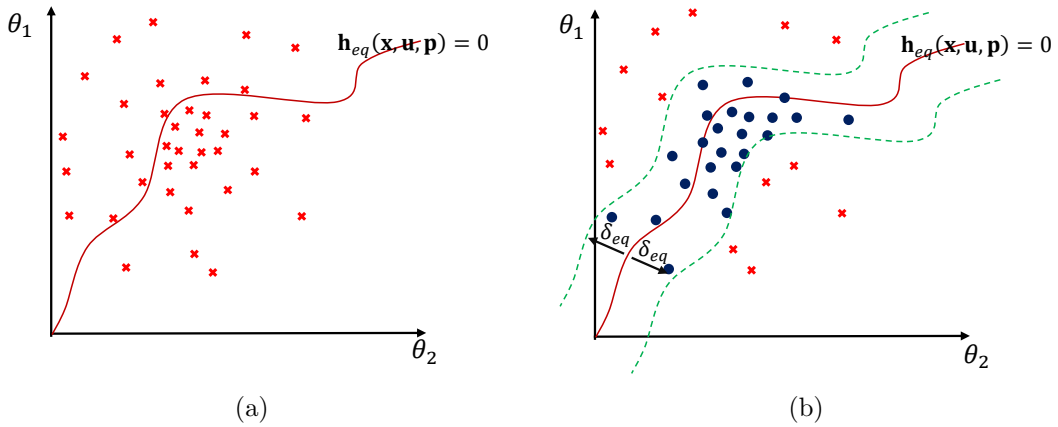


Figure 4.1: Illustration of *soft* equality constraints $\mathbf{h_{eq}}(\mathbf{x}, \mathbf{u}, \mathbf{p}) = 0$. For the sake of explanation, a two-dimensional random space with uncertain parameters θ_1 and θ_2 is used. Samples satisfying the constraints are shown by blue-filled circles ●, while samples that violate the constraints are shown by red cross ✖. (a) The probability of samples that satisfy the equality constraint (red line —) is equal to zero for the continuous random space; (b) the equality constraint and its relaxed boundaries (green dashed line - -) with width δ_{eq}. The probability of satisfying the equality constraints is given by the percentage of samples, i.e., ●, which are located within the boundaries.

that are generated from the probability distributions of the uncertain model parameter. Based on the evaluations, Equations (4.21)–(4.26) calculate the statistical moments of the objective function and constraints, which are used in Equations (4.11), (4.18) and (4.19). Although Equations (4.18) and (4.19) demonstrate the approximation with the fourth moment method, it is easy to switch to the second moment method by adapting the structure with Equation (4.6).

$$\min_{\mathbf{x}(\cdot),\mathbf{u}(\cdot)} \mathbf{E}[M(\mathbf{x}_{t_f})] + \alpha \mathbf{Var}[M(\mathbf{x}_{t_f})], \tag{4.11}$$

subject to:

$$i = 1, \ldots, 2d^2 + 1, \qquad m = 1, 2, 3 \tag{4.12}$$

$$\boldsymbol{\theta}_i = [\mathbf{p}_i, \mathbf{x}_{0,i}]^T, \mathbf{x}_i = [\mathbf{x}_{\mathbf{d},i}, \mathbf{x}_{\mathbf{a},i}]^T, \mathbf{x}_{\mathbf{d},i}(0) = \mathbf{x}_{0,i}, \mathbf{x}_{t_f,i} = \mathbf{x}_i(t_{final}), \tag{4.13}$$

$$\dot{\mathbf{x}}_{\mathbf{d},i}(t) = \mathbf{g}_{\mathbf{d}}(\mathbf{x}_i(t), \mathbf{u}(t), \mathbf{p}_i), \quad \mathbf{0} = \mathbf{g}_{\mathbf{a}}(\mathbf{x}_i(t), \mathbf{u}(t), \mathbf{p}_i), \tag{4.14}$$

$$\overline{h}_{1,i} = -h_{nq}(\mathbf{x}_i(t), \mathbf{u}(t), \mathbf{p}_i), \tag{4.15}$$

$$\overline{h}_{2,i} = h_{eq}(\mathbf{x}_i(t), \mathbf{u}(t), \mathbf{p}_i) + \delta_{eq}, \tag{4.16}$$

$$\overline{h}_{3,i} = -h_{eq}(\mathbf{x}_i(t), \mathbf{u}(t), \mathbf{p}_i) + \delta_{eq}, \tag{4.17}$$

$$F\left(-\frac{3(\alpha_{\overline{h}_1,4} - 1)(\frac{\mu_{\overline{h}_1}}{\sigma_{\overline{h}_1}}) + \alpha_{\overline{h}_1,3}((\frac{\mu_{\overline{h}_1}}{\sigma_{\overline{h}_1}})^2 - 1)}{\sqrt{(9\alpha_{\overline{h}_1,4} - 5\alpha_{\overline{h}_1,3}^2 - 9)(\alpha_{\overline{h}_1,4} - 1)}}\right) \leq \varepsilon_{nq}, \tag{4.18}$$

$$F\left(-\frac{3(\alpha_{\overline{h}_2,4} - 1)(\frac{\mu_{\overline{h}_2}}{\sigma_{\overline{h}_2}}) + \alpha_{\overline{h}_2,3}((\frac{\mu_{\overline{h}_2}}{\sigma_{\overline{h}_2}})^2 - 1)}{\sqrt{(9\alpha_{\overline{h}_2,4} - 5\alpha_{\overline{h}_2,3}^2 - 9)(\alpha_{\overline{h}_2,4} - 1)}}\right) +$$

$$F\left(-\frac{3(\alpha_{\overline{h}_3,4} - 1)(\frac{\mu_{\overline{h}_3}}{\sigma_{\overline{h}_3}}) + \alpha_{\overline{h}_3,3}((\frac{\mu_{\overline{h}_3}}{\sigma_{\overline{h}_3}})^2 - 1)}{\sqrt{(9\alpha_{\overline{h}_3,4} - 5\alpha_{\overline{h}_3,3}^2 - 9)(\alpha_{\overline{h}_3,4} - 1)}}\right) \leq \varepsilon_{eq}, \tag{4.19}$$

$$\mathbf{u}_{min} \leq \mathbf{u} \leq \mathbf{u}_{max}, \tag{4.20}$$

$$\mathbf{E}[M(\mathbf{x}_{t_f})] = w_0 M(\mathbf{x}_{t_f,1}) + w_1 \sum_{i=2}^{2d+1} M(\mathbf{x}_{t_f,i}) + w_2 \sum_{j=2d+2}^{2d^2+1} M(\mathbf{x}_{t_f,j}), \tag{4.21}$$

$$\mathbf{Var}[M(\mathbf{x}_{t_f})] = w_0(M(\mathbf{x}_{t_f,1}) - \mathbf{E}[M(\mathbf{x}_{t_f})])^2 + w_1 \sum_{i=2}^{2d+1} (M(\mathbf{x}_{t_f,i}) - \mathbf{E}[M(\mathbf{x}_{t_f})])^2 + w_2 \sum_{j=2d+2}^{2d^2+1} (M(\mathbf{x}_{t_f,j}) - \mathbf{E}[M(\mathbf{x}_{t_f})])^2, \tag{4.22}$$

$$\mu_{\overline{h}_m} = w_0 \overline{h}_{m,1} + w_1 \sum_{i=2}^{2d+1} \overline{h}_{m,i} + w_2 \sum_{j=2d+2}^{2d^2+1} \overline{h}_{m,j}, \tag{4.23}$$

$$\sigma^2_{\overline{h}_m} = w_0(\overline{h}_{m,1} - \mu_{\overline{h}_m})^2 + w_1 \sum_{i=2}^{2d+1} (\overline{h}_{m,i} - \mu_{\overline{h}_m})^2 + w_2 \sum_{j=2d+2}^{2d^2+1} (\overline{h}_{m,j} - \mu_{\overline{h}_m})^2, \tag{4.24}$$

$$\alpha_{\overline{h}_m,3} = \frac{w_0(\overline{h}_{m,1} - \mu_{\overline{h}_m})^3 + w_1 \sum_{i=2}^{2d+1} (\overline{h}_{m,i} - \mu_{\overline{h}_m})^3 + w_2 \sum_{j=2d+2}^{2d^2+1} (\overline{h}_{m,j} - \mu_{\overline{h}_m})^3}{\sigma^3_{\overline{h}_m}}, \tag{4.25}$$

$$\alpha_{\overline{h}_m,4} = \frac{w_0(\overline{h}_{m,1} - \mu_{\overline{h}_m})^4 + w_1 \sum_{i=2}^{2d+1} (\overline{h}_{m,i} - \mu_{\overline{h}_m})^4 + w_2 \sum_{j=2d+2}^{2d^2+1} (\overline{h}_{m,j} - \mu_{\overline{h}_{nq}})^4}{\sigma^4_{\overline{h}_m}}, \tag{4.26}$$

4.4 Case Studies

In this section, the performance of the proposed framework is demonstrated with two case studies. In Case Study 1, an optimal jacket temperature profile is design for a tubular reactor considering two uncertain and correlated model parameters. Additionally, a robust equality constraint for the product temperature at the reactor outlet is assumed to incorporate process intensification aspects in the design problem. In Case Study 2, a penicillin fermentation process is analyzed as it is of interest in the pharmaceutical industry. A fed-batch bioreactor model is used to design an optimal feeding profile under parameter uncertainties. GSA is applied to determine the influence of parameter uncertainties on the process states and to offer a more tractable problem, i.e., a reduced number of uncertain model parameters, which have to be considered in the robust process design.

GSA and the RO problem were solved in MATLAB®(Version 2017b, The MathWorks Inc., Natick, Massachusetts, USA). Parameter sensitivities for the independent case were calculated with UQLAB (Version 1.0, ETH Zurich, Switzerland) [122]. The RO problem

for the first case study was solved with the MATLAB function *fmincon*, while the RO problem for the second case study, which is more complex, was solved by the simultaneous approach [130] and implemented in the symbolic framework CasADi (Verion 3.3.0, KU Leuven, Belgium) for numerical optimization [131] using the NLP solver IPOPT [132] and the MA57 linear solver [133].

4.4.1 Case Study 1: A Jacket Tubular Reactor

Here, the design of a tubular reactor, where an irreversible first-order reaction in Equation (4.27) takes place, is considered as the first benchmark case study [30].

$$A \rightarrow B + C. \tag{4.27}$$

The reactor, which is operated under the steady-state condition, is described by the following governing Equations [30]:

$$\frac{d\mathbf{x}_1}{dz} = \frac{\alpha_{kin}}{v}(1-\mathbf{x}_1)e^{\frac{\gamma \mathbf{x}_2}{1+\mathbf{x}_2}}, \tag{4.28}$$

$$\frac{d\mathbf{x}_2}{dz} = \frac{\alpha_{kin}\delta}{v}(1-\mathbf{x}_1)e^{\frac{\gamma \mathbf{x}_2}{1+\mathbf{x}_2}} + \frac{\beta}{v}(\mathbf{u}-\mathbf{x}_2), \tag{4.29}$$

where z is the relative position along the reactor, $0 \leq z \leq 1$. The states $\mathbf{x}_1$ and $\mathbf{x}_2$ are the dimensionless forms of the reactant concentration of A and the reactor temperature, respectively. The jacket temperature is the control input given in its dimensionless form $\mathbf{u} = (T_j - T_{in})/T_{in}$ and is adjusted to meet the desired performance and robustness requirements. The control input is discretized into 25 equidistant elements constrained within 280 K and 400 K. The kinetic coefficient α_{kin} and the heat transfer coefficient β are assumed to be uncertain [30]. They both follow Gaussian distributions with mean at their nominal value and standard deviation equal to 10 % of their nominal values. The implemented parameter values and operating conditions are summarized in Table 4.1. For additional details of the proposed reactor model, the interested reader is referred to [30], from which the case study is taken. The conversion of the reactor C_f, as well as the reactor temperature T_r, can be calculated from their dimensionless form via:

$$C_f(z=1) = x_1(z=1), \tag{4.30}$$

$$T_r(z) = x_2(z) \times T_{in} + T_{in}. \tag{4.31}$$

In this case study, the final conversion of reactant A is maximized while fulfilling the given constraints on the reactor temperature [30]. In particular, an upper boundary is

added to the reactor temperature to avoid undesired side reactions [30]. The results of the deterministic optimal design are depicted on the left of Figure 4.2. As can be seen, the reactor temperature increases rapidly to the upper boundary to ensure the maximum reaction rate and final conversion of 0.996, respectively. However, numerous violations of the temperature boundary occur when the parameter uncertainties are taken into account. In contrast to the deterministic process design, a robust optimal design that includes parameter uncertainties is conducted next. Here, a weight factor $\alpha = 3$ and a tolerance value $\varepsilon_{nq} = 1\%$ are used for the robust objective and inequality constraints. Please note that the weight factor α indicates the amount of trade-off between process performance and robustness of objective function, and is selected based on one previous study [68]. The tolerance value $\varepsilon_{np} = 1\%$ means the robust solution has to guarantee that the violation probability of inequality constraints should not be larger than 1%, and could be changed depending on robustness required for the inequality constraint. The robust inequality constraints are approximated with the second moment method as in Equation (4.6). The corresponding results are given on the right of Figure 4.2. As can be seen from the results, the jacket temperature profile is different from the nominal design, especially from location 0.3 to the end of the reactor. Moreover, the reactor temperature profile of the robust design remains below its upper limit with a probability of 99% with the loss in the reactor performance; i.e., final conversion decreases to 0.985.

Table 4.1: Parameters for the tubular reactor model.

Parameters	Unit	Nominal Value	Uncertainty
$x_1(0)$	-	0	-
$x_2(0)$	-	0	-
α_{kin}	s^{-1}	0.058	$\mathcal{N}(0.058, 0.0058^2)$
β	s^{-1}	0.2	$\mathcal{N}(0.2, 0.02^2)$
v	$\mathrm{m\,s}^{-1}$	0.1	-
γ	-	16.66	-
δ	-	0.25	-

4.4.1.1 Robust Design with Parameter Correlation

Next, the influence of parameter correlation on the robust process design is investigated. The two uncertain parameters α_{kin} and β are assigned with the marginal distributions shown in Table 4.1 and the additional Pearson correlation coefficient of 0.8. Deterministic sample points for the correlated parameters are generated with Algorithm 2.1 and

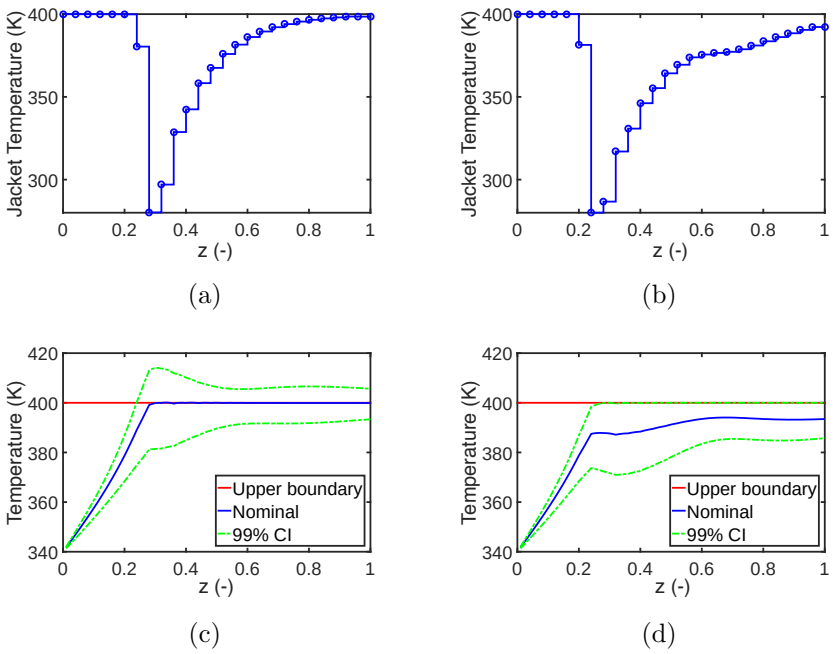
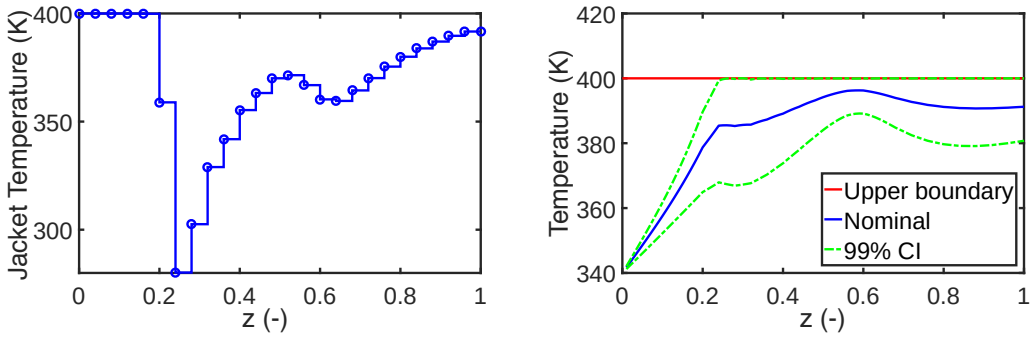


Figure 4.2: Results for nominal design (left) and robust design (right). (a,b) are the optimal profiles of the jacket temperature; (c,d) are the evolution of the reactor temperature and the 99% confidence interval (CI). The mean and standard deviation of the conversion of reactant A have values of [0.996, 0.004] and [0.985, 0.010] for the nominal and robust design, respectively.

the generalized PEM is used to calculated the statistic moments used in the robust objective function and constraints. The structure of the RO problem is similar to that for independent parameters with a weight factor $\alpha = 3$ and a tolerance value $\varepsilon_{nq} = 1\%$. Here, too, the second moment method is applied. In Figure 4.3, results for the optimal design with parameter correlation are given. As can be seen, the profile of the jacket temperature has considerable differences compared to the independent case; see Figure 4.2. Especially, the drop in the jacket temperature between position $z = 0.5$ and $z = 0.8$ results from the parameter dependencies effect.



(a) (b)

Figure 4.3: Results for robust design with parameter correlation: (a) the optimal jacket temperature profile and (b) the reactor temperature and its 99% confidence interval (CI). The mean and standard deviation of the conversion of reactant A are 0.986 and 0.008, respectively.

4.4.1.2 Performance of the Fourth Moment Method

Thus far, only the second moment method has been used to approximate the robust inequality constraints. The resulting confidence intervals of the reactor temperature are illustrated with green dashed curves in Figures 4.2d and 4.3b. As can be observed, the upper boundaries of the confidence intervals stay the same with the upper limit of the reactor temperature once they approach it. However, the confidence intervals are approximated by taking into account only the first and second statistical moments and are insufficient if the probability distribution of the reactor temperature is non-Gaussian. Reference values based on MC simulations with 10,000 sample points are summarized in Table 4.2. In the case of the second moment method, the violation probabilities are 4.7% and 3.6%, respectively, which exceeds the tolerance value $\varepsilon_{nq} = 1\%$. The reason for this

mismatch is mainly due to the approximation error of the robust inequality constraints while considering only the first two statistical moments.

Table 4.2: The number of constraint violations from 10,000 Monte Carlo simulations, where the robust inequality constraints are approximated by the second and fourth moment methods for process designs with independent and correlated parameters.

	Second Moment Method		Fourth Moment Method	
Number of violations	Independent 470	Correlated 357	Independent 440	Correlated 385
Probability	0.047	0.036	0.044	0.039

As discussed in Section 4.2.2, the fourth moment method uses more statistical information than the second moment method and has a lower approximation error. The same RO problem is solved again with the fourth moment approach, and the violation probabilities are estimated and listed in the right of Table 4.2. However, the expected improvement could not be validated. In fact, the violation probability for the correlated scenario increases in case of the fourth moment method. The reason for this unexpected performance is mainly due to the estimation error of higher order statistical moments. When comparing the first four statistical moments estimated by the PEM and the MC simulations (as reference), it can be seen in Figure 4.4 that the PEM provides useful approximations for the first and second moments and deteriorates considerably for the higher order moments. As has been mentioned, the PEM is accurate if the system can be accurately approximated with the sum of monomials up to order of five, and as such its accuracy deteriorates with the increasing order of the statistical moments. The comparison indicates that the fourth moment method might not be suitable for the PEM-based robust optimization framework, especially for practical applications where the systems might be already highly nonlinear and complex. The fourth moment approach, in turn, is still a promising way to approximate probability distributions if the higher order moments can be estimated accurately, e.g., using PCE. Based on this finding, in the following section, the second moment method is exclusively implemented.

Alternatively, one might adjust the tolerance value for the robust inequality constraints to mitigate the effect of approximation errors when using the second moment method. The violation probabilities of the inequality constraints for the robust design with different tolerance values are given in Figure 4.5. As can be seen from the figure, the

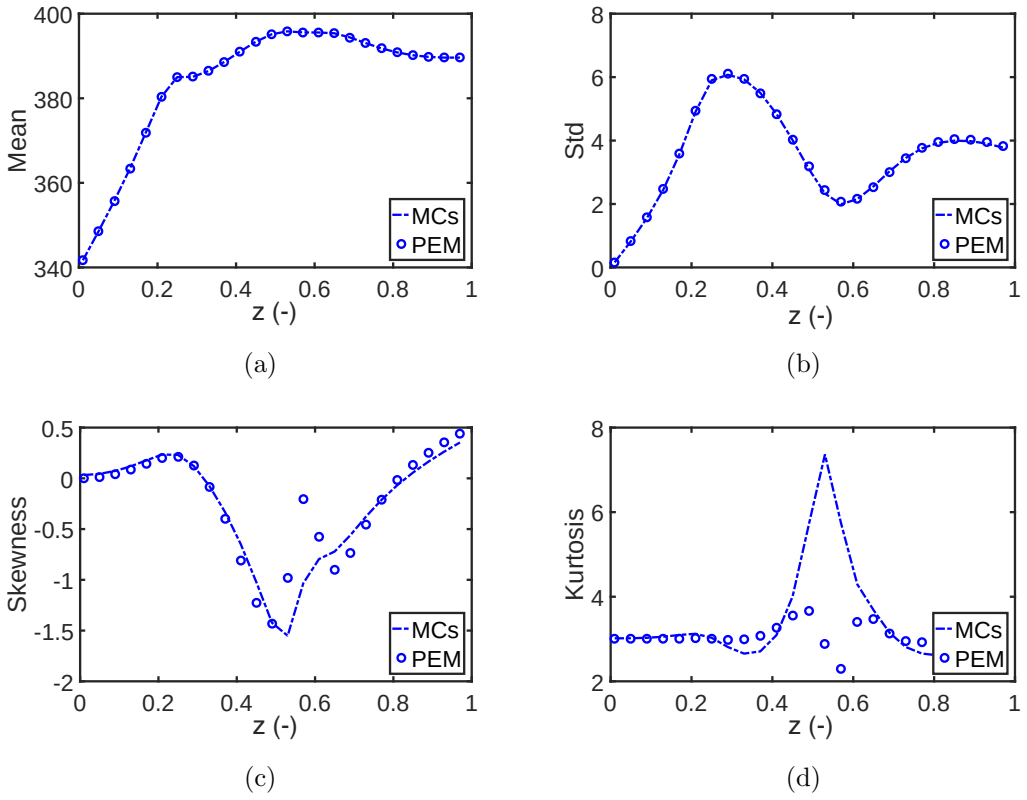


Figure 4.4: A comparison of the first four statistical moments, (a) mean value, (b) standard deviation, (c) skewness, and (d) kurtosis, estimated with the point estimate method (PEM) and Monte Carlo (MC) simulations for the reactor temperature in Case Study 1.

probability can achieve 0.01 by setting the tolerance value to 0.002 for the independent and correlated cases, while the average conversion of reactant A is slightly decreased.

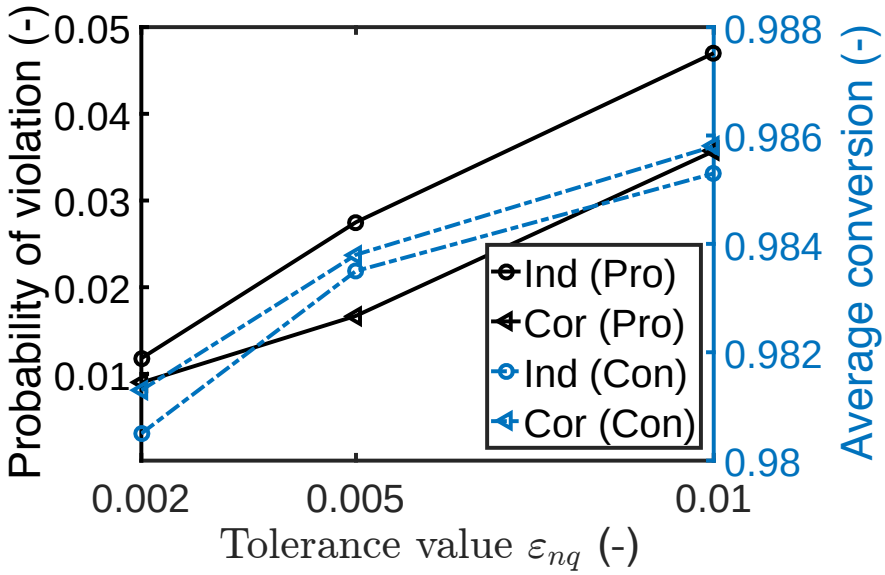


Figure 4.5: The violation probability of the reactor temperature (Pro) and the average conversion of the reactant A (Con) for process designs with different tolerance values. Ind and Cor indicate the results for the independent and correlated scenarios.

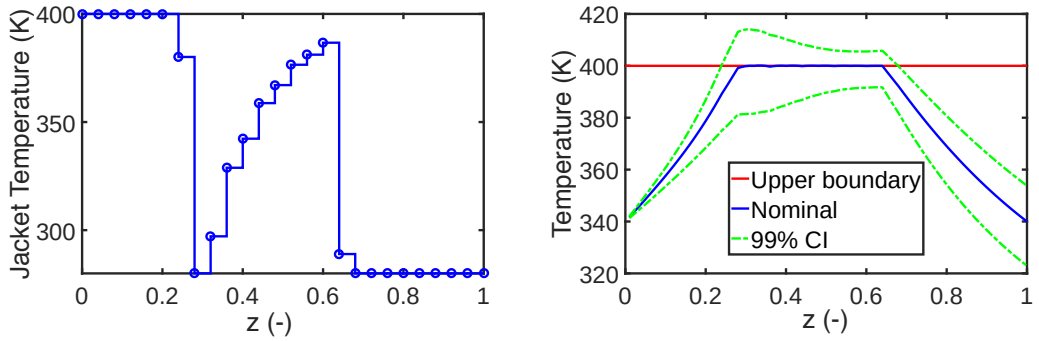
4.4.1.3 Impact of Robust Equality Constraints

Here, the effect of robust equality constraints that might result from process specifications is investigated. Process integration and intensification for energy and raw material saving purpose is added as additional design target of the continuous processes [30]. For instance, a cooling part could be integrated into the reactor design and bring down the reactor outlet temperature, so that extra cooling expenses for downstream process is avoided. To this end, an equality constraint is added to lower the outlet temperature to the value of the inlet temperature [30]:

$$\left|T_r(z=1) - T_r(z=0)\right| = 0. \tag{4.32}$$

With this additional *soft* equality constraint, there exists a trade-off between maximizing the reactant conversion while minimizing the temperature difference. First, the results of the reactor design, where parameter uncertainties are neglected, are given in

Figure 4.6a. The jacket temperature drops sharply to its lower limit for the second half of the reactor, and the outlet temperature returns exactly to 340 K. Consequently, the reactant conversion decreases with 2% compared to the nominal design without the equality constraint (Figure 4.2). Next, the effect of parameter uncertainties on the nominal design is illustrated in Figure 4.6b with the green dotted line. Here, the case where uncertain parameters are correlated is considered. In this case, a strong violation of inequality and equality constraints exists and has to be tackled properly. The robust optimization framework proposed in Section 4.3 is used to solve this problem. An identical setting ($\alpha = 3$ and $\varepsilon_{nq} = 1\%$) is used for the objective function and inequality constraints here. Different scenarios with different relaxation factors δ_{eq} and tolerance factors ε_{eq} are used to demonstrate the effect of robust equality constraints on the process performance. Values for the relaxation factors and results are summarized in Figure 4.7. It can be seen that the probability distribution of the outlet temperature narrows quickly once the relaxed region and the violation probability are reduced, while the performance of the reactor (the reactant conversion) deteriorates considerably. Therefore, the process engineer has to decide on the trade-off between product performance and energy expense. Note that the robust inequality constraints in these scenarios are always satisfied, and thus, are not explicitly shown here.



(a) (b)

Figure 4.6: Results for the nominal design with terminal equality constraints: (a) the optimal jacket temperature profile and (b) the reactor temperature with its 99% confidence interval (CI). The mean and standard deviation of the conversion of reactant A are 0.980 and 0.016, respectively.

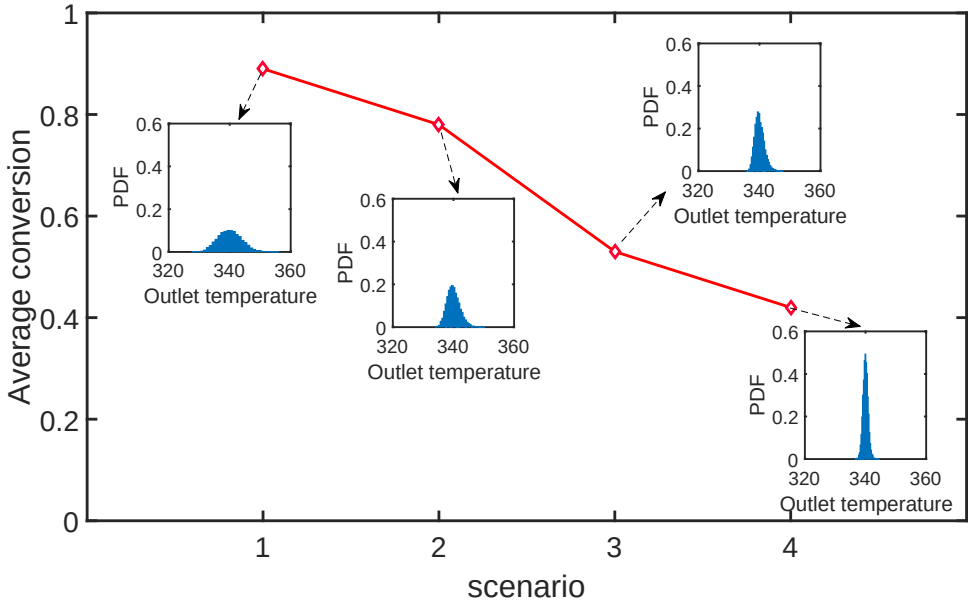


Figure 4.7: The average conversion of reactant A and the probability density function of the outlet temperature for four scenarios that have different relaxation factors δ_{eq} and tolerance factors ε_{eq}. 1: $\delta_{eq} = 5, \varepsilon_{eq} = 10\%$, 2: $\delta_{eq} = 5, \varepsilon_{eq} = 1\%$, 3: $\delta_{eq} = 2, \varepsilon_{eq} = 10\%$, 4: $\delta_{eq} = 2, \varepsilon_{eq} = 1\%$.

4.4.2 Case Study 2: Fed-Batch Bioreactor for Fermentation of Penicillin

The performance of the PEM-based robust optimization framework is also demonstrated with a fermentation process as illustrated in Figure 4.8. Fermentation processes have received great interest in the pharmaceutical industry, and in this study, the focus is to optimize the penicillin fermentation [134]. To this end, a substrate feeding profile is designed to ensure the optimal performance and robustness of the bioreactor which is the essential unit of the fermentation process. A fed-batch reactor model is used to describe the bioreactor based on the following assumptions: (1) ideal mixing of all components in the bioreactor; (2) isothermal condition in the reactor; and (3) the effect of the oxygen transfer can be neglected by considering an upper limitation on the biomass and substrate concentrations. The mathematical expression of the model reads as [134]:

$$\frac{dX}{dt} = \mu X - \frac{F}{V}X \tag{4.33}$$

$$\frac{dS}{dt} = -\frac{\mu}{Y_x}X - \frac{\theta_p}{Y_p}X - m_x X + \frac{F}{V}(S_f - S) \tag{4.34}$$

$$\frac{dP}{dt} = \theta_p X - KP - \frac{F}{V}P \tag{4.35}$$

$$\frac{dV}{dt} = F, \tag{4.36}$$

where the state variables, X, S, P and V, indicate the concentration of the biomass, substrate, product and the volume of components in the bioreactor, respectively. The feeding stream of the substrate has a constant concentration S_f and a time-dependent flow rate F. The specific growth rate of the biomass μ and the product θ_p is described by the contois kinetics and substrate inhibition kinetic with the following form [134]:

$$\mu = \frac{\mu_m S}{S + K_x X} \tag{4.37}$$

$$\theta_p = \frac{\theta_m S}{S + K_p + S^2/K_i}. \tag{4.38}$$

in which μ_m is the maximum specific growth rate, K_x is the Contois Saturation constant, θ_m is maximum specific production rate, K_p is the Michaelis-Menten constant and Ki is the dissociation constant. The initial conditions of the state variables and the nominal value of the other kinetic parameters are summarized in Table 4.3. Further details about the model are given in [134].

First, the process is optimized assuming that all parameters are estimated precisely; i.e., parameter uncertainties are neglected. The goal is to maximize the final concentra-

Table 4.3: Nominal values of the model parameters and the initial conditions for the fed-batch model.

Parameters	Unit	Nominal Value	Parameters	Unit	Nominal Value
μ_m	1/h	0.11	m_x	1/h	0.029
K_x	-	0.006	S_f	g/L	400
θ_m	1/h	0.004	t	h	0–80
K_p	g/L	0.0001	$X(0)$	g/L	1
K_i	g/L	0.1	$S(0)$	g/L	0.5
K	1/h	0.01	$P(0)$	g/L	0
Y_x	-	0.47	$V(0)$	L	250
Y_p	-	1.2			

tion of product P within a given time range while the concentration of biomass X and substrate S should be below 40 g/L (limited by the oxygen transfer capacity) and 0.5 g/L (to avoid side reactions) for the entire time horizon, respectively [135]. The control variable F is parametrized with 100 elements, which are bounded within the range of [0, 10]. The resulting dynamic optimization problem is firstly solved with the nominal value of all parameters, and the results are shown in Figure 4.9. Here, the feed rate of the substrate is adjusted to keep the substrate concentration equal to 0.5 g/L at which the maximum growth rate of the biomass is achieved at the beginning. After the biomass concentration reaches its upper limit, the substrate concentration drops nearly to zero to cease the self-reproduction of the biomass. At the same time, the substrate is fed at low rate and is consumed by the biomass to produce the desired product, i.e, penicillin.

However, due to imperfect measurement data and model simplifications, the estimates of the model parameters may have a considerable error as well as being correlated. Based on the results given in [53], the uncertainties in the nine parameters are assigned with a multivariate Gaussian distribution, where their marginal distributions have the mean values equal to the nominal values and standard deviations equal to 10% of the nominal values. To investigate the effect of parameter correlations, two scenarios are analyzed: (1) the parameter dependencies are neglected, and the correlation matrix Σ is set to the identity matrix; (2) the correlation coefficients of μ_m, θ_m, Y_x, and m_x in Σ are set to 0.95. The effect of parameter uncertainties on the process performance is also shown in Figure 4.9 with the blue dotted lines. Strong violation of the constraints and large variation in the product quality are observed, and thus, the parameter uncertainty has to be considered in the robust design of the fermentation process. Please note that the negative

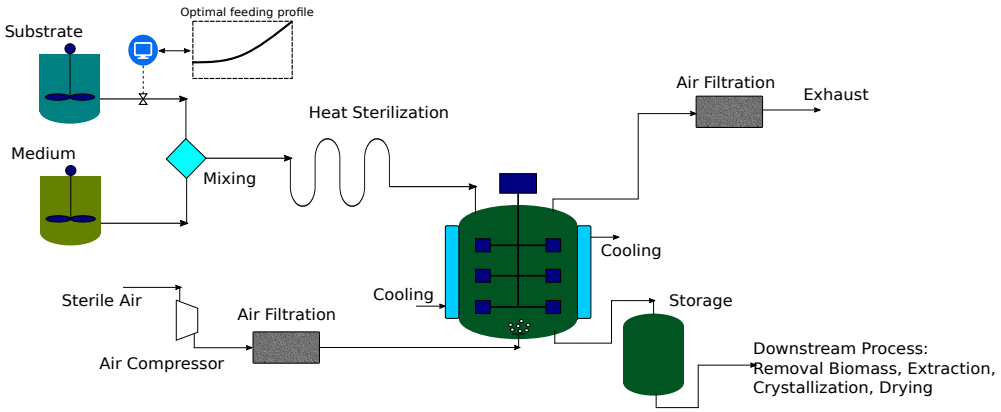


Figure 4.8: Scheme of a fermentation process with a fed-batch bioreactor.

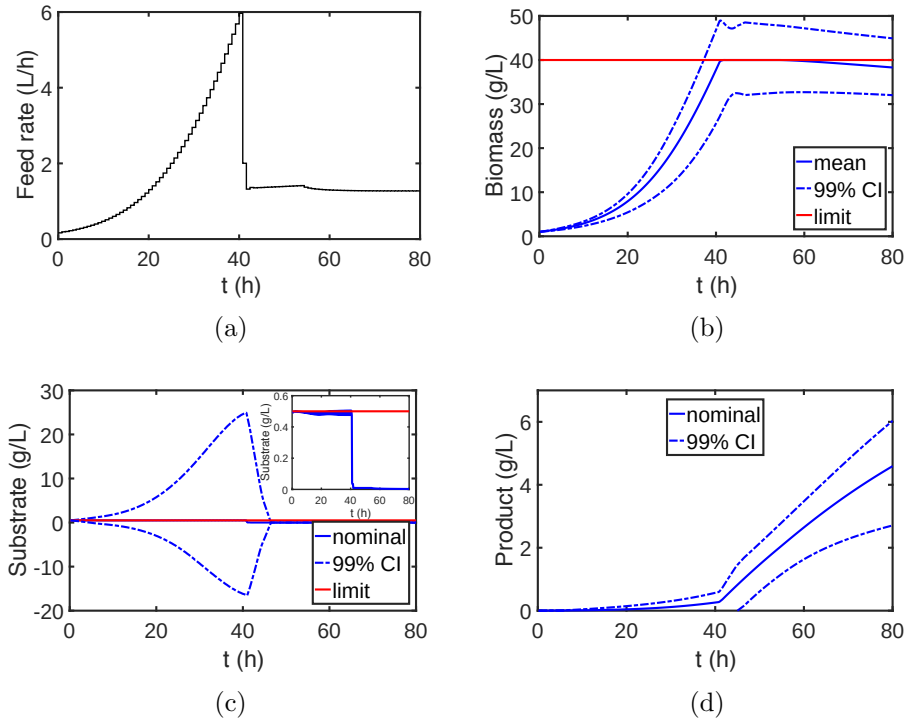


Figure 4.9: (a) Feeding profile; evolution of the (b) biomass; (c) substrate and (d) product obtained from the nominal design, where the parameter uncertainties are neglected. In turn, the blue dotted lines illustrate the effect of the parameter uncertainties.

confidence interval (CI) of the substrate concentration stems from the assumption that the CIs are symmetric and directly derived from the mean and variance of the states.

4.4.2.1 Global Sensitivity Analysis

Before solving the RO problem for the fermentation process, it is desired to decrease its computational cost by deciding which parameters are not relevant and can be neglected in the robust process design. Thus, the corresponding time-dependent sensitivity indices of the parameters are calculated for the biomass and substrate concentrations in addition to the product concentration at the final time point, i.e., for those quantities involved in either the objective function or the constraints of the optimization problem. Figures 4.10a,c and 4.11a show the sensitivity results for the independent case. As can be seen, the biomass and product concentrations are strongly affected by parameters μ_m, θ_m, Y_x, and m_x, while the other parameters have a minor impact. Moreover, by summing up the first-order sensitivity indices, the interaction among the parameters are negligible. Next, the correlative (S^C) and total covariance-based (S^{cov}) first-order sensitivity regarding the biomass, substrate, and product concentration are calculated; see Figures 4.10b,d and 4.11b. Here, the results for the structural sensitivity indices and all the total sensitivity indices are not shown. The reason is that the model structure does not change with the existence of parameter correlations, and thus, the structural sensitivity indices and parameter interactions are similar to those for the independent case. Nevertheless, an evident effect of parameter correlations on the sensitivity analysis result can be observed: The parameter sensitivities have a completely different trend compared to the independent case. The sensitivity results from the correlated case also suggest considering the uncertainties and correlations from μ_m, θ_m, Y_x, and m_x for the RO problem. By using the information from the sensitivity analysis, the number of required PEM points for the RO problem can be significantly reduced. The number of model evaluations for each optimization iteration decreases from $2 \times 9^2 + 1 = 163$ to $2 \times 4^2 + 1 = 33$ for both independent and correlated cases. The performance of the RO with parameter uncertainties of appreciable sensitivities is studied in the following section.

4.4.2.2 Robust Optimization

The RO is solved with the framework proposed in Section 4.3. To this end, a weight factor $\alpha = 3$ and a tolerance value $\varepsilon_{nq} = 1\%$ are used for the robust objective and inequality constraints. The PEM points for RO are generated only for parameters with

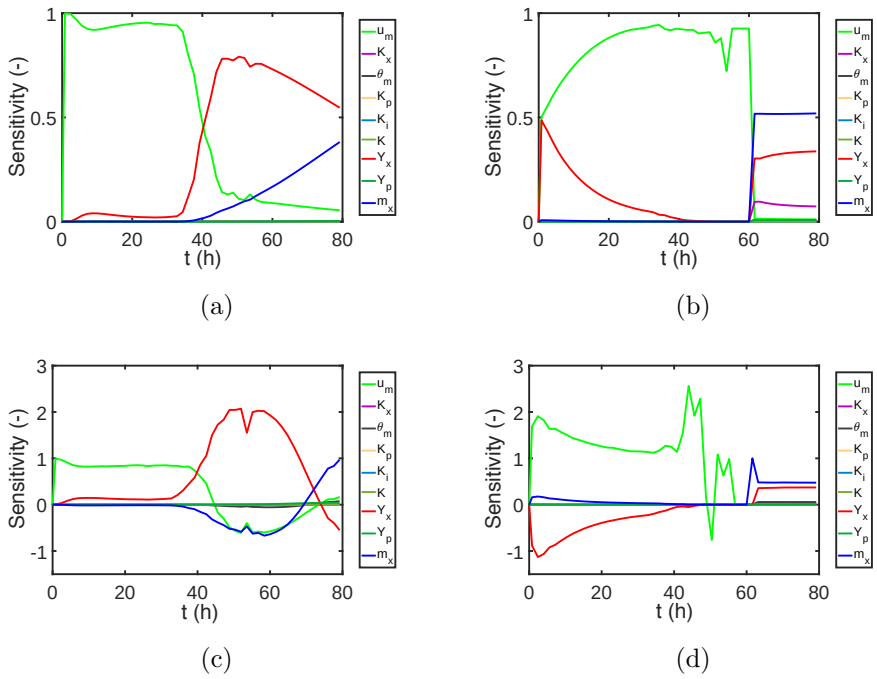


Figure 4.10: Sensitivity results of the nine parameters on the biomass and substrate concentrations for the independent case: (a) first-order sensitivity indices for the concentration of biomass X; (b) first-order sensitivity indices for the concentration of substrate S; and correlated case (c) total covariance-based first-order sensitivity indices for biomass X; (d) total covariance-based first-order sensitivity indices for substrate S.

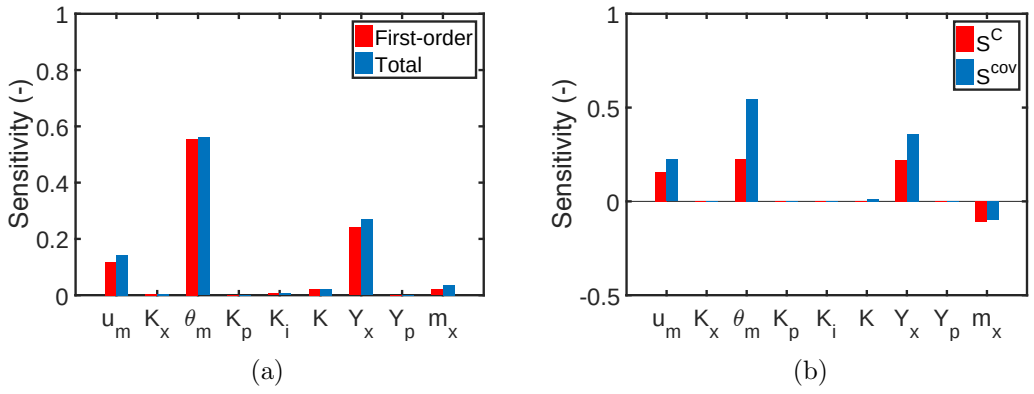
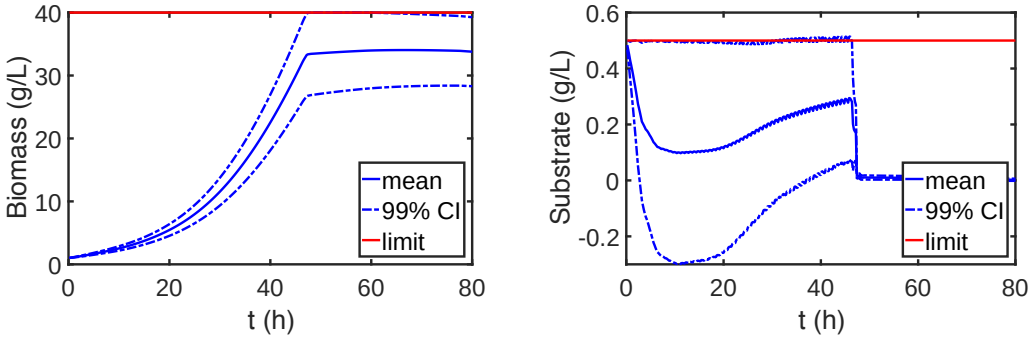


Figure 4.11: Sensitivity results of nine parameters on the final product concentrations for the independent (a) and correlated (b) case.

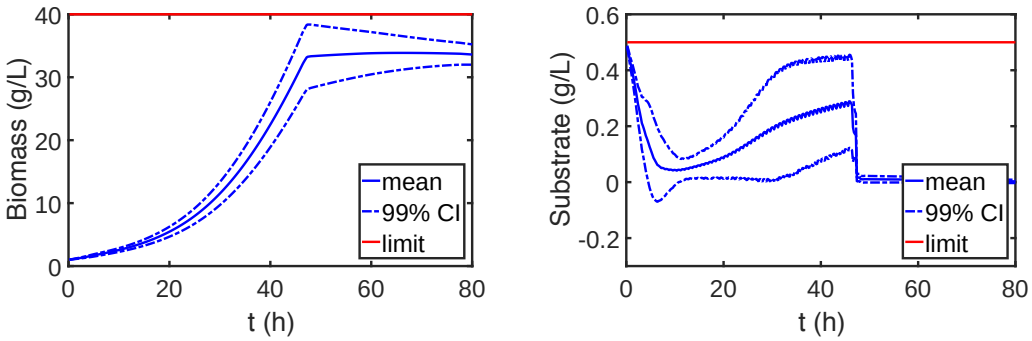
appreciable sensitivities, i.e., four parameters are considered. First, the RO is solved for the simplifying assumption of the independent parameters. The evolution of the mean and 99% CIs for the biomass and the substrate are illustrated in Figure 4.12. Please note that the CIs in all the plots are quantified with considering the uncertainties from all nine parameters. As it can be seen from Figure 4.12, the biomass grows rapidly until its CI approaches the upper boundary to maximize the productivity, while the CI of the substrate remains at its upper boundary at the beginning and decreases to a low value to activate the production phase. However, the result of the RO ignoring parameter dependencies is too conservative. The effect of parameter dependencies is shown in Figure 4.13 for the previous optimized setting, i.e., assuming independent parameters. The shape of the CIs of the biomass and the substrate are quite different from those in Figure 4.12 and do not reach their upper boundaries, which leaves some space for improvement. Therefore, the RO is repeated with considering the parameter dependencies accordingly and show the results in Figure 4.14. As it can be seen, the CIs of both biomass and substrate concentration reach the upper boundaries and are less conservative compared to the results in Figure 4.13. The optimized feeding profile of the substrate for the independent and correlated cases are compared in Figure 4.15a. The substrate for the correlated case is fed with a higher rate and descended a bit earlier than that for the independent scenario. The PDFs of the product concentrations at the final time point shown in Figures 4.13 and 4.14 are compared in more detail in Figure 4.15b. The product concentration is improved considerably as the dashed curve, which represents the correlated parameter case, is a bit narrowed and shifted to higher concentrations.

As mentioned above, the negative CIs of the substrate concentration in all the figures are due to the assumption of symmetric distributions of the states. This also indicates that the CIs might not be accurate, and thus, they are validated by checking the number of constraint violations with 10,000 Monte Carlo simulation for the independent and correlated case, where the corresponding optimal feeding profiles are applied. The results are listed in Table 4.4. As it can be seen from the second row, the violation frequencies are higher than expected, $\varepsilon_{nq} = 1\% = \frac{100}{10{,}000}$, especially for the substrate concentration. Although the violation frequencies might be acceptable for practical applications, the reliability of the RO can be improved by using a smaller tolerance factor as introduced in Section 4.4.1.2. Corresponding results are shown in the third row of Table 4.4. All violation numbers are improved, while the reactor performance regarding the penicillin productivity is slightly decreased.



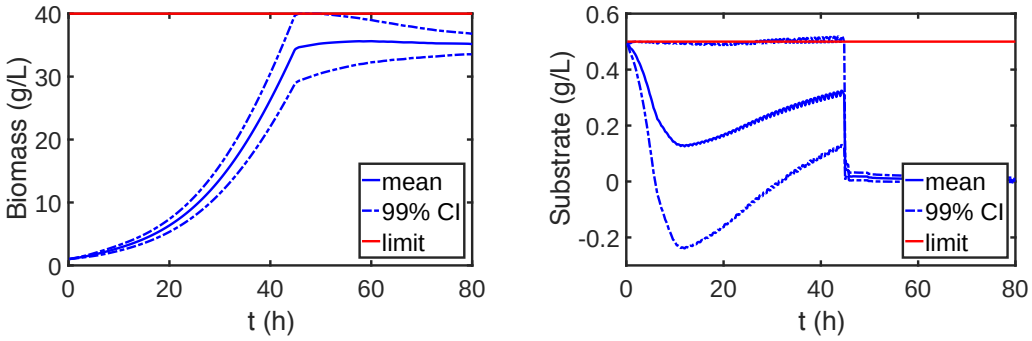
(a) (b)

Figure 4.12: Evolution of the mean and 99% confidence interval (CI) of the (a) biomass and (b) substrate concentrations for the robust design of the fed-batch bioreactor, where the uncertain parameters are independent. The feeding profile from the robust design with independent uncertain parameters is applied.



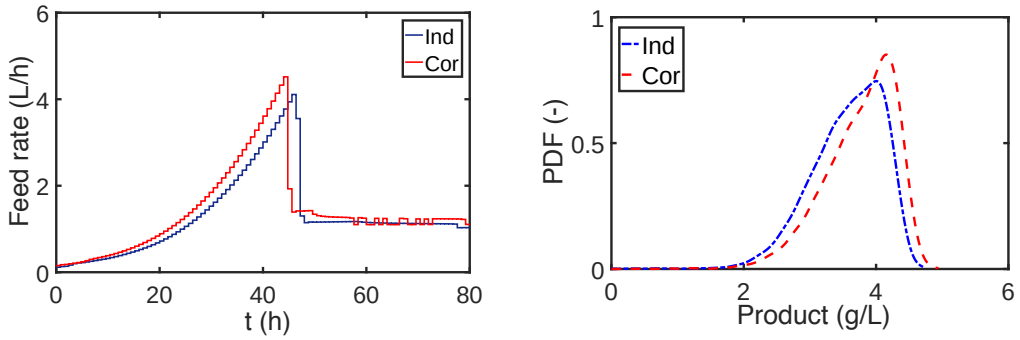
(a) (b)

Figure 4.13: Evolution of the mean and 99% confidence interval (CI) of the (a) biomass and (b) substrate concentrations for the robust design of the fed-batch bioreactor, where the uncertain parameters are correlated. The feeding profile from the robust design with independent uncertain parameters is applied.



(a) (b)

Figure 4.14: Evolution of the mean and 99% confidence interval (CI) of the (a) biomass and (b) substrate concentrations for the robust design of the fed-batch bioreactor, where the uncertain parameters are correlated. The feeding profile from the robust design with correlated uncertain parameters is applied.



(a) (b)

Figure 4.15: Results for the robust design of the fed-batch bioreactor, where the uncertain parameters are either independent or correlated. (a) control sequence for substrate feeding; and (b) final concentration of the product, respectively.

Table 4.4: The number of constraint violations from 10,000 Monte Carlo simulations, where the tolerance factor $\varepsilon_{nq} = 1\%$ and $\varepsilon_{nq} = 0.14\%$ for both designs with independent and correlated parameters. The performance indicates the mean value of the production concentration at the end.

		Independent	Correlated
$\varepsilon_{nq} = 1\%$	X	146	35
	S	572	554
performance		3.63	3.76
$\varepsilon_{nq} = 0.14\%$	X	19	2
	S	378	369
performance		3.53	3.67

4.5 Chapter Summary

In this Chapter, a new framework for solving robust optimization problems using the point estimate method is proposed. Here, a sampling strategy derived from an isoprobabilistic transformation was used to include parameter dependencies and *soft* equality constraints of practical relevance. In parallel, methods including higher-order statistical moments are investigated to approximate robust equality and inequality constraints. To include only the most relevant model parameters and to reduce the computation costs, the global parameter sensitivities are calculated before the robustification step.

Two case studies, which include chemical and biological production processes, were used to demonstrate the performance of the proposed framework. The first case study attempts to maximize the conversion of a reactant while simultaneously satisfying the constraints on the reactor temperature of a tubular reactor. The proposed method addresses the trade-off between performance and robustness for the reactor under parameter uncertainties. An evident influence of parameter correlation on the designed control profile and confidence intervals of the system states is observed. Performances of the second and fourth moment methods for approximating the robust inequality constraints were also examined. The fourth moment method has a more rigorous structure compared to the second moment approach. However, the performance of the fourth moment method is limited by the accuracy of the PEM. Thus, it could be concluded that the second moment method might be more favorable in this particular case. Furthermore, the approximation error could be compensated by using more conservative tolerance values, which resulted in slight deterioration of the reactor performance. To save energy costs, an equality constraint is added to the outlet temperature. The robust equality

constraint had to be relaxed deliberately to be solvable. The process performance deteriorated dramatically with lower relaxation factors. The second example is about optimal design of a bioreactor for the penicillin fermentation process. Global sensitivity analysis was used to determine the relevant parameters and to ease the computational expense of the robustification framework. This is extremely useful for large-scale problems with a high number of uncertain parameters. Moreover, the effect of parameter correlations on the robust process design was also observed. Here, the PEM still performs reasonably well and retains a relatively low computational cost.

In conclusion, the proposed framework provides a comprehensive strategy for robust optimization problems and covers features that have not been considered in previous works from literature. It is able to achieve suitable robust design in the absence and presence of parameter correlations at low computational costs. As discussed, the PEM might fail in estimating higher-order statistical moments, especially for systems with strong nonlinearities. This is also the main reason why the performance of the fourth moment method did not provide the expected improvement in robustification. Alternatively, the accuracy of the PEM can be increased using extended sample-generating rules, i.e., higher sample number results in more precise approximation at the cost of efficiency, or different methods for uncertainty quantification might be studied.

Chapter 5

Robust Process Design with Non-Gaussian Parameter Uncertainties [5]

In this chapter, the framework proposed in Chapter 4 is further implemented for the robust design of a downstream process, i.e., the freeze-drying process (lyophilization), in which non-Gaussian parameters are considered. Moreover, a novel idea of combining Gaussian mixture distributions (GMD) with the PEM is presented as an alternative to address parameter uncertainties with non-Gaussian distributions. The different approaches are explained and compared in detail.

The chapter is organized as follows. The mathematical background of the methods and the structure of the robust process design problem are introduced in Section 5.2. The first-principle model of the freeze-drying process, as well as the assumed parameter uncertainties, is presented in Section 5.3. Results and discussion about the performance of the proposed GMD-PEM algorithm for robust process design are provided in Section 5.4. In Section 5.5, the conclusions are given.

5.1 Motivation

Probability-based approaches have been implemented extensively in various studies to solve robust optimization problems. These approaches provide an explicit description of uncertainties in terms of probability distributions, and thus, probability-based approaches ensure less conservative solutions compared with scenario-based concepts [29, 136, 137]. However, probability-based approaches suffer from high computational costs of the propagation and quantification of uncertainties [138, 30]. The point estimate method (PEM) introduced in Chapter 2 [74, 3] aims to avoid costly sampling and simulation runs by analyzing exclusively statistical moments of low order. In doing so, fewer

[5] Part of this chapter has been published in (Xie et al., Chem. Eng. Sci., 207, 805-819, 2019 [10])

sample points compared with Monte Carlo simulations (MCs) and polynomial chaos expansion (PCE) have to be evaluated in robust process design [93, 139]. Moreover, the PEM is superior than other cubature methods, in terms of approximation accuracy, as concluded by [84]. The original PEM, however, is limited to Gaussian probability distributions [80].

The typical method for calibrating models is regressing the model to the experimental data, in which the Gaussian likelihood function is used to find the most reliable parameter values [6]. Furthermore, the associated confidence region of parameter value is derived with the inverse of the Fisher information matrix (FIM) [6]. However, the mentioned concept and derived parameter uncertainties is reliable only under the assumption of linear relationship between the parameters and observations (measurements) in the model and Gaussian distributed uncertainties [140, 6]. In practically application, these two assumptions are highly likely not consistent with the real process and associated measurement data, which are highly nonlinear and complex [141, 142]. Shi[140] compared estimated parameter uncertainties based on the two methods, which are linear regression method and Bayesian technique. The results point out that the parameter uncertainties could be non-Gaussian distributions [140]. Kalyanaraman[143] also obtained non-Gaussian distributions for parameter uncertainty estimated with experimental data using the Bayesian approach. A similar conclusion is given in [144] where the parameter uncertainties, which are approximated by the bootstrap method, have non-Gaussian distributions. Therefore, it is necessary to consider probability distributions with non-Gaussian distributions for robust process design in the field of pharmaceutical manufacturing [145, 146, 93]. Robust process design that includes parameter uncertainties of non-Gaussian, data-centric probability density functions is still missing in the literature and is the focus of this chapter. To apply the PEM for these non-Gaussian uncertainties, an adapted sampling scheme within a Gaussian-mixture framework is presented for proper uncertainty quantification.

Commonly, not only the non-Gaussian shape of parameter distributions but also parameter correlations are ignored to simplify robust process design problems [30, 84]. This assumption, in turn, leads to an inevitable loss of information and might result in sub-optimal process designs. Alternatively, in the case of the PEM, non-Gaussian distributions including parameter correlations can be incorporated via a transformation step, as presented in Chapter 2. In the last chapter, the Nataf transformation is used to map the original PEM samples to the samples for parameters associated with multivariate Gaussian distribution and parameter dependencies. The transformation step, however, leads to additional complexity in the approximation which might reduce the accuracy of the PEM. Moreover, the Nataf transformation is capable only of describing

linear correlations between the parameters. In the current chapter, Gaussian mixture distributions (GMD) is propose to represent non-Gaussian and correlated parameter uncertainties with the weighted sum of a limited number of Gaussian distributions. The advantage of using GMD concept is that it does not introduce extra complexity in the approximation with PEM and is available for nonlinear parameter correlation problems.

GMD has been extensively applied in the field of pattern recognition and machine learning to cluster data into subgroups [147]. Rossner[145] came up with the idea of using GMD to decompose a one-dimensional Gaussian distribution into several component distributions to approximate the uncertainties in the model output more accurately even if the resulting model uncertainties are non-Gaussian. Technically, a least square estimation approach with additional constraints on the width of the Gaussian mixture components was used to determine the weight factors of the GMD. Here, GMD is also used for proper uncertainty quantification but follow a different philosophy. In this chapter, first, the multivariate distribution is decomposed directly instead of doing that individually for each marginal distribution which alleviates the computational costs. Second, the expectation-maximization (EM) algorithm is implemented to cluster the data and to estimate the weight factors of the GMD. Note that for the traditional least square estimation as suggested in [145], the optimal weight factors are difficult to obtain, especially for high-dimensional parameter problems [148]. With this novel GMD-PEM algorithm for robust process design, the shapes of the non-Gaussian parameter distributions and the respective model output distributions can be captured realistically.

The novel GMD-PEM algorithm in robust process design is motivated and applied to the freeze-drying process. The freeze-drying process provide the solution for preservation and transportation of medicines[149, 150]. However, comprehensive and sophisticated design have to be investigated to minimize the degradation effects during the drying steps and ensure cake integrity [151]. Moreover, the current status of freeze-drying process is time-consuming and noncontinuous, which requires further design focus on maximizing the efficiency [151]. Here, the primary drying step, in which most water is removed from the product, is the most time-consuming and failure-prone step [151]. As APIs used for freeze-drying are highly valuable products, the model-based design of the primary drying step may help to ensure product quality at a competitive cost. A mathematical model for primary drying was investigated in [151, 152] to mimic the mass and energy transfers during the freeze-drying process. [70] adapted the model for the first steps in model-based process design. A grid-based approach was implemented to increase the efficiency of the primary drying step while guaranteeing the CQAs of the dried product [48]. Moreover, Monte Carlo simulations were also used to quantify the effect of parameter

uncertainties which might result from measurement noise and model simplifications [70]. However, optimization with the Monte Carlo simulations at individual time points might lead to a sub-optimal solution and is computationally expensive by definition, especially in combination with Monte Carlo simulations for uncertainty quantification.

Here, the aim is to optimize the freeze-drying process under uncertain model parameters with advanced process system engineer techniques. Uncertainties with non-Gaussian distributions are assumed for two critical process parameters: the mass transfer resistance coefficient of the product and the heat transfer coefficient of the vial and the product [153, 70]. The non-Gaussian distributed parameters are taken into account in the freeze-drying simulation, and the resulting uncertainties in the sublimation mass and the temperature at the sublimation interface are quantified as illustrated in Fig. 5.1. The novel GMD-PEM algorithm is implemented to efficiently describe and quantify the non-Gaussian uncertainties in the model parameters and outputs, and these uncertainties are integrated into the robust process design to ensure product quality standards and process efficiency simultaneously. Moreover, the superiority of the novel GMD-PEM algorithm is demonstrated by comparing the derived robust process design results from different approaches.

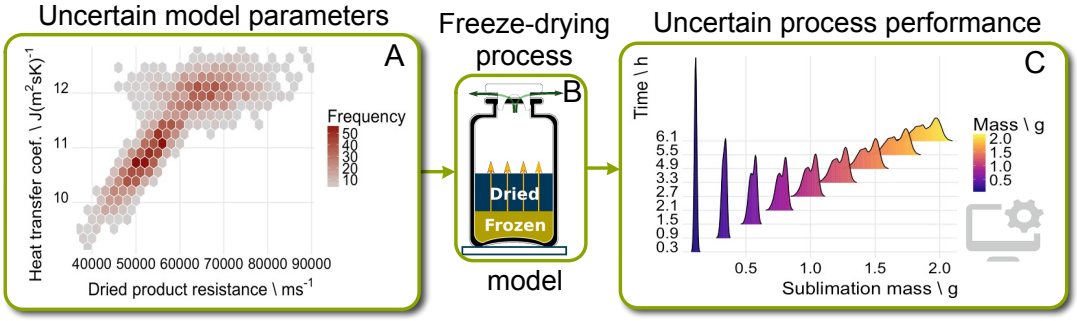


Figure 5.1: Illustration of uncertain model parameters in the freeze-drying process: Two correlated model parameters with a non-Gaussian distribution, see the scatter plot **(A)**, of the freeze-drying process **(B)** lead to uncertain non-Gaussian simulation results, e.g., sublimation mass distributions **(C)**.

This figure is originated from [10] and partially created by René Schenkendorf

5.2 Methodology

In this section, the basics and the mathematical formulations of the GMD-PEM algorithm for robust process design is described. First, it starts with a brief review of the PEM, the single Gaussian approach, and the nonlinear-transfer approach. Second, the EM algorithm is introduced to calibrate the GMD-PEM algorithm, i.e., to iteratively determine the maximum likelihood of the parameters even when the actual structure of the Gaussian mixture distribution is unknown.

Robust process design requires statistical information regarding the quantities of interest, e.g., yield, conversion and costs, to ensure meaningful designs based on simulation studies. To this end, the parameter uncertainties that are propagated through the model have to be transferred to the simulation results, and the resulting statistics are then quantified. Two options are available to calculate the needed statistical information. One option is to estimate the probability distribution directly with Monte Carlo simulations and kernel density estimators [78], which requires a vast number of samples to cover the relevant parameter space [119]. The second option is to calculate statistical moments instead, e.g., mean, variance, skewness, and kurtosis, and to parameterize the probability distributions with these statistical moments [125]. As the statistical moments can be approximated with the PEM [80, 74, 3], the computational cost is low compared to the direct approximation of the resulting probability density functions. The basics of the PEM are presented in Chapter 2.

Assuming an n_ξ-dimensional random parameter vector $\boldsymbol{\xi} \in I_{\boldsymbol{\xi}} \subset \mathbb{R}^{n_\xi}$, the corresponding multivariate normal distribution with the joint density function $\boldsymbol{f(\xi)}$ reads as:

$$\boldsymbol{f(\xi)} = \frac{1}{\sqrt{(2\pi)^{n_\xi}|\Sigma|}} exp(-\frac{1}{2}(\boldsymbol{\xi} - E(\boldsymbol{\xi}))^T \Sigma^{-1}(\boldsymbol{\xi} - E(\boldsymbol{\xi}))), \tag{5.1}$$

where $E(\boldsymbol{\xi})$ and Σ are the vector of the mean values and the covariance matrix, respectively. Moreover, a nonlinear function $\boldsymbol{g(\xi)} : \mathbb{R}^{(n_\xi)} \rightarrow \mathbb{R}^{n_x}$ is used to represent the mathematical process model. The process model maps the model parameters to the model outputs; i.e., $\boldsymbol{y} = \boldsymbol{g(\xi)}$. The n^{th} statistical moment of $\boldsymbol{y}$ is given as:

$$\boldsymbol{\mu}_1 = \textstyle\int_{I_{\boldsymbol{\xi}}} \boldsymbol{y f(\xi)} d\boldsymbol{\xi}, \tag{5.2}$$

$$\boldsymbol{\mu}_n = \textstyle\int_{I_{\boldsymbol{\xi}}} (\boldsymbol{y} - \boldsymbol{\mu}_1)^n \boldsymbol{f(\xi)} d\boldsymbol{\xi} \qquad n \geq 2. \tag{5.3}$$

According to the PEM presented in Chapter 2, the integral terms in Eqs. (5.2) and (5.3) can be approximated with the weighted sum of the model simulations which are

evaluated at deterministic parameter sample points:

$$\boldsymbol{\mu}_1 \approx w_0 \boldsymbol{g}(GF[0]) + w_1 \sum \boldsymbol{g}(GF[\pm\vartheta]) + w_2 \sum \boldsymbol{g}(GF[\pm\vartheta, \pm\vartheta]), \tag{5.4}$$

$$\begin{aligned} \boldsymbol{\mu}_n \approx\ & w_0(\boldsymbol{g}(GF[0]) - \boldsymbol{\mu}_1)^n + w_1 \sum (\boldsymbol{g}(GF[\pm\vartheta]) - \boldsymbol{\mu}_1)^n + \\ & w_2 \sum (\boldsymbol{g}(GF[\pm\vartheta, \pm\vartheta]) - \boldsymbol{\mu}_1)^n, \end{aligned} \tag{5.5}$$

where $\vartheta = \sqrt{3}, w_0 = 1 + \frac{n_\xi^2 - 7n_\xi}{18}, w_1 = \frac{4-n_\xi}{18}, w_2 = \frac{1}{36}$.

As pointed out in Chapter 2, the PEM can calculate integral terms of monomials only up to order of 5 accurately. In other words, the complexity of function $\boldsymbol{g}(\boldsymbol{\xi})$, as well as the order of the calculated statistical moments, determines the accuracy of Eqs. (5.4) and (5.5) when the original PEM is used. According to the results from Chapter 4, satisfactory estimations of the first- and second-order moments can be found, while higher-order moments with $n \geq 3$ might be beyond the capability of the PEM. Note that the approximation is available only if $\boldsymbol{\xi}$ follows a standard multivariate normal distribution; i.e., $E(\boldsymbol{\xi}) = \mathbf{0}$ and $\Sigma = I$ [80, 3]. Thus, for non-Gaussian parameter uncertainties, The PEM has to be modified.

5.2.1 Non-Gaussian Parameter Uncertainties

5.2.1.1 Methods from Literature and Previous Chapter

As mentioned in [145], the distribution of parameter uncertainties estimated with experimental data from various resources might have arbitrary shapes; i.e., the parameter uncertainties cannot be described properly by Gaussian density functions. Thus, Eqs. (5.4) and (5.5) are ill-posed and have to be refined. To employ the PEM for non-Gaussian parameter distributions, the deterministic sample points have to be modified as illustrated in Fig. 5.2. Here, it is assumed that $\boldsymbol{\theta} \in I_{\boldsymbol{\theta}} \subset \mathbb{R}^{n_\theta}$ is the vector of the parameters with non-Gaussian uncertainties, and it aims to estimate the mean and variance of $\boldsymbol{g}(\boldsymbol{\theta})$ which could be the cost function or the constraint function for the process design. The first idea, which is what is called the single Gaussian approach in the left column, is to approximate the arbitrary distribution with a multivariate normal distribution of which the mean $E(\boldsymbol{\theta})$ and variance $\Sigma(\boldsymbol{\theta})$ are estimated according to:

$$E(\boldsymbol{\theta}) = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\theta}_i, \tag{5.6}$$

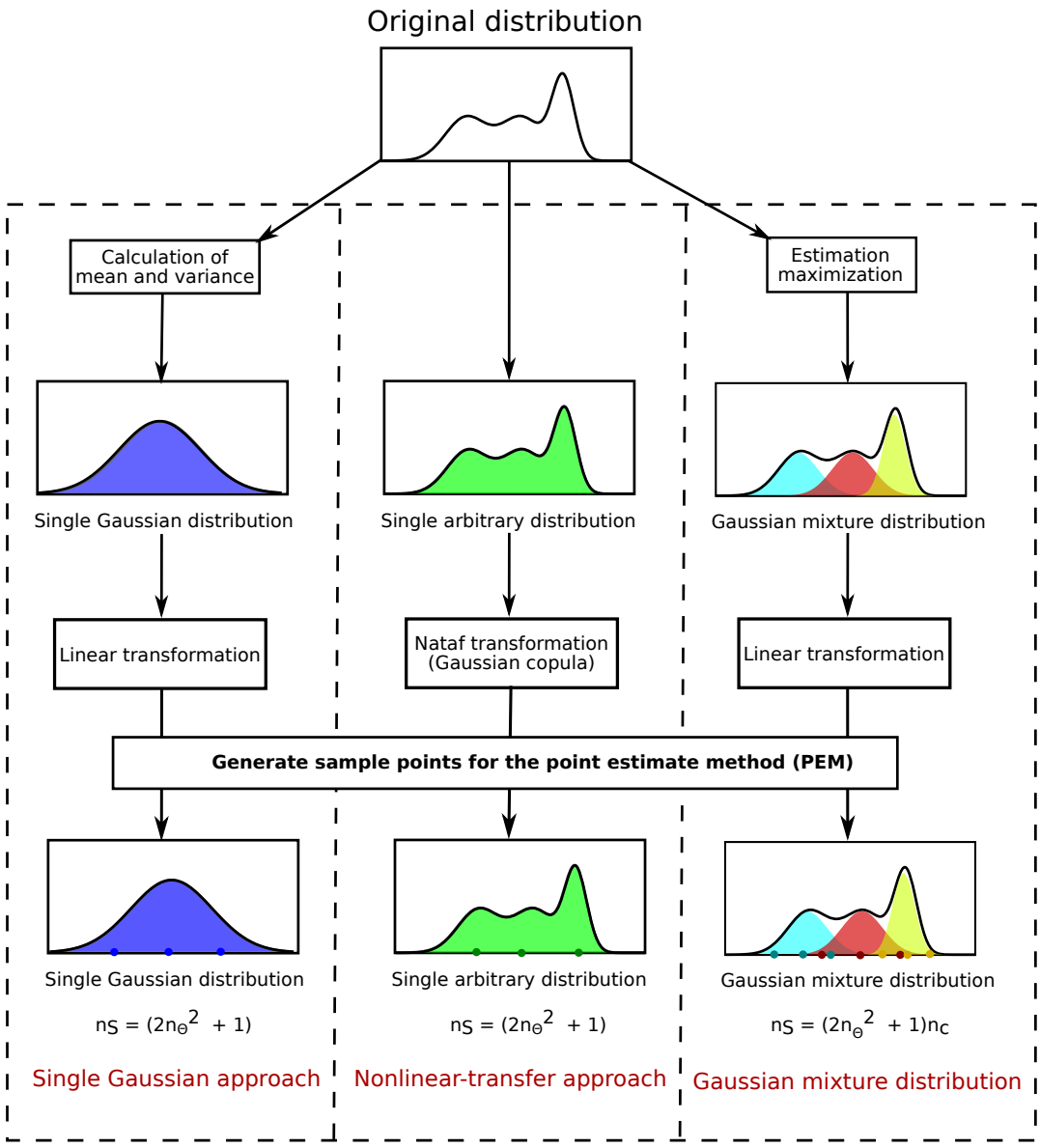


Figure 5.2: A schematic diagram for the single Gaussian approach (section 2.2), the nonlinear-transfer approach (section 4.2), and the Gaussian mixture distribution (GMD) with an example of a one-dimensional probability distribution. n_S and n_θ are the number of sample points for the PEM and the number of parameters. n_c is the number of component mixtures for the GMD.

$$\Sigma(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{\theta}_i - E(\boldsymbol{\theta}))(\boldsymbol{\theta}_i - E(\boldsymbol{\theta}))^T. \tag{5.7}$$

in which N is the number of parameter samples. Note that the off-diagonal elements $\Sigma(\boldsymbol{\theta})$ are typically neglected and set to zero; i.e., the parameter uncertainties are considered a multivariate Gaussian distribution without parameter correlations. The original PEM samples generated by function $GF[\cdot]$ are mapped onto the the multivariate normal distribution described by $E(\boldsymbol{\theta})$ and $\Sigma(\boldsymbol{\theta})$ with the linear transformation given below:

$$\boldsymbol{\theta}_i^{PEM} = A\boldsymbol{\xi}_i^{PEM} + E(\boldsymbol{\theta}), \qquad i = 1 \dots 2n_\theta^2 + 1, \tag{5.8}$$

where A is the lower triangle matrix from the Cholesky decomposition of $\Sigma(\boldsymbol{\theta})$. According to Eq. (2.16), the mean and variance of $\boldsymbol{g}(\boldsymbol{\theta})$ can be derived with the transformed samples based on Eqs. (5.4) and (5.5):

$$\boldsymbol{\mu}_1(\boldsymbol{g}(\boldsymbol{\theta})) \approx \sum_{i=1}^{2n_\theta^2+1} w_i \boldsymbol{g}(\theta_i^{PEM}), \tag{5.9}$$

$$\boldsymbol{\mu}_2(\boldsymbol{g}(\boldsymbol{\theta})) \approx \sum_{i=1}^{2n_\theta^2+1} w_i (\boldsymbol{g}(\theta_i^{PEM}) - \boldsymbol{\mu}_1(\boldsymbol{g}(\boldsymbol{\theta})))^2. \tag{5.10}$$

For the sake of simplicity, in the robust process design literature, it is commonly assumed that the parameter uncertainties follow a Gaussian distribution and are independent [154, 30, 73, 155]. However, this assumption leads to a certain loss of information regarding parameter uncertainties and includes extra deviations in the approximations of Eqs. (5.9) and (5.10).

The second approach is one presented in Chapter 2 and implemented in Chapter 4 for handling parameter dependencies, the so-called nonlinear-transfer approach (NTA). It is illustrated in the middle column of Fig. 5.2. Instead of using a Gaussian distribution, the PEM samples are directly mapped to the actual parameter distribution. To this end, the iso-probabilistic transformation given in Proposition 2.1 is used:

$$\boldsymbol{\theta}_i^{PEM} = F_{\boldsymbol{\theta}}^{-1}(F_\xi(\boldsymbol{\xi}_i^{PEM})). \tag{5.11}$$

In Eq. (5.11), $F_{\boldsymbol{\theta}}$ and $F_{\boldsymbol{\xi}}$ are the joint cumulative density function (CDF) for parameters θ and standard Gaussian random parameters $\boldsymbol{\xi}$. The joint CDF $F_{\boldsymbol{\theta}}$ for multivariate distributions are usually complex and implicit, especially when the distributions are non-Gaussian and correlated. Moreover, it is also challenging to map directly from one

parameter space to the other with Eq. (5.11) due to the complexity of $F_{\boldsymbol{\theta}}$. Therefore, $F_{\boldsymbol{\theta}}$ is approximated using Gaussian copulas with marginal CDFs F_{n_θ} for individual parameters and the correlation matrix Σ_ρ, as shown in Section 3.2:

$$F_{\boldsymbol{\theta}} = F_n[F^{-1}(F_1), \cdots, F^{-1}(F_{n_\theta}); \Sigma_\rho]. \tag{5.12}$$

Furthermore, the inverse Nataf transformation, which is an alternative numerical algorithm based on Eqs. (5.11) and (5.12) and has been presented in Algorithm 2.1, is used to transform the PEM samples of a standard multivariate Gaussian distribution to the samples of the target distribution of $\boldsymbol{\theta}$. With the modified samples, Eqs. (5.9) and (5.10) can be used to calculate the mean and the variance of the model outputs.

This method retains almost all of the parameter distribution information in the modified PEM sample points. However, the iso-probabilistic transformation might be highly nonlinear and induces extra complexity in the simulation results $\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{g}(F_{\boldsymbol{\theta}}^{-1}(F_\xi(\boldsymbol{\xi})))$. The additional complexity from the transformation step might deteriorate the approximation accuracy of PEM, and a such lead to sub-optimal robust process designs.

5.2.1.2 Gaussian Mixture Distributions (GMD)

GMD are an essential part of the proposed algorithm for robust process design. Unlike the two previous approaches where either a non-Gaussian distribution is simplified into an independent Gaussian distribution or the original PEM samples are transformed to a non-Gaussian distribution with nonlinear transformation, the GMD concept represents the non-Gaussian distribution with the superposition of a limited number of Gaussian distributions. The PEM samples are then mapped to these Gaussian distributions with a linear transformation step, as illustrated in the right column in Fig. 5.2.

The GMD is structured as

$$\boldsymbol{\theta} \sim \sum_{j=1}^{N_C} \omega_j \mathcal{N}(E_j, \Sigma_j), \tag{5.13}$$

where $\mathcal{N}(E_j, \Sigma_j)$ means the Gaussian distribution with mean E_j and covariance matrix Σ_j for the jth component. Here, ω_j is the non-negative weight for ith component with $\sum_{j=1}^{N_C} \omega_j = 1$. N_C is the total number of Gaussian distributions. The probability density

function (PDF) of $\boldsymbol{\theta}$ is equal to:

$$\boldsymbol{f}(\boldsymbol{\theta}) = \sum_{j=1}^{N_C} \omega_j \boldsymbol{f}_j(\boldsymbol{\theta}), \tag{5.14}$$

where $\boldsymbol{f}_j(\boldsymbol{\theta})$ is the PDF of the j^{th} component distribution given as

$$\boldsymbol{f}_j(\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^{n_\theta}\left|\Sigma_j\right|}} exp(-\frac{1}{2}(\boldsymbol{\theta} - E_j)^T {\Sigma_j}^{-1}(\boldsymbol{\theta} - E_j)). \tag{5.15}$$

Note that the sum of Gaussian distributed random variables and the weighted sum of the Gaussian distributions of random variables are different concepts, i.e., the sum of the random variables results in a Gaussian distribution, whereas the weighted sum of the Gaussian distributions can represent non-Gaussian distributions needed for robust process design. With the approximation in Eq. (5.13), the original PEM samples are mapped individually to the component distributions with the linear transformation given in Eq. (5.8). Here, $N_C \times (2n_\theta^2 + 1)$ samples are obtained as:

$$\boldsymbol{\theta}_{ji}^{PEM} = A_j \boldsymbol{\xi}_i^{PEM} + E_j, \qquad i = 1 \dots 2n_\theta^2 + 1, \quad j = 1 \dots N_C, \tag{5.16}$$

where A_j is the lower triangle matrix from the Cholesky decomposition of the full covariance matrix Σ_j. To determine the samples with Eq. (5.16), it still needs information about the structure of the GMD in Eq. (5.13). In other words, the values for ω_j, E_j, Σ_j and N_C have to be estimated.

Assuming $\boldsymbol{z}$ are the realizations from a non-Gaussian distribution. Generally, the unknown parameters ω_j, E_j, and Σ_j can be determined by maximizing the marginal likelihood function of the unknown parameters given realizations $\boldsymbol{z}$ [156, 145]. However, it is challenging to numerically maximize the marginal likelihood function here, because of the existence of latent variables [147], which are in this case, the distribution group, to which the individual sample belongs. Note that the latent variables are discretized and are assigned with the index of component distributions from which the realizations stem. Alternatively, the EM approach, which is commonly used to maximize the marginal likelihood function of parameters in cases where latent variables exist [147], is implemented. The EM approach includes two steps:

- The expectation step: the conditional distribution of the latent variables is determined with the parameter values; i.e., ω_j, E_j, and Σ_j obtained from the last iteration.

- The maximization step: the parameter values are updated with the results from maximizing the expected value of the log-likelihood function for the conditional distribution of the latent variables determined in the expectation step.

The expectation and maximization steps are iterated until the terminal condition is fulfilled. For more details regarding the derivation, the proof, and the mathematical equations of the EM approach, the interested reader is referred to [147] and [157]. Moreover, to identify a meaningful number of Gaussian distributions N_C to build the GMD, the Bayesian information criterion (BIC) is applied to determine the optimum number of component distributions [158, 159].

Unlike the single Gaussian distribution and the nonlinear-transfer approach, the mean and the variance of $\boldsymbol{g}(\boldsymbol{\theta})$ are derived with the weighted sum of the mean and the variance of $\boldsymbol{g}(\boldsymbol{\theta})$ on the component distributions and are formulated as:

$$\boldsymbol{\mu}_1(\boldsymbol{g}(\boldsymbol{\theta})) \approx \sum_{j=1}^{N_C} \omega_j \boldsymbol{\mu}_{1j}(\boldsymbol{g}(\boldsymbol{\theta})), \tag{5.17}$$

$$\boldsymbol{\mu}_2(\boldsymbol{g}(\boldsymbol{\theta})) \approx \left(\sum_{j=1}^{N_C} \omega_j (\boldsymbol{\mu}_{2j}(\boldsymbol{g}(\boldsymbol{\theta})) + \boldsymbol{\mu}_{1j}(\boldsymbol{g}(\boldsymbol{\theta}))^2) \right) - \boldsymbol{\mu}_1(\boldsymbol{g}(\boldsymbol{\theta}))^2, \tag{5.18}$$

where the mean and the variance of $\boldsymbol{g}(\boldsymbol{\theta})$ of the component distributions are calculated as:

$$\boldsymbol{\mu}_{1j}(\boldsymbol{g}(\boldsymbol{\theta})) \approx \sum_{i=1}^{2n_\theta^2+1} w_i \boldsymbol{g}(\theta_{ji}^{PEM}), \tag{5.19}$$

$$\boldsymbol{\mu}_{2j}(\boldsymbol{g}(\boldsymbol{\theta})) \approx \sum_{i=1}^{2n_\theta^2+1} w_i (\boldsymbol{g}(\theta_{ji}^{PEM}) - \boldsymbol{\mu}_{1j}(\boldsymbol{g}(\boldsymbol{\theta})))^2. \tag{5.20}$$

In summary, the proposed GMD-PEM approach retains most of the information of a non-Gaussian parameter distribution and does not introduce additional nonlinearities from the transformation step in Eq. (5.16) to Eqs. (5.17) to (5.20). Note that the computational cost of GMD-PEM approach increases proportionally to the number of component distributions, and thus, the BIC is applied to ensure that only a low number of component distributions are used. In the next section, the structure of a robust process design with GMD-PEM approach is presented.

5.2.2 Robust Process Design with Non-Gaussian Uncertainties

The general structure of the probability-based robust process design is given in Eq. 4.1. The structure of the robust process design where the single Gaussian approach and the nonlinear-transfer approach are implemented could be easily adapted from the formulations in Section 4.3. The structure of the GMD-PEM algorithm for robust process design is shown in Eq. 5.21. The proposed GMD-PEM algorithm provides meaningful results not only for non-Gaussian parameter uncertainties but also for non-Gaussian simulation results. The failure probability of the inequality constraints is approximated with the weighted sum of the failure probability results of the individual component distributions as shown in Eq. (5.21f). The second moment method is used, as it outperforms the fourth moment method when PEM is used to approximate the statistical moments. Moreover, the formulation of robust equality constraints is actually similar to that of robust inequality constraints, as they have to be relaxed to circumvent the ill-posed structure. Therefore, they are not shown in Eqs. 5.21.

Note that when the SGA or NTA is implemented, the approximation is accurate only if the probability distribution of the constraints is Gaussian. Thus, the SGA- and NTA-based approaches might lose essential information about the model output uncertainties in the case of non-Gaussian distributions, which might be the result of non-Gaussian model parameter uncertainties or due to model nonlinearities. The performance of the different algorithms and their accuracy are compared with the case study in the next section.

$$\min_{\mathbf{x}(\cdot),\mathbf{u}(\cdot)} \ \boldsymbol{\mu}_1(M(\mathbf{x}_{t_f})) + \alpha\boldsymbol{\mu}_2(M(\mathbf{x}_{t_f}))^{0.5}, \tag{5.21a}$$

subject to:

$$i = 1, \ldots, 2n_\theta^2 + 1 \quad j = 1, \ldots, N_C, \tag{5.21b}$$

$$\boldsymbol{\theta}_{ji} = [\mathbf{p}_{ji}, \mathbf{x}_{0,ji}]^T, \mathbf{x}_{ji} = [\mathbf{x}_{\mathbf{d},ji}, \mathbf{x}_{\mathbf{a},ji}]^T, \mathbf{x}_{\mathbf{d},ji}(0) = \mathbf{x}_{0,ji}, \mathbf{x}_{t_f,ji} = \mathbf{x}_{ji}(t_f), \tag{5.21c}$$

$$\dot{\mathbf{x}}_{\mathbf{d},ji}(t) = \mathbf{g}_{\mathbf{d}}(\mathbf{x}_{ji}(t), \mathbf{u}(t), \mathbf{p}_j i), \quad \mathbf{0} = \mathbf{g}_{\mathbf{a}}(\mathbf{x}_{ji}(t), \mathbf{u}(t), \mathbf{p}_{ji}), \tag{5.21d}$$

$$\overline{\boldsymbol{h}}_{ji} = -\mathbf{h}_{\mathbf{nq}}(\mathbf{x}_{ji}(t), \mathbf{u}(t), \mathbf{p}_{ji}) \tag{5.21e}$$

$$\sum_{j=1}^{N_C} \omega_j F\left(-\frac{\boldsymbol{\mu}_{1j}(\overline{\boldsymbol{h}})}{\boldsymbol{\mu}_{2j}(\overline{\boldsymbol{h}})^{0.5}}\right) \leq \varepsilon_{nq}, \tag{5.21f}$$

$$\boldsymbol{\mu}_1(M(\mathbf{x}_{t_f})) = \sum_{j=1}^{N_C} \omega_j \boldsymbol{\mu}_{1j}(M(\mathbf{x}_{t_f})), \tag{5.21g}$$

$$\boldsymbol{\mu}_2(M(\mathbf{x}_{t_f})) = \left(\sum_{j=1}^{N_C} \omega_j (\boldsymbol{\mu}_{2j}(M(\mathbf{x}_{t_f})) + \boldsymbol{\mu}_{1j}(M(\mathbf{x}_{t_f}))^2) \right) - \boldsymbol{\mu}_1(M(\mathbf{x}_{t_f}))^2, \quad (5.21h)$$

$$\boldsymbol{\mu}_{1j}(M(\mathbf{x}_{t_f})), \boldsymbol{\mu}_{2j}(M(\mathbf{x}_{t_f})), \boldsymbol{\mu}_{1j}(\overline{\boldsymbol{h}}), \boldsymbol{\mu}_{2j}(\overline{\boldsymbol{h}}) \text{are calculated with Eqs. (5.19) and (5.20)} \quad (5.21i)$$

$$\mathbf{u}_{min} \leq \mathbf{u} \leq \mathbf{u}_{max}, \quad (5.21j)$$

5.3 Case Study: the Freeze-Drying Process

To demonstrate the performance of the GMD-PEM algorithm, the freeze-drying process is used as a case study. First principle model of the freeze-drying process is introduced in this section.

The freeze-drying process, also known as lyophilization, is a solution for the storage and transportation of APIs, e.g, vaccines [149]. For these APIs, the traditional methods of preservation in aqueous solutions are not applicable for long term purpose [150]. However, shortages of such process are the high energy consumption, high time cost and critical operation conditions [152]. Freezing, primary drying and secondary drying are the major steps for a batch chamber based freeze-drying process [152]. The schematic diagram of the freeze-drying process is illustrated in Fig. 5.3. The primary drying step is recognized as the most time-consuming and error-prone step [151]. Thus, in this study, the robust design of the primary drying step is considered, in which parameter uncertainties are taken into account and the risk level of defective dried APIs is minimized.

Typically, during the primary drying step, the shelf temperature is increased, and the chamber pressure is decreased to a certain level so that sublimation of the water in the vials is initiated and continued at the sublimation surface, which moves downward during the drying step. The water vapor is transferred from each vial to the condenser chamber and is discharged via the vacuum pump to keep the pressure in the chamber at a specific level. Assuming the mass and energy transfer occur in the vials is identical, the model of the primary drying step is based on a single vial as shown in Fig. 5.3. The mathematical model for the primary drying in a single vial is adapted from [70].

The mass transfer equation [152], which describes the dynamics of the sublimation process at the sublimation surface, is given as:

$$\frac{\mathrm{d}m_{sub}}{\mathrm{d}t} = A_p \frac{P_i - P_c}{R_p}, \quad (5.22)$$

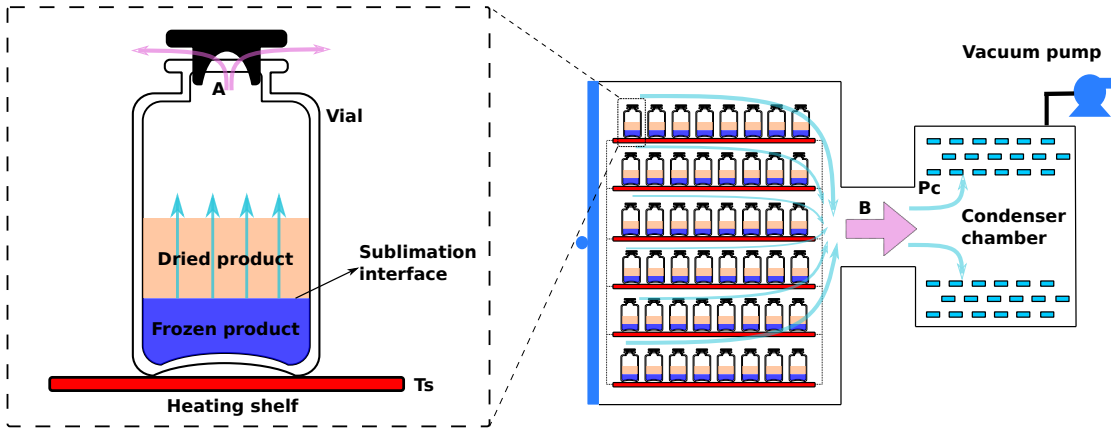


Figure 5.3: Schematic diagram of a lyophilizer where the temperature of the heating shelf (T_s) and the pressure of the chamber (P_c) are optimized to achieve maximum performance. On the left side is the enlarged view of the vial with the frozen and dried API. The light blue and red arrows indicate the flow of the vapor. A and B are the locations where a choked flow might exist.

where A_p is the cross-sectional area of the product, P_c is the chamber pressure, and R_p is the dried product resistance to the vapor flux. P_i is the vapor pressure at the sublimation interface which can be calculated according to the equation in [160]:

$$P_i = \exp(9.55 - \frac{5720}{T_i} + 3.53\ln(T_i) - 0.00728T_i). \tag{5.23}$$

T_i is the temperature at the sublimation interface and is calculated with the energy balance equation given in [152]. Here, the heat used for sublimation is assumed to be equal to the heat transferred from the heating shelf:

$$K_v(T_s - T_B)A_v = \Delta H_s \frac{\mathrm{d}m_{sub}}{\mathrm{d}t}, \tag{5.24}$$

where K_v, A_v, and T_s denote the heat transfer coefficient, the outer cross-sectional area of the vial, and the shelf temperature, respectively. ΔH_s is the specific heat of sublimation as given in [160] and reads as:

$$\Delta H_s = 4.68 \times 10^4 + 35.9T_i - 0.0741T_i^2 + 542\exp(-(\frac{T_i}{124})^2). \tag{5.25}$$

T_B is the temperature at the bottom of vial and is equal to

$$T_B = T_i + \Delta T, \tag{5.26}$$

where ΔT is the temperature difference across the frozen layer and is calculated with the following equation given by [70]:

$$\Delta T = \frac{889200\frac{(L_f(P_i - P_c))}{R_p} - 0.0102 L_f (T_s - T_i)}{1 - 0.0102 L_f}. \tag{5.27}$$

Here, L_f is the height of the frozen layer and has the following relation with m_{sub}:

$$m_{sub} = (L_{total} - L_f)\rho_I \epsilon A_p, \tag{5.28}$$

where L_{total}, ρ_I, and ϵ are the total height of the product layer, the density of the ice, and the volume of the ice fraction, respectively. The model parameters and the size of the vials, are taken from [70]. Note that the value of R_p and K_v in Table 5.1 are the averages of their values at different conditions. More details regarding the determination and the structure of coefficients R_p and K_v can be found in [161] and [70]. Since these two coefficients are estimated from experiments, they also suffer from the existence of measurement noise and inaccuracy in the parameter estimation strategy. Therefore, uncertainties in these two parameters have to been considered for the design of primary-drying step.

The mathematical model is used in this case study to maximize the efficiency of the primary drying step under parameter uncertainties in R_p and K_v while ensuring the product quality at the same time. Thus, the objective function is to maximize the total mass of the ice removed by sublimation per operating time, i.e., minimize the operation time. To avoid irreversible product damage and have an acceptable API cake appearance, the production temperature at the sublimation interface should be carefully maintained below the critical collapse temperature T_c, which in this case is assumed to be equal to the glass transition temperature T_g [70]. Additionally, an upper boundary is given for the sublimation rate $\frac{\mathrm{d}m_{sub}}{\mathrm{d}t}$, which is due to the choke flow phenomenon at the vial neck as explained in [162]. The choked sublimation rate is calculated using the following equation as in [70]:

$$\dot{m}_{sub,choke,vial} = \frac{0.3 A_{v,n}\sqrt{\frac{kT_r R}{M}}M}{RT_r} P_c, \tag{5.29}$$

Table 5.1: Nominal values of the model parameters and the initial conditions for the primary drying model.

Parameters	Unit	Nominal Value
A_p	m^2	3.80×10^{-4}
A_v	m^2	4.15×10^{-4}
$A_{v,n}$	m^2	1.25×10^{-4}
R_p	m/s	5.57×10^4
K_v	J/(m^2sK)	11.47
L_{total}	m	0.00658
ρ_I	kg/m^3	919
ϵ	–	0.97
M	kg/mol	0.018
k	–	1.33
R	J/(Kmol)	8.314
T_g	°C	-34

where $A_{v,n}$ is the cross area at the vial neck, the equation inside the square root operation is used for the calculation of the speed of sound in an idea gas, T_r is the absolute temperature at the vial neck. Chamber pressure P_c and shelf temperature T_s are manipulated within the range from [5 Pa 30 Pa] and [-40 °C 30 °C], respectively. Note that the lower boundary of P_c is normally not set below 5 Pa, as a very low chamber pressure may have problems with product contamination and heterogeneous heat transfer [153]. The uncertainties of R_p and K_v are assumed to follow a non-Gaussian distribution and is characterized by the samples in the scatter plot in Fig. 5.4. The result structure of the nominal optimization of the primary drying process is given in A.2. Moreover, the structure of the robust optimization of the primary freeze-drying step can be straightforwardly derived with Eq. 5.21 and equations in Section 4.3 for the different approaches, respectively.

The case study is coded in MATLAB®. The robust process design was solved with the simultaneous approach [130] which was implemented in the symbolic framework CasADi for numerical optimization [131] using the NLP solver IPOPT [132] and the MA57 linear solver [133]. The EM algorithm is initialized by k-means [163].

5.4 Results and Discussion

This section discusses the robust process design results of the primary freeze-drying step. First, the results for the nominal process design are given. Next, the adverse effect of

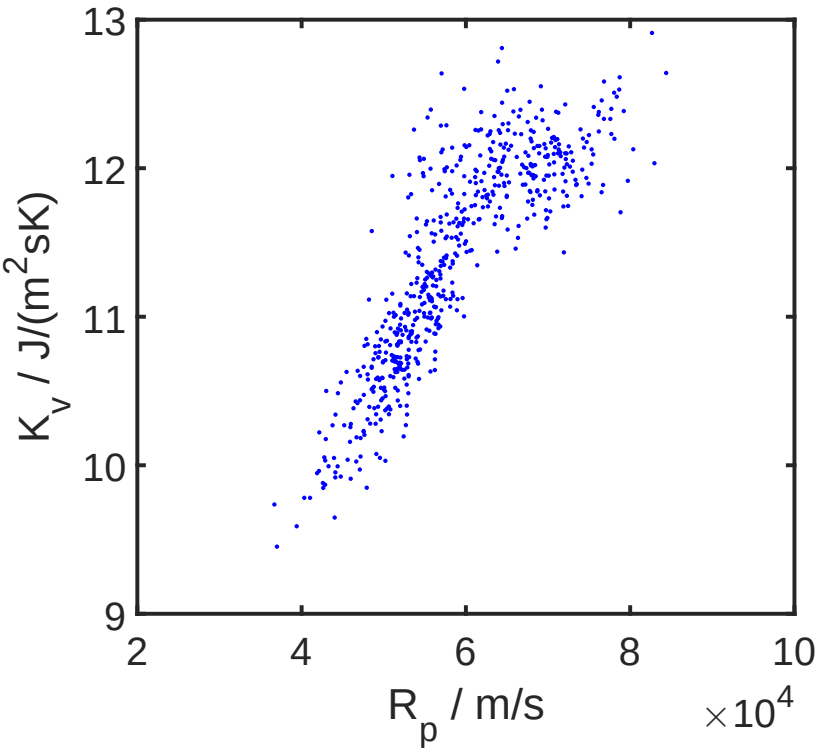


Figure 5.4: Scatter plot of the sample points for two uncertain parameters R_p and K_v.

the parameter uncertainties on the nominal process design is presented. To this end, the accuracy of approximated statistical moments and distributions are analyzed, and the results from the robust process design for the primary freeze-drying step are compared and discussed for the different approaches presented in the previous sections.

5.4.1 Results for the Nominal Process Design

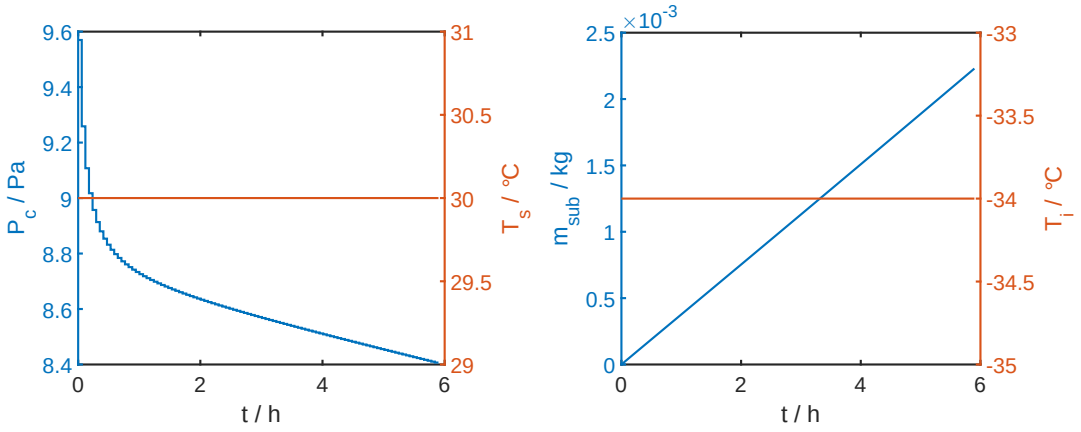


Figure 5.5: Control profile of chamber pressure P_c and shelf temperature T_s (a), and evolution profile of the mass of the ice removed by sublimation and the temperature at the sublimation interface (b) from the nominal design.

First, it aims to optimize the efficiency of the primary drying process in the absence of parameter uncertainties. As be described in the last section, shelf temperature T_s and chamber pressure P_c are designed to achieve the maximum amount of ice removed by sublimation within the shortest drying time, while the appearance of the dried product is guaranteed by limiting temperature T_i at the sublimation interface below critical temperature T_c. The designed temperature and pressure profiles are shown in Fig. 5.5. As in Fig. 5.5(a), T_s is kept at its upper boundary, i.e., 30 °C, to provide more energy for the sublimation. In Fig. 5.5(b), T_i is also kept at its upper boundary to ensure higher vapor pressure P_i at the sublimation interface to accelerate the sublimation process according to Eqs. (5.22) and (5.23). At the beginning, P_c is set to 9.6 bar to achieve a higher sublimation speed and is decreased gradually to compensate the influence of the decreasing height of the frozen layer in accordance with Eq. (5.27). Note that the complete expressions of R_p and K_v from [70] are used for the nominal process design. Note that Mortier[70] also attempted to optimize the primary drying in the freeze-drying

process with a grid-based approach. In contrast to our results, the grid-based design compensates the influence of the decreasing height of the frozen layer by decreasing the shelf temperature, which leads to a certain loss in the sublimation speed. Consequently, it requires less than 6 h rather than almost 7 h in [70] to complete the primary drying process as indicated by the curve of m_{sub} in Fig. 5.5(b). Note that the transition phase at the beginning, i.e., heating the shelf and vacuuming the chamber to the design value, is neglected in this work because the period is quite short compared to the entire primary drying step. In addition, the choked flow limit at the vial neck calculated with Eq. 5.29 is not activated because the sublimation speed is far below the limit, which is also shown in the following subsection.

5.4.2 Effect of Parameter Uncertainties on the Nominal Process Design

The optimal process design for the nominal case above is based on the assumption that the model parameters are accurate. However, due to measurement imperfections and model simplifications, model parameters derived from noisy measurement data are imprecise and might be described best via arbitrarily distributed random variables [145]. In this particular case, it is assumed that the two parameters R_p and K_v are uncertain as indicated in the scatter plot in Fig. 5.4. Next, the adverse effect of these parameter uncertainties on the performance of the primary drying process obtained by ignoring these parameter uncertainties is analyzed. The results are illustrated in Fig. 5.6. As can be observed, the curves for the evolution of m_{sub} and T_i in Figs. 5.6(a) and 5.6(c) deviate from the nominal values and vary in certain ranges expressed with the confidence intervals (CIs) [164]. The primary drying process with smaller (higher) values for R_p and K_v has a lower (higher) T_i. In contrast, with smaller R_p and larger K_v values, the efficiency of the process is higher than what is expected regarding the nominal process design. As T_i exceeds the critical temperature $T_c = -34°\mathrm{C}$, this leads to an undesired collapse of the API cake. In Fig. 5.6(d), there is a high risk that the API product will be wasted as almost half of the probability distribution is on the right side of the red line, i.e., above the critical temperature. Therefore, it is necessary to consider parameter uncertainties in the design of the primary drying process as part of the robust process design.

5.4.3 Deterministic Samples

As discussed in Section 5.2, the GMD-PEM algorithm can be used to propagate and to quantify uncertainties in the field of robust process design efficiently. The resulting

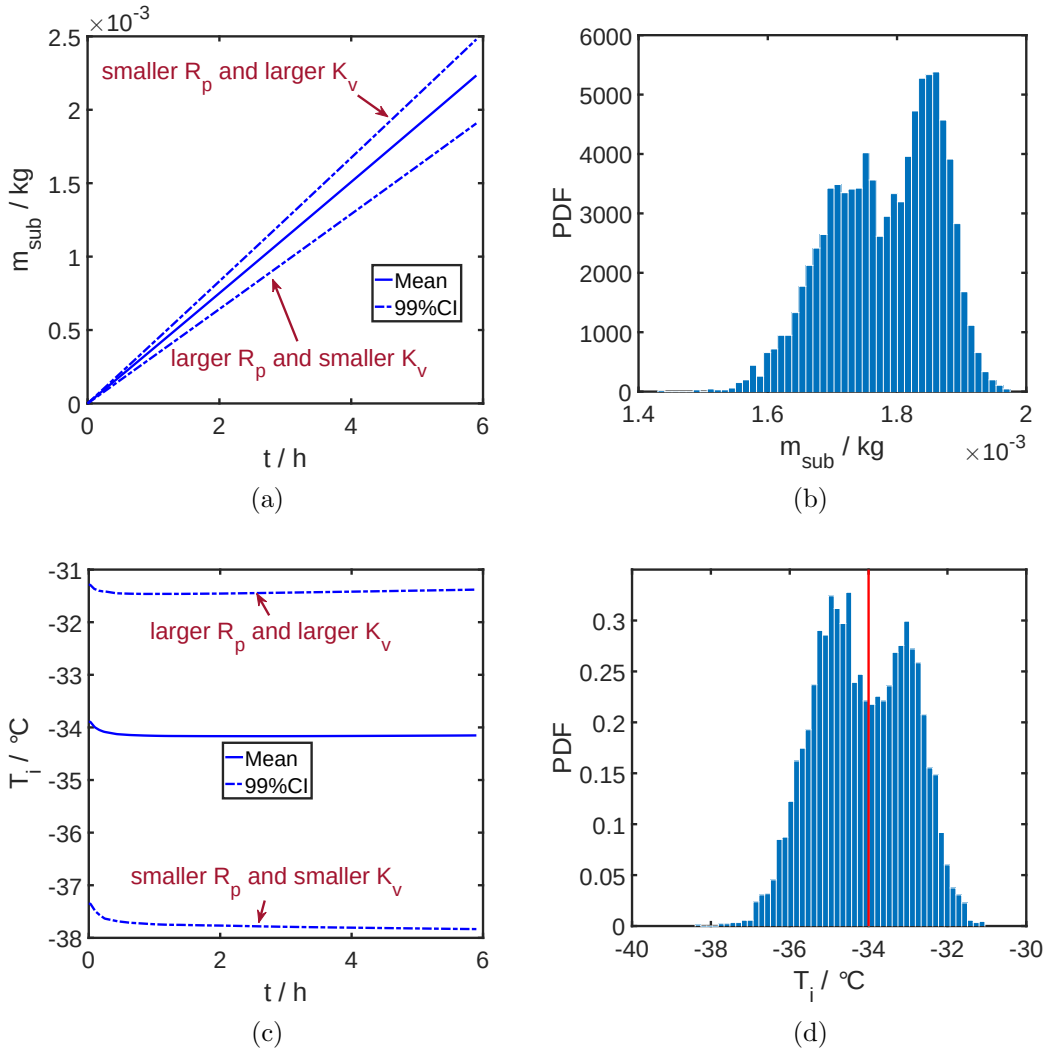


Figure 5.6: Evolution of the mean and 99% confidence interval (CI) of the mass of the ice removed by sublimation m_{sub} (a) and the temperature at sublimation interface T_i (c). (b) and (d) are the plots of the probability density function (PDF) of m_{sub} and T_i at t = 5 h, respectively.

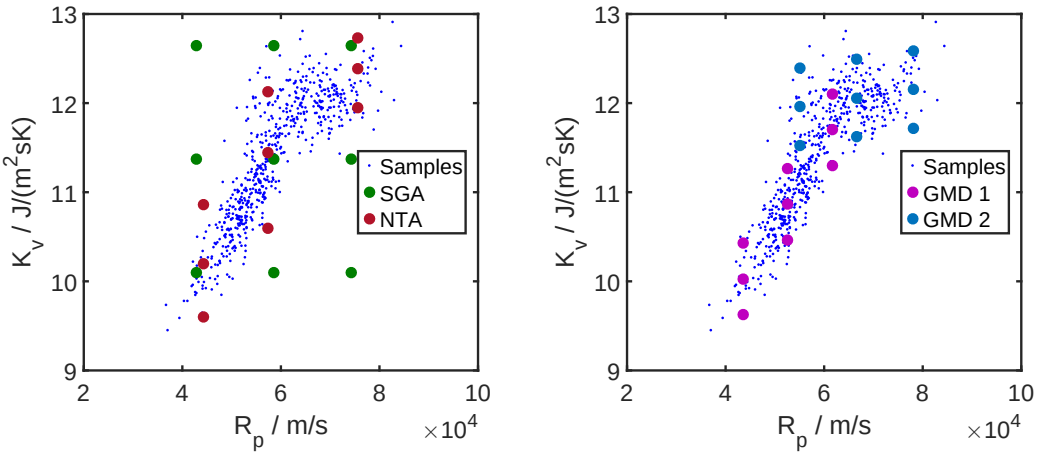


Figure 5.7: Illustration of the original samples and the deterministic samples derived from the three approaches. The green and red points in (a) are the deterministic samples from the single Gaussian approach (SGA) and the nonlinear-transfer approach (NTA), respectively. The purple and blue points in (b) are the deterministic samples for two component distributions of the Gaussian mixture distribution (GMD).

samples are plotted in Fig. 5.7 in comparison with sample points derived with the SGA and the NTA. The green sample points in Fig. 5.7(a) are derived with the SGA. They are symmetrically distributed in the parameter space as only the mean and the variance of the individual parameters are used; see Section 5.2.1. The red sample points in Fig. 5.7(a) are derived with the NTA. The NTA sample points are asymmetrically located in the parameter space due to the nonlinear transformation step addressing the correlation and the non-Gaussian shape of the parameter distributions. The EM algorithm and the BIC as explained in Section 5.2.1 are used to determine the number and the coefficients of the Gaussian component distributions. In Fig. 5.7(b), two component distributions are illustrated which have the best match with the parameter uncertainties. The corresponding deterministic samples for different component distributions are highlighted individually in Fig. 5.7(b) with purple and blue sample points, respectively. Although the number of GMD-PEM samples is twice as many as those for SGA or NTA, they can represent the shape of the original parameter uncertainties more accurately with lower computational cost compared to Monte Carlo simulations.

5.4.4 Approximation Accuracy

In this section, the accuracy of the novel GMD-PEM algorithm regarding the approximated mean and the variance of the freeze-drying process relevant state variables, i.e., m_{sub} and T_i, is analyzed. The results are summarized in Fig. 5.8. Here, the approximated mean and variance values are compared with the references from the Monte Carlo simulations. As shown in Figs. 5.8(a) and 5.8(b), the mean of m_{sub} can be estimated accurately with the SGA, NTA, and GMD-PEM algorithm, while the variances are estimated more accurately with the NTA and the GMD-PEM algorithm. Note that the approximated variances derived with the NTA deviate slightly from the reference, but the deviation is considerably smaller compared to the estimation based on the SGA; see Fig. 5.8(b). The same analysis for T_i is illustrated in Figs. 5.8(c) and 5.8(d). The estimated means and variances from the GMD-PEM algorithm are in good accordance with the references. The estimated mean based on the SGA or NTA is slightly lower than the references. However, the estimated variance deviates considerably from the reference, especially the variance approximated by the SGA. The GMD-PEM algorithm, in turn, not only provides a more accurate estimate of the mean and the variance but also captures the non-Gaussian shape of the model output distributions which is essential for approximating the robust inequality constraints given in Eq. (5.21f). For instance, the non-Gaussian distribution can be observed from the temperature at the sublima-

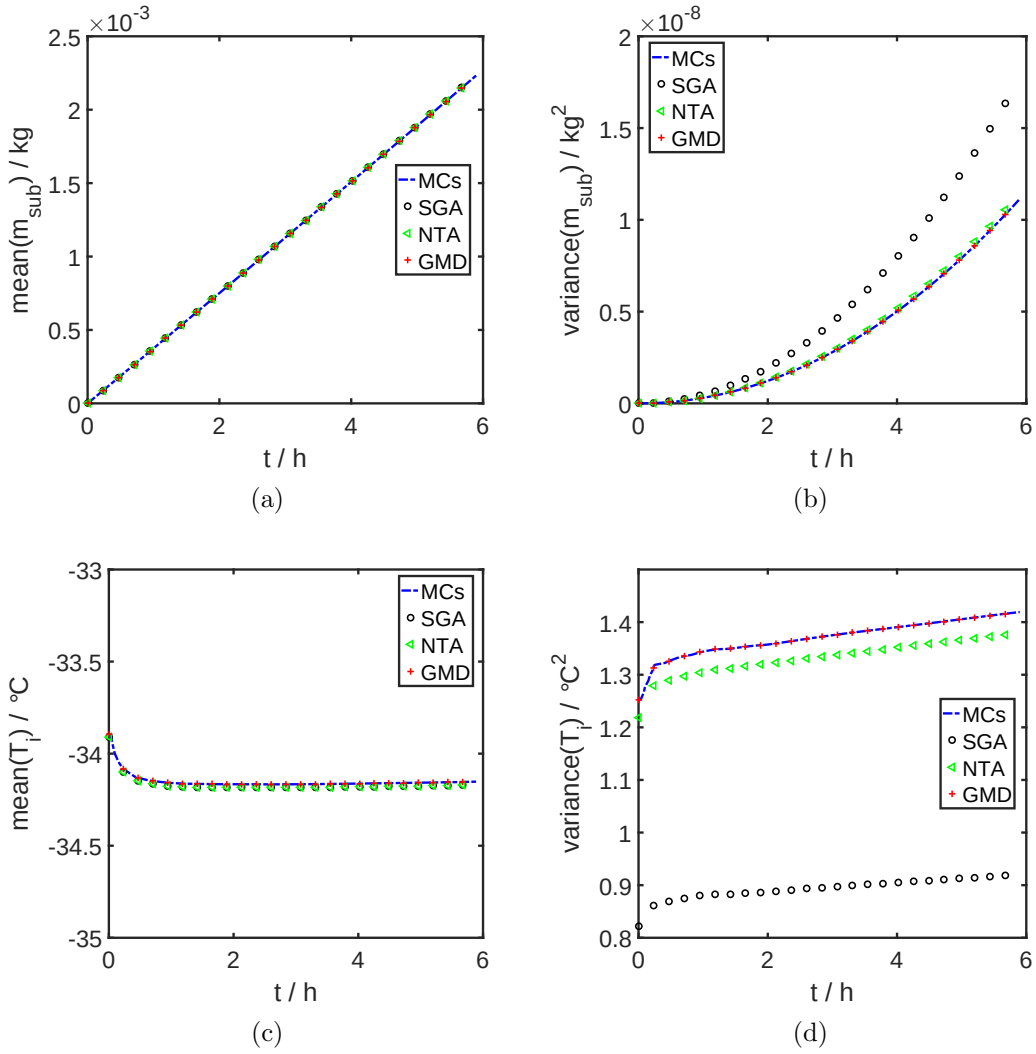
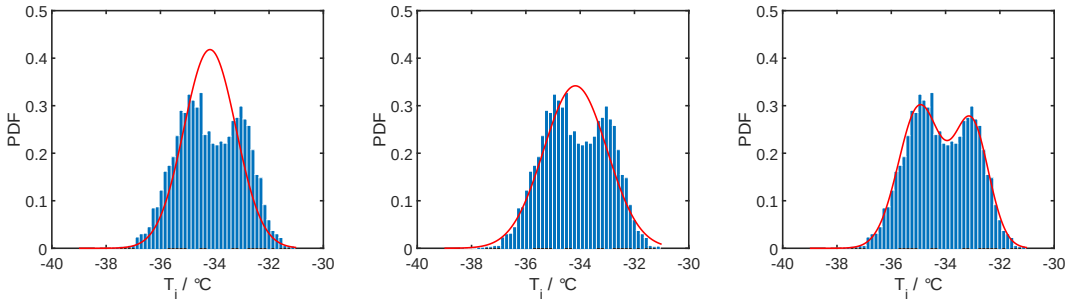


Figure 5.8: Comparison of the mean and the variance of the mass of the ice removed by sublimation m_{sub} and the temperature at sublimation interface T_i, which are estimated with the single Gaussian approach (SGA), the nonlinear-transfer approach (NTA), and the Gaussian mixture distribution (GMD). Values from Monte Carlo simulations (MCs) are used as references.



(a) SGA (b) NTA (c) GMD

Figure 5.9: Comparison of the original probability density function of the temperature at sublimation interface T_i at a single time point with the ones approximated by the single Gaussian approach (SGA), the nonlinear-transfer approach (NTA), and the Gaussian mixture distribution (GMD).

tion interface T_i given in Fig. 5.9 and is accurately approximated with the GMD-PEM algorithm.

5.4.5 Robust Process Design Results

The different robust process design strategies presented in Section 5.2.2 are implemented with the associated samples shown in Fig. 5.7 to design the primary drying process. The tolerance factor ε_{nq} is set to 1%, so that the risk of failure in the API cake appearance is equal or lower than 1%. The designed profiles for the control variables P_c and T_s are compared in Fig. 5.10. As it can be seen, the optimal shelf temperature is lower than the upper boundary to force T_i to be lower than the critical collapse temperature with 99% probability. In parallel, the chamber pressure decreases to its lower boundary 5 bar to accelerate the sublimation. While the chamber pressure is fixed, T_s is gradually decreased to compensate for the influence of decreasing height of the frozen layer. As discussed in Section 5.4.4, the approximation accuracy of the SGA, NTA, and GMD-PEM is different, and thus, the optimized T_s is different as well. To compare the performance of the results obtained from the SGA, NTA, and GMD-PEM algorithm, the primary drying process is simulated with 6000 samples that are derived from the original parameter distributions; see Fig. 5.4. The mean value of m_{sub} at the final time point and the number of constraint violations are calculated based on the simulation results which are related to the 6000 parameter samples. The results are listed in Table 5.2. The nominal design has the best efficiency: a low freeze-drying time and high m_{sub}. However, almost 50% of the

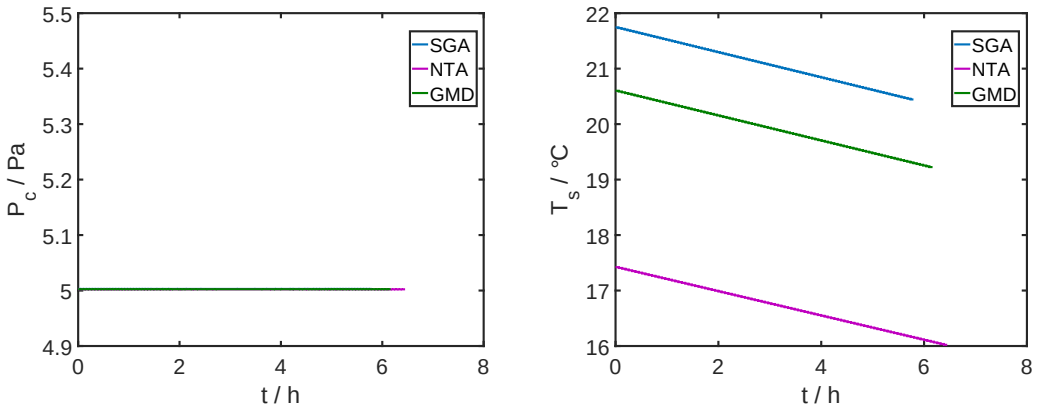


Figure 5.10: Results from a robust process design with the deterministic samples from the single Gaussian approach (SGA), the nonlinear-transfer approach (NTA), and the Gaussian mixture distribution (GMD). (a) and (b) are the profiles for chamber pressure P_c and shelf temperature T_s, respectively.

simulations violate the temperature constraint; i.e., T_i exceeds the critical temperature. The SGA result, in turn, has a better performance compared to the results from the NTA and the GMD-PEM algorithm, but the constraint violation is $\frac{94}{6000} = 1.6\%$ which is almost twice as high as the target value $\varepsilon_{nq} = 1\%$. The resulting NTA setting has much fewer constraint violations, but the efficiency is low with 6.45 h for the primary drying process. In other words, the design is too conservative. In contrast, the derived robust design from the GMD-PEM algorithm has the best trade-off between process efficiency and product quality, i.e., the API cake appearance and integrity. In Table 5.2, the GMD-PEM design achieves the same m_{sub} within 6 h and results in a failure probability of $\frac{62}{6000} = 1\%$. The choked flow limit is investigated for the GMD-PEM design, and the sublimation speed is considerably slower than the choked flow limit, as shown in Fig. 5.11. Thus, the flow rate of the vapor at the vial neck is not fast enough to reach the speed of sound and to trigger the choked flow phenomenon [162].

5.5 Chapter Summary

In this chapter, a novel GMD-PEM algorithm for propagating and quantifying parameter uncertainties with non-Gaussian distributions was introduced. The GMD-PEM algorithm is benchmarked with two approaches which are either commonly used in lit-

Table 5.2: Results from the nominal design and the robust design with the deterministic samples from the single Gaussian approach (SGA), the nonlinear-transfer approach (NTA), and the Gaussian mixture distribution (GMD). 6000 samples generated from the original samples are used to validate the number of constraint violations for the different methods. Samples column in the table means the number of sample needed for robust process design with different approaches, which is proportional to the computational cost.

	drying time [h]	mean(m_{sub})$\times 10^3$ [kg]	constraint violations	samples
Nominal	5.90	2.2	50%	-
SGA	5.80	2.0	1.6%	9
NTA	6.45	2.0	0.15%	9
GMD	5.98	2.0	1%	18

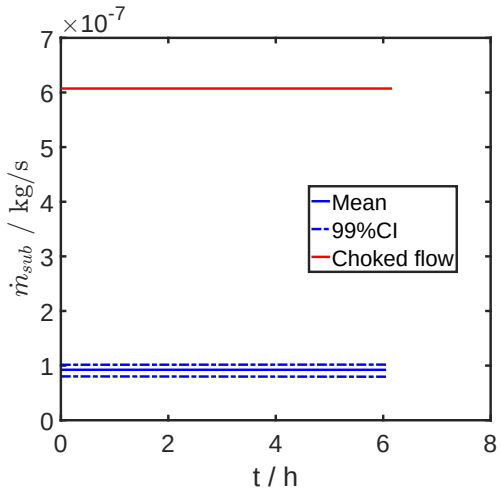


Figure 5.11: Comparison of the sublimation rates during the the primary drying process and the choked flow limitation.

erature or introduced in the previous chapter. The robust process design framework proposed in this chapter is constructed based on GMD-PEM algorithm in which the component distributions are quantified with the expectation-maximization algorithm. PEM samples for the component distributions are generated with linear transformation and used to approximate the statistic information in both the objective function and constraints. A case study on freeze-drying process is investigated in this chapter. The primary drying step of the freeze-drying process is firstly optimized while the parameter uncertainties were ignored. A design with higher drying efficiency was obtained compared to the literature but might result in low-quality products in practice due to the neglected parameter uncertainties. Afterwards, the proposed robust design approach is used to design the primary drying step with taking into account the non-Gaussian distributed parameter uncertainties. Moreover, robust design framework with other approaches, i,e, SGA and NTA, are also implemented to design the same primary drying step. As demonstrated, the proposed algorithm can approximate the non-Gaussian model parameters and the output distributions adequately, which is essential in fulfilling inequality constraints under non-Gaussian parameter uncertainties. Monte Carlo simulations were used to evaluate the final design. In comparison to SGA and NTA approach, the GMD-PEM algorithm provides the best trade-off between process efficiency and the acceptable limit of constraint violations on the product-relevant sublimation interface temperature.

Chapter 6

Stochastic Back-off Robust Process Design[6]

In this chapter, limitations on strategies for robust process design introduced in Chapter 4 and 5 are addressed. A novel stochastic back-off strategy is proposed to handle the potential flaws of the aforementioned strategy for robust process design. PCE is used to compute the statistical information needed for the novel back-off strategy. Moreover, a back-off strategy, in which PEM is used, is also presented and compared with the novel approach. The proposed approaches are demonstrated on a separation process of pharmaceutical manufacturing, i.e., a plug-flow crystallizer.

The chapter is organized as follows. A brief introduction and motivation of this chapter is provided in Section 6.1. In Section 6.2, the basics of the traditional back-off strategy are presented. Next, the novel stochastic back-off strategy is introduced in Section 6.3. A case study involving the continuous crystallization of ibuprofen is detailed in Section 6.4. Results and discussions are provided in Section 6.5. Section 6.6 summarizes the chapter.

6.1 Motivation

The probability-based robust process design provides less conservative solution compare to the scenario-based method [30, 3]. In Chapter 4 and 5, probability-based approaches are proposed to incorporate the parameter uncertainty with any types of probability distributions and mutual dependencies in process design to derive expected solutions. The PEM, as the most efficient UQ method indicated in Chapter 2, is used as the core of these approaches to compute the statistical moments needed. However, within the optimization iterations, the statistical moments have to be calculated and updated in each iteration. This means that the computational demand of the approaches depends not only on the UQ method, but also on the number of optimization iterations. In case

[6]Part of this chapter has been published in (Xie et al., Comput. Chem. Eng., 124, 80-92, 2019 [11])

of inadequate optimization formulation, e.g., poor initial guesses, the computational demand of robust design becomes burdensome.

There are at least two alternatives for circumventing this issue. The first one is introducing a surrogate models to replace the process model and alleviate the computational burden for calculating the statistical moments, as illustrated in [165, 1, 166, 167, 5]. The accuracy of the surrogate model, however, is a crucial problem with this approach, especially for a large-scale optimization problem [1]. Alternatively, Srinivasan[154] introduced an iterative back-off strategy to shrink the feasible region of the optimization problem and thus, making the optimal operating conditions robust. The merit of this iterative back-off strategy is that the back-off terms are determined outside the optimization, as such robust process design does not include additional complexity from calculating the statistical information and remains the same computational efficiency as solving a nominal optimization problem. With the benefit of its computational efficiency, this approach has been further investigated and implemented in model-predictive control [168], optimal experimental design [169] and robust process design [73, 170, 8].

In this chapter, the structure of the conventional iterative back-off strategy is briefly summarized and the procedure is improved by introducing a quantitative update rule for the back-off factor η which controls the conservativeness of a robust process design. With implementing the adapted iterative back-off strategy, the desired robustness for the constraints in robust process design can be guaranteed, even for non-Gaussian distributions, without trial-and-error updating of η factor. However, the two nested loops for determining back-off terms and the η factor introduce additional difficulties in the convergence of the conventional iterative back-off strategy. Moreover, the validation and calculation of the η factor limit the strategy's overall efficiency. Therefore, the main contribution in this chapter is to introduce a novel stochastic back-off strategy which has simpler structure. The key idea is that the back-off terms are derived by the entire probability distribution rather than the standard deviations of the constraints. Moreover, the computational efficiency of the stochastic back-off strategy is improved considerably by using PCE to replace the CPU-intensive process model and to approximate the probability distributions of the constraints [82, 171].

To demonstrate the performance of the stochastic back-off strategy, a continuous crystallization process of ibuprofen is used as a case study. Ibuprofen is one of the most commonly used medicines for treating pain, fever, and inflammatory diseases [172]. The mechanisms and kinetics of the crystallization of ibuprofen, e.g., the nucleation and growth rates, were recently investigated [172]. The identified value of the kinetics parameters and the associated uncertainties from [172] are used to design a continuous

plug-flow crystallizer for ibuprofen. The stochastic back-off strategy is implemented to ensure the supersaturation level of the solution stays below the primary nucleation threshold to avoid unstable primary nucleation. Moreover, this strategy is compared with other iterative back-off strategies to highlight the superiority of the proposed stochastic concept in terms of computational efficiency.

6.2 Basics of the standard back-off strategy

Assume there is an inequality constraint for the process design

$$h(x, \theta) \leq 0, \tag{6.1}$$

where function $h : \mathbb{R}^{n_x \times n_\theta} \rightarrow \mathbb{R}^{n_h}$, x is the vector of the state variables, and θ is the vector of the uncertain parameters with the joint probability density function (PDF) $f(\theta)$. Technically, to ensure the robustness of the inequality constraint, Eq. (6.1) must be satisfied for all or at least most possible parameter realizations of PDF $f(\theta)$, as presented in Chapter 4. However, straightforward implementation of Eq. (4.1f) for all possible realizations results in a semi-infinite optimization problem (similar to the worst-case scenario) which might be NP-hard. Alternatively, the inequality constraint can be robustified by introducing back-off terms to ensure the reliability of the inequality constraint under uncertainty. The back-off terms b are added to the left hand side of the inequality constraint in Eq. (6.1)

$$h(x, \theta_n) + b \leq 0, \tag{6.2}$$

where θ_n represents the vector of the nominal parameter values. Here, Eq. (6.2) has the same complexity as the inequality constraint for the nominal design, and the back-off term b, which could be a constant or time-varying variable, moves the inequality constraint at the nominal condition away from its boundary to ensure a sufficient safe margins for counteract with the influence from parameter uncertainties.

Pioneering work in [173] on the back-off concept has been advanced to select control structures for different chemical processes, in which linearized models are used to calculate the back-off terms [174, 175]. The original back-off strategy has also been used to guarantee the dynamic feasibility of joint process design and control optimization problems [176, 177, 178]. [179] and [180] approximated back-off terms with constraint derivatives and implemented for robust design of a fed-batch fermentation process and a batch distillation process, respectively. However, the back-off terms used were de-

termined by linearization of the model and the constraints, and thus, are not reliable for the robust design of highly non-linear pharmaceutical processes and might lead to robustification that is either too conservative or not robust enough with excessive constraint violations. Moreover, [154] proposed the idea of iterative calculation of back-off terms; i.e., in addition to the optimization loop an outer back-off validation loop is used based on Monte Carlo simulations. This iterative back-off concept was also implemented and refined by [73] for robust optimization of a polymerization process, in which Monte Carlo simulations were used to propagate the parameter uncertainties through the original process model and to approximate the resulting mean and variance values of the constraints for the back-off calculation as summarized in Eqs. (6.3) to (6.5):

$$b = \eta\sqrt{Var(h(x,\theta))}, \tag{6.3}$$

$$\begin{aligned} Var(h(x,\theta)) &= \int_{I_\theta} (h(x,\theta) - E(h(x,\theta)))^2 f(\theta) d\theta \\ &\approx \sum_{i=1}^{n_{MC}} \frac{1}{n_{MC}} (h(x,\theta_i^{MC}) - E(h(x,\theta)))^2, \end{aligned} \tag{6.4}$$

$$E(h(x,\theta)) = \int_{I_\theta} h(x,\theta) f(\theta) d\theta \approx \sum_{i=1}^{n_{MC}} \frac{1}{n_{MC}} h(x,\theta_i^{MC}), \tag{6.5}$$

where n_{MC} is the number of Monte Carlo simulation samples, $E(h(x,\theta))$ and $Var(h(x,\theta))$ indicate the mean and the variance of the constraint function h, and η is the factor for adapting back-off terms to controls the conservativeness/robustness of the robust constraints. Shi[73] also demonstrated that the back-off terms could be improved with an iterative update based on Eqs. (6.3) to (6.5), in which the optimal design from the previous iteration is used. And the results from the back-off strategy are equivalent to those from multi-scenario optimization once the back-off terms converge. To reliably approximate the mean and variance values with Eqs. (6.4) and (6.5), n_{MC} has to be large, which might be prohibitive especially for large-scale systems and complex process models [8]. Emenike[8] came up with the idea of calculating back-off terms with the point estimate method (PEM) as a more efficient alternative to the conventional Monte Carlo simulations. Based on their results [8], a dramatical improvement of computational efficiency of robust process design is observed with using the PEM-based back-off terms. In addition, Koller[181] investigated the impact of different η values on the performance of a back-off strategy by applying it to a simultaneous design, control, and scheduling problem of multi-product systems. Maussner[170] introduced an additional iteration loop to update the η value to ensure that the number of constraint violations is

below an acceptable level. This particular double-loop back-off strategy is summarized in the workflow diagram in Fig. 6.1.

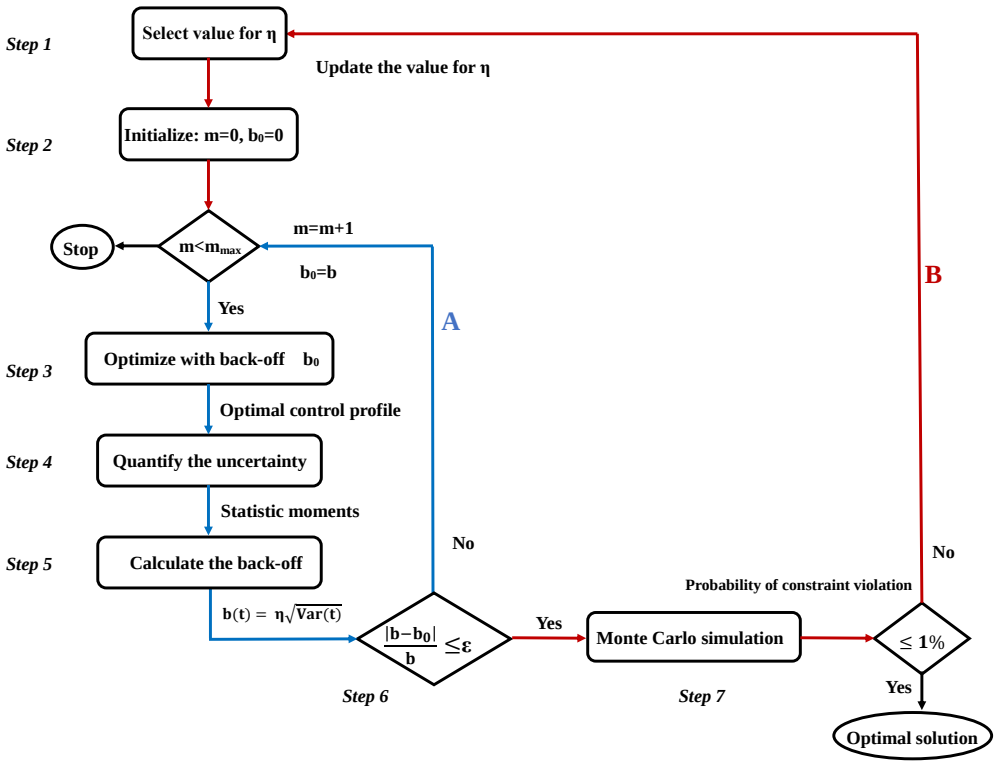


Figure 6.1: Flow diagram for the double-loop back-off strategy for robust process design.

The double-loop back-off strategy given in Fig. 6.1 has an internal loop (A) for the convergence of the back-off terms and an external loop (B) for updating the value of the factor η. More details about the double-loop back-off strategy presented in [8] are summarized below.

Step 1 (External loop (B) start) Specify the initial η value. The quantile of the standard Gaussian distribution with the desired probability is taken. For example, the quantile is set to 2.33 to ensure that 99% of the inequality constraints are satisfied.

Step 2 Initialize the internal loop (A).

Step 3 (Internal loop (A) start) Optimize with the inequality constraints; i.e., $h(x, \theta) + b_0 \leq 0$, where b_0 is the back-off term. Note that the optimization in the

first iteration with $b_0 = 0$ is equivalent to nominal optimization and has the same computational complexity as the nominal optimization problem in general.

Step 4 Quantify the impact of the parameter uncertainties on the constraints. The PEM is used to estimate the mean and variance of the constraints as given in Eqs. (6.6) and (6.7) based on the optimal design result from Step 3:

$$E(h(x,\theta)) \approx \sum_{i=1}^{2n_\theta^2+1} w_i h(x,\theta_i^{PEM}), \tag{6.6}$$

$$Var(h(x,\theta)) \approx \sum_{i=1}^{2n_\theta^2+1} w_i (h(x,\theta_i^{PEM}) - E(h(x,\theta)))^2, \tag{6.7}$$

where n_θ is the number of uncertain parameters, and w_i and θ_i^{PEM} are the weight factors and the deterministic parameter samples, respectively. [3].

Step 5 Calculate the back-off term with η and the variance calculated in Step 4 by using Eq. (6.3). The time-varying back-off term $b(t)$ is used as it provides more flexibility and better solution [73].

Step 6 (Internal loop (A) end) Check if the back-off term converges; if not, b_0 is replaced with the new back-off term b and steps 3 to 6 are repeated until it converges.

Step 7 (External loop (B) end) With the converged back-off term, the optimal design is validated with Monte Carlo simulations. To this end, 10,000 realizations generated from the distribution of the parameter uncertainty are evaluated and used to calculate the probability of a constraint violation. If the violation probability is smaller than 1%, the optimal solution is exported. If not, a new value for η has to be selected and the whole algorithm is repeated. Note that any other violation probability might be feasible but affects the initial η value selection procedure in Step 1.

As can be seen, the two loops are essential to fulfill the probability of the given constraint violation limits. Note that the internal loop with the initial guess of η from the quantile of a standard Gaussian distribution might be sufficient but only if the probability distribution of the constraint function follows a Gaussian distribution. However, models for pharmaceutical processes are complex and highly nonlinear, and thus, the distribution of the constraint function is typically non-Gaussian [182]. Maussner[170] provided candidate values for η based on expert guessing to update the η factor. In this work, a more systematic and problem-specific procedure is proposed for updating η

according to:

$$\eta_i = \eta_{i-1} \frac{\text{norminv}(99\%)}{\text{norminv}(1 - e_c)}, \tag{6.8}$$

where norminv means the inverse cumulative density function of a standard Gaussian distribution, and e_c is the probability of a constraint violation calculated in Step 7. As fewer η values are tested as a result, the overall computational efficiency of the double-loop back-off concept can be improved.

With the additional external loop, the double-loop back-off strategy is capable of handling constraints with non-Gaussian distributions. However, the high computational costs and the redundant structure might be critical for many practical problems in robust process design. Although the PEM is used to reduce the cost of the internal loop considerably, the external loop still needs a vast number of Monte Carlo simulations to validate the probability of a constraint violation. For this reason, the computational efficiency of the double-loop back-off strategy deteriorates dramatically if the external loop converges slowly. To circumvent the redundant structure and the heavy computational burden for updating η, a novel, highly effective stochastic back-off strategy is proposed.

6.3 The stochastic back-off strategy

Before outlining the novel approach, as a motivation, it deserves to explain why Eq. (6.3) is not an appropriate formulation for calculating the back-off terms first. In Fig. 6.2, the calculation of back-off terms is illustrated assuming a Gaussian distribution (A) and a non-Gaussian distribution (B). In particular, the back-off term b is determined with the distance between the nominal value of the inequality constraint and its 99% quantile. In the case of a Gaussian distributed inequality constraint, the mean value is equal to the nominal value, and the back-off term b is equal to the confidence interval; i.e., $b = \eta_{99\%} \sqrt{Var(h(x, \theta))}$. However, in the case of a non-Gaussian distribution, these two aspects do not hold, and thus, there is an external loop in the double-loop back-off strategy to approximate the real value of the back-off terms by adapting the value of η iteratively; see Eq. (6.8). As mentioned, the iterative update of η is inefficient and might increase the computational cost dramatically. Alternatively, it is suggested to calculating the back-off terms directly with the distance between its normal value and the empirical quantile at 99% based on the probability distribution of inequality constraints. And the stochastic back-off strategy is constructed based on the idea, which is explained in more detail below.

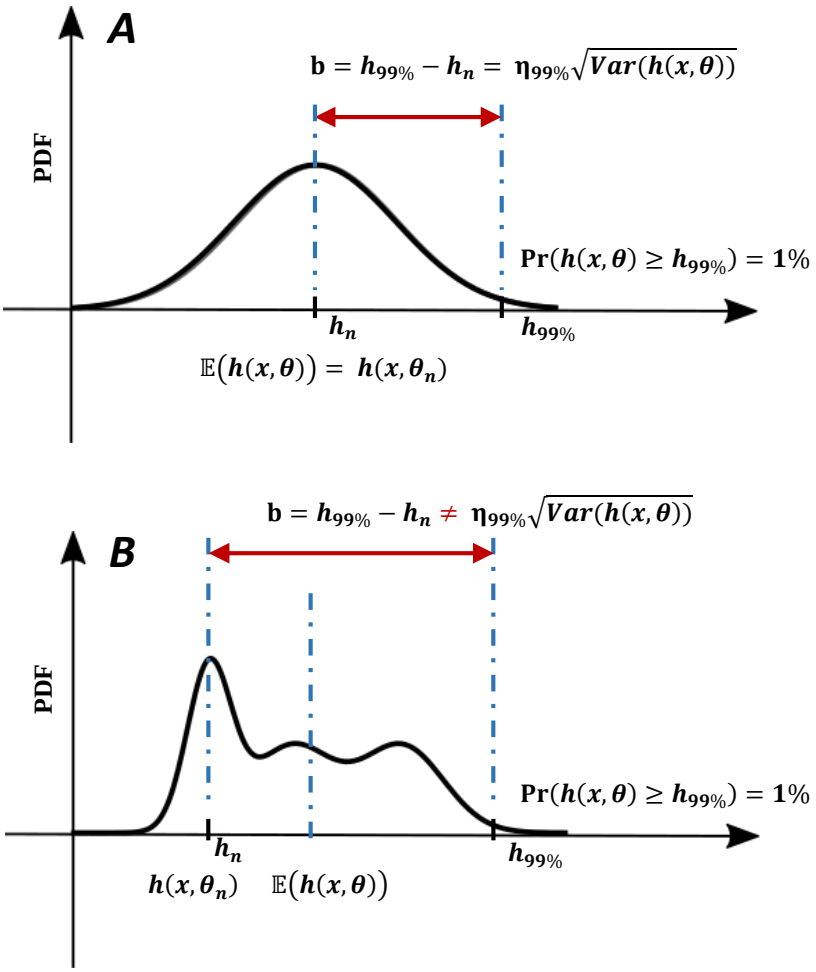


Figure 6.2: Calculation of the back-off terms for the cases where probability distribution of constraint function $h(x,\theta)$ is Gaussian (A) and non-Gaussian (B), respectively. $h_n = h(x,\theta_n)$ and $h_{99\%}$ indicate the value of constraint function at nominal point and point with cumulative density equal to 99%.

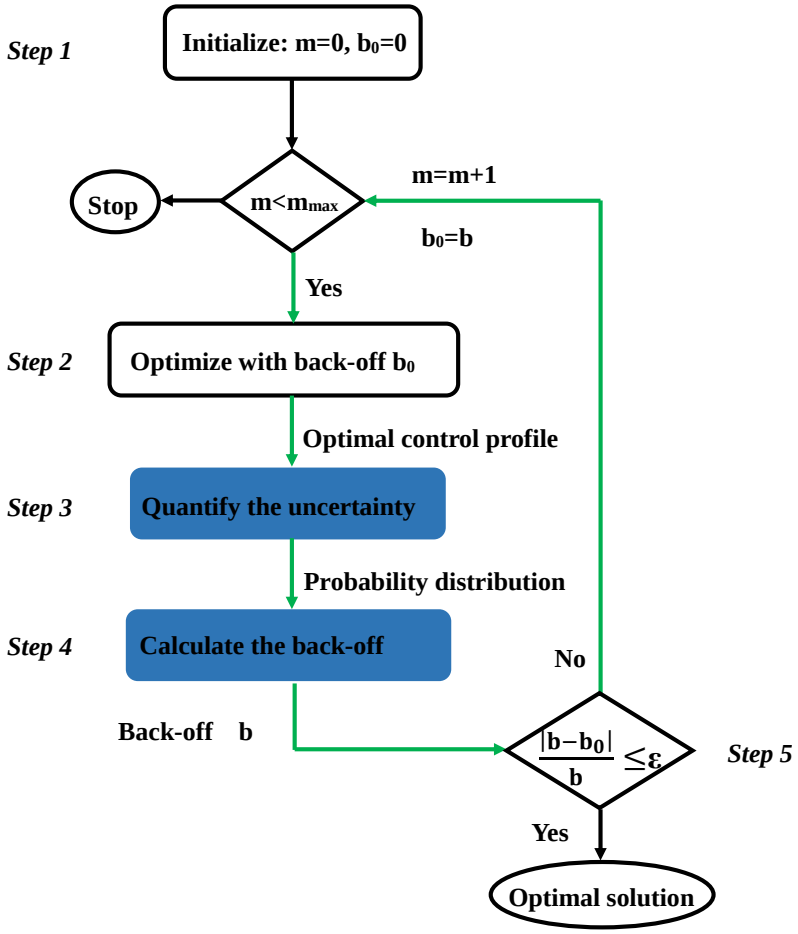


Figure 6.3: Workflow of the stochastic (single-loop) back-off strategy for robust process design.

In Fig. 6.3, the workflow diagram of the novel stochastic back-off strategy is illustrated. The structure of the stochastic back-off strategy is similar to the internal loop of the double-loops back-off strategy in Fig. 6.1 but has significant differences for the uncertainty quantification and the approximation of the back-off terms:

Step 1 (Loop start) Initialize of the stochastic back-off strategy.

Step 2 Optimize under inequality constraints; i.e., $h(x, \theta) + b_0 \leq 0$. Note that the optimization in the first iteration with $b_0 = 0$ is actually equivalent to nominal optimization, and thus, has the same computational complexity as nominal optimization.

Step 3 Propagate and quantify the constraint uncertainties with the optimal design from Step 2. Instead of the variance, the exact shape of the probability distribution of the constraints is approximated with KDE and PCE.

To avoid a repetitive evaluation of the original CPU-intensive model, the polynomial chaos expansion (PCE) approach presented in Section 2.3 is used as an alternative concept. Thus, the original process model is replaced by a surrogate model that can be evaluated with low computational costs [1]. The number of samples for estimating the PCE model depends on the complexity of the model and the number of uncertain parameters but is typically negligible compared to the number of samples needed for approximating the probability distribution.

Note that Monte Carlo simulations as a traditional sample-based method could also be used here for the uncertainty quantification step. However, the deficiency of Monte Carlo simulations, as was addressed in the double-loop approach [8], might be prohibitive for the stochastic approach.

Step 4 Calculate of the back-off terms with the probability distribution of the constraints. As shown in Fig. 6.4, the sample evaluations from the PCE model are processed to approximate the probability distribution of the constraint function making use of the KDE in Step 3. The resulting probability distribution is subsequently used to calculate the back-off terms. As illustrated in Fig.6.2, it is more appropriate to calculate the back-off terms with the empirical quantile distance than just the standard deviation. Therefore, the back-off terms for robust design with a probability of constraint violation $\leq 1\%$ is determined by

$$b = D(h_n, h_{99\%}) = \hat{F}_h^{-1}(99\%) - h_n, \tag{6.9}$$

where $D(\cdot)$ means the distance function, and $\hat{F}_h^{-1}(\cdot)$ is the inverse cumulative density function of the constraints adapted from Eq. (3.50).

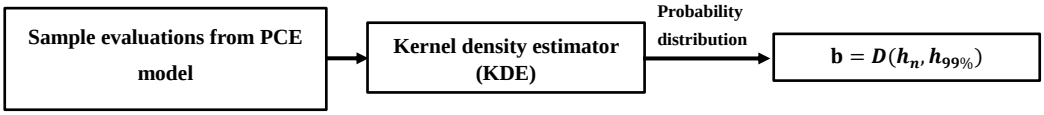


Figure 6.4: Calculation of the back-off term, b, for the stochastic back-off strategy. $D(h_n, h_{99\%})$ is the distance function that calculates the difference between h_n and $h_{99\%}$. h_n and $h_{99\%}$ indicate the values of the constraint function at the nominal point and the point with a cumulative density equal to 99%, respectively.

Step 5 (Loop end) Check if the back-off terms converge. If not, replace b_0 with the new back-off b and repeat steps 2 to 5 until the procedure converges.

To conclude, the novel concept of a stochastic back-off implementation has a simpler structure in comparison to the double-loop back-off strategy while PCE ensures low computational costs at the same time. More details about the performance of the novel stochastic back-off concept are described in the following case study.

6.4 Case study: a continuous plug-flow crystallizer for ibuprofen

The case study aims to design a continuous plug-flow crystallizer (PFC) for the API ibuprofen. The PFC has the advantage of higher efficiency and narrower crystal distributions compared with the commonly used mixed suspension mixed product removal crystallizer (MSMPRC) and has been used to crystallize various APIs [185, 186]. The focus of this work is to optimize the steady-state operation of the PFC and to maximize the mass-based mean crystal size (d_{43}) under the condition of uncertainty.

6.4.1 Mathematical model

The scheme of a continuous PFC is illustrated in Fig. 6.5. The model for the PFC, which is adapted from [186], consists of the population balance equation (Eq. (6.10)) that describes the evolution of the crystal size distribution (CSD) and the mass balance equation (Eq. (6.11)) that describes the mass balance between the liquid and solid phases. Only the mass balance equation in liquid phase is shown in Eq. (6.11), since

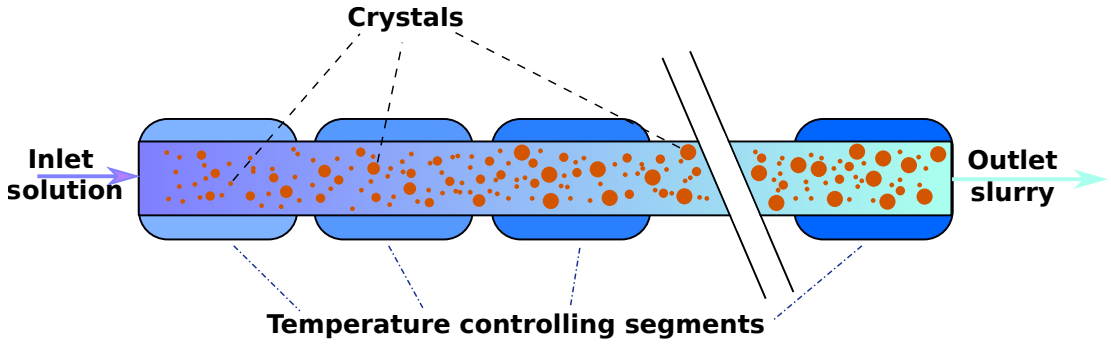


Figure 6.5: Schematic diagram of a continuous plug flow crystallizer for ibuprofen. The temperature controlling segments are used to realize the cooling profile for the crystallizer.

the mass balance equation in the solid phase is conjugated to that and it is not necessary to show both of them in the governing equations. Note that the assumption here is no dissolution, agglomeration, and breakage happening during the crystallization of ibuprofen as discussed in [172]. Another assumption is that no mixing effect exists in the PFC and that dispersion of the crystal density and the API concentration exist only in the axis direction. The governing equations for the steady-state PFC model are:

$$0 = \frac{\partial(vn)}{\partial z} + \frac{\partial(Gn)}{\partial L} \tag{6.10}$$

$$\text{Liquid:} \quad 0 = \frac{\partial vC}{\partial z} + k_v \rho_s (BL_0^3 + 3\int_0^\infty GL^2 n dL), \tag{6.11}$$

where z is the axis coordinate of PFC, m; L is the characteristic crystal size, m; n is the size density of crystals per kilogram of slurry, #/kg/m; B is the nucleation rate, #/kg/s; G is the crystal growth rate, m/s; C is the mass of solute per kilogram slurry, kg/kg; k_v and ρ_s are the shape factor and the density of the crystals, kg/m^3, respectively, and v is the superficial velocity of the slurry along the PFC, m/s. The mass of the solution and the solids is considered, and it is assumed that the formation of the solids does not change the volume of the slurry [186]. Therefore, the superficial velocity v is considered as constant for the entire PFC. Note that the mass equation for the solid phase is already implicitly included in Eq. (6.10). The boundary conditions at $z = 0$ and $L = L_0$, where

L_0 is the size of the nuclei, m, of the model are:

$$n(0, L) = n_{feed}(L) \tag{6.12}$$

$$n(z, L_0) = \frac{B}{G} \tag{6.13}$$

$$C(0) = C_{feed}. \tag{6.14}$$

As it can be seen, the steady-state PFC model consists of partial differential equations (PDEs) and has to be discretized or modified to be solved by a common ordinary differential equation (ODE) solver. The classical method of moments (MOM) can be used to transfer the PDEs into several ODEs because the growth rates G are assumed to be size independent [172, 186]. For the classical MOM, size density n is multiplied with the k^{th} order of crystal size L and subsequently integrated over the entire crystal size domain to compute its k^{th} moment, i.e, μ_k. Typically, the first six moments, $k = 0, \ldots, 5$, are used to represent the key information included in size density n [186]:

$$\mu_k = \int_0^\infty L^k n dL, \quad k = 0, \ldots, 5. \tag{6.15}$$

The resulting ODE systems reads as:

$$\frac{d\mu_0}{dz} = \frac{B}{v} \tag{6.16}$$

$$\frac{d\mu_k}{dz} = \frac{BL_0^k}{v} + \frac{kG\mu_{k-1}}{v} \quad k = 1, \ldots, 5 \tag{6.17}$$

$$\text{Liquid:} \quad \frac{\partial C}{\partial z} = -\frac{k_v \rho_s}{v}(BL_0^3 + 3G\mu_2). \tag{6.18}$$

Alternatively, to calculate the probability density of the crystal number, a high-resolution scheme based on the finite-volume method (FVM), proposed in [187], is used to solve the PDEs with discretization of the characteristic crystal length L, as illustrated in Fig. 6.6. The resulting ODEs are given in Eq. (6.19) that could be solved directly with Eq. (6.11) with standard ODE solvers to calculate the probability density of the crystal number. The cell-face fluxes $n_{L_{i\pm1/2}}$ are computed with a robust upwind discretization method [187]:

$$\frac{\partial(n_i)}{\partial z} + \frac{G}{v\Delta L}(n_{L_{i+1/2}} - n_{L_{i-1/2}}) = 0, \quad i = 1, \ldots, N. \tag{6.19}$$

The classical MOM generates less complex ODE systems (= 7), and therefore, is more suitable for the optimal (robust) crystallizer design regarding computational costs. In

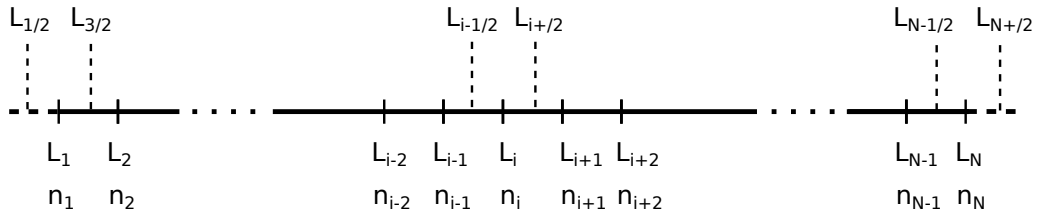


Figure 6.6: Cell centered discretization scheme of finite-volume method for internal domain, i.e., crystal characteristic length L, adapted from [187]. L_i refers the cell centers, and $L_{i-1/2}, L_{i+1/2}$ represent cell faces. n_i is the average value of the number density of crystals in cell i.

contrast, the FVM needs a fine mesh (≥ 100) for the characteristic crystal length and is too redundant to be embedded in the optimization algorithm. Therefore, the FVM is used only to generate the reference probability density of the crystal number for illustration and validation in what follows.

6.4.2 Crystallization kinetics of ibuprofen

The kinetics for the crystallization of ibuprofen in absolute ethanol are adapted from [172]. The main driving force for crystallization is the degree of supersaturation S, which is defined with the difference between solution concentration C_{sol} and solubility C^* as:

$$S(z) = C_{sol}(z) - C^*(T(z)). \tag{6.20}$$

The solubility of ibuprofen in absolute ethanol is a function of the temperature, and thus, can be used to design the crystallization process [172]:

$$C^* = 0.495 + 0.001026T^2. \tag{6.21}$$

All the quantities above have the same unit kilogram of solute per kilogram of ethanol. The concentration of the solution can be derived from the slurry mass solution and the solid concentration C and C_s with the following relations:

$$C_{sol} = \frac{C}{1 - C - C_s}, \tag{6.22}$$

$$C_s = k_v \rho_s \mu_3. \tag{6.23}$$

Table 6.1: Nominal values and units of the parameters for ibuprofen crystallization and the plug-flow crystallizer [172].

Parameters	Unit	Nominal value
k_{b_0}	#/kg/s/(kg solute/kg ethanol)	1.73×10^8
k_{g_0}	m/s/(kg solute/kg ethanol)	5.3
T_g	°C	42
n_b	–	1
n_g	–	1
k_v	–	$\pi/6$
ρ_s	kg/m^3	1100
v	m/s	0.007
z_f	m	20.16

Technically, the solubility of ibuprofen changes not only with the temperature but also with the composition of the solution. Water as an antisolvent can be added to decrease the solubility of ibuprofen as investigated by [172]. However, ibuprofen induces phase separation in the water-ethanol mixture especially at $40°C$ [172]. Moreover, the information about ibuprofen solubility in the water-ethanol mixture is not complete. Therefore, the focus is about the crystallization of ibuprofen in absolute ethanol. The kinetics for crystallization, i.e., the nucleation rate and the growth rate, in absolute ethanol are given below, and the values for the kinetic parameters in Eqs. (6.24) and (6.25) are listed in Table 6.1:

$$B = k_{b_0} S^{n_b} \tag{6.24}$$

$$G = k_{g_0} \exp(\frac{T}{T_g}) S^{n_g}. \tag{6.25}$$

6.4.3 Optimization problem

In this section, the structure of the nominal optimization problem of the PFC is introduced in Eq. (6.26). The objective function in Eq. (6.26a) is to maximize the critical quality attribute (CQA) of the crystallization process, i.e., the mass-based mean crystal size d_{43} (see Eq. (6.26f)) at the outlet of the PFC. The MOM implementation of the PFC model and the kinetics of ibuprofen crystallization are used in Eq. (6.26b) to calculate the objective function and the constraints of the optimization problem. There are three inequality constraints and one equality constraint included in the optimization problem.

The first inequality constraint, Eq. (6.26c), ensures that supersaturation S is within the metastable zone to avoid primary nucleation. This inequality constraint is important for two reasons: i) only the kinetics of secondary nucleation is provided by [172], and ii) the primary nucleation, which is a spontaneous process that happens in the region above the primary nucleation threshold (PNT) in the phase diagram as discussed by [172]. For industrial applications, engineers always attempt to avoid the *unstable* primary nucleation to prevent the creation of enormous amount of fine crystals which restrains the growth of crystals. Moreover, the PNT measured by [172] is used here. The second inequality constraint, Eq. (6.26d), ensures the yield of the crystallization process is above 95%. The third inequality constraint, Eq. (6.26e), avoids temperature increase as the kinetics for dissolution is not included in the model. The equality constraint Eq. (6.26f), in turn, calculates the value of d_{43} for the objective function. The PFC consists of 20 temperature controlling segments, where each segment is almost 1 m long, and the temperature of each segment is bounded within the range where the solubility information is available, i.e.; Eq. (6.26g) is fulfilled.

$$\min_{\mathbf{T}(\cdot)} \quad -d_{43}(z_f), \tag{6.26a}$$

subject to:

$$\text{Mathemical model:} \quad \text{Eqs.(6.16) to (6.18), (6.20) to (6.25)} \tag{6.26b}$$

$$\text{Inequality constraints:} \quad \frac{S(z)}{C^*(T(z))} \leq PNT \quad \forall z \in [0, z_f] \tag{6.26c}$$

$$\frac{C(0) - C(z_f)}{C(0) - C^*(10^\circ C)} \geq 95\% \tag{6.26d}$$

$$\frac{dT(z)}{dz} \leq 0 \quad \forall z \in [0, z_f] \tag{6.26e}$$

$$\text{Equality constraints:} \quad d_{43}(z_f) = \frac{\mu_4(z_f)}{\mu_3(z_f)} \tag{6.26f}$$

$$\text{Bounds:} \quad 10^\circ C \leq T(z) \leq 40^\circ C \quad \forall z \in [0, z_f] \tag{6.26g}$$

The case study is coded in MATLAB®(Version 2017b, The MathWorks Inc., Natick, Massachusetts, USA). The PFC model and the optimization problem are solved by using the functions $ode15s$ and $fmincon$, respectively. The PCE model is built with UQLaB (Version 1.0, ETH Zurich, Switzerland).

6.5 Results and discussion

First, the results of the nominal design are discussed. Then, the adverse effects of the parameter uncertainties, which are assumed based on the estimated value and CI from experiments, on the nominal process design are shown. To alleviate the influence of parameter uncertainties, the stochastic back-off strategy is then used for the robust design of the PFC. The double-loop back-off strategy and the stochastic back-off strategy using Monte Carlo simulations are also implemented as references. Finally, the convergence and computational demands of the different approaches are compared.

6.5.1 Nominal design

The saturated solution of ibuprofen in pure ethanol at $40°C$ is fed into the PFC with a total length of $z_f = 20.16$ m. The optimal temperature profile for the PFC, which includes 20 controlling segments, is derived by solving the nominal optimization problem given in Eq. (6.26). Results derived from the nominal design are depicted in Fig. 6.7. On the left side of Fig. 6.7, it is shown that the complete mass-based CSD at the reactor position $z_f = 20.16$ that was derived from Eq. (6.19) with $N = 100$. On the right side of Fig. 6.7, the evolution profiles of d_{43}, solution concentration C, and temperature T along the axis of the crystallizer are illustrated. To gain a better understanding of the results from the nominal design, the supersaturation profile is given in Fig. 6.8. The first part of the PFC has a relatively high reactor temperature and low supersaturation; i.e., fewer nuclei are generated, and the growth rate is maintained at a comparatively high value as indicated by the slope of the curve in Fig. 6.7B. Consequently, the consumption of ibuprofen in solution is also low, and solute concentration C does not decrease too much at the beginning. However, to achieve the desired yield at the end of the PFC, the temperature decreases gradually at higher amplitudes, and the supersaturation increases to its upper limit. As a result, the consumption of ibuprofen in the solution is increased. The predicted yield at the end of the PFC is 99.31%. With the nominal design, the final mass-based mean crystal size is maximized while all given constraints are satisfied. Another important CQA for the crystallization process, i.e, the coefficient of variation (CV) of the crystal size distribution according to Eq. (6.27), is also calculated. CV is the relation between the (mass-based) standard deviation and (mass-based) mean value of crystal size [186]. It provides an indicator of the spread of the crystals size distribution. It provides an indication about the consistence in the crystal size. The CV value for the PFC is equal to 0.21, which is much smaller than 0.5 for a single-stage MSMPRC, and reveals another important benefit of the PFC. In principle, the CV could also be used

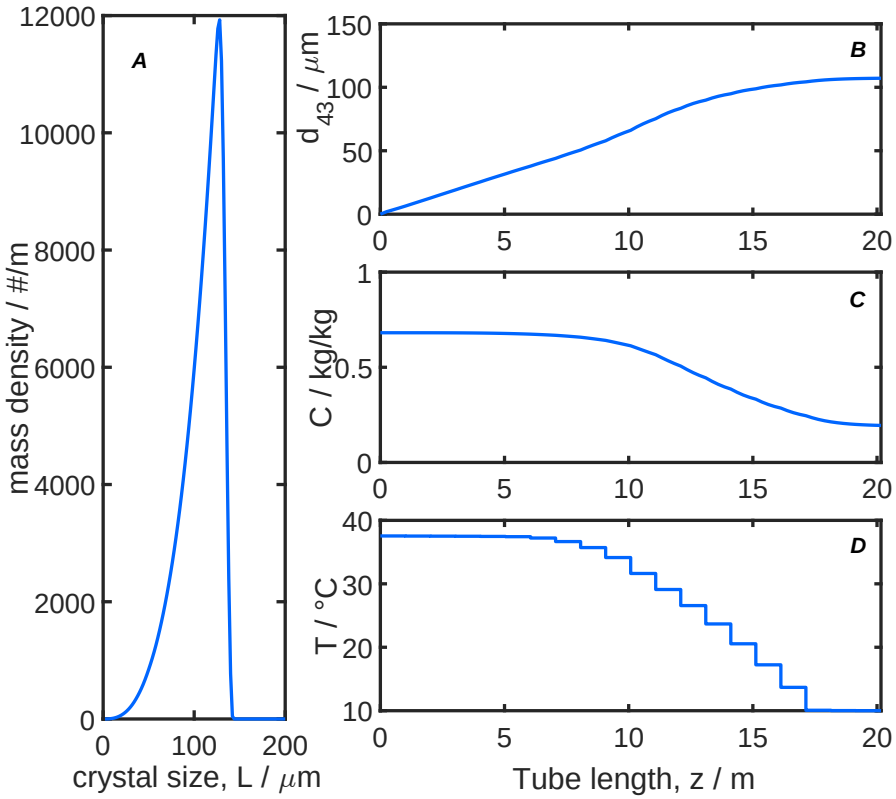


Figure 6.7: Results from nominal design of plug-flow crystallizer (PFC) for active pharmaceutical ingredient ibuprofen. ***A*** is the mass-based crystal size distribution (CSD), which is defined as $\rho_s L^3 n$ at the outlet of PFC. ***B***, ***C*** and ***D*** are the evolution profile of mass-based mean crystal size (d_{43}), mass concentration of solute ibuprofen (C) and operation temperature along the axis coordinate, respectively.

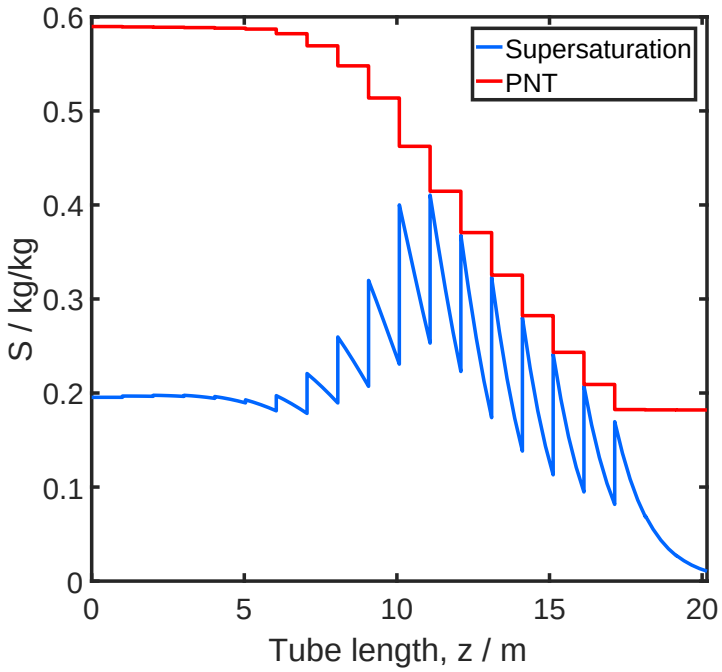


Figure 6.8: Profile of supersaturation (S) along the axis of the plug-flow crystallizer. The primary nucleation threshold PNT is the primary nucleation threshold.

Table 6.2: Uncertainties and feasible ranges of the kinetic parameters based on [172].

Parameters	Uncertainty	Range
k_{b_0}	$\mathcal{N}(1.73 \times 10^8, 2.6 \times 10^7)$	$[0, \infty]$
k_{g_0}	$\mathcal{N}(5.3, 0.69)$	$[0, \infty]$
T_g	$\mathcal{N}(42, 12.6)$	$[20, 65]$

directly as an objective function, but that is beyond the scope of this work.

$$CV = \sqrt{\frac{\mu_5 \mu_3}{\mu_4^2} - 1} \tag{6.27}$$

6.5.2 Effect of parameter uncertainties on the nominal design

According to [172], experimental data of the crystallization process are considerably affected by measurement noise and environmental conditions. The resulting data uncertainties lead to strong deviations in the estimation of the kinetic parameters for the nucleation and growth rates. Based on the estimated kinetic parameters and their confidence intervals (CIs) from [172], the parameter uncertainties of k_{b_0}, k_{g_0}, and T_g are summarized in Table 6.2 and are assumed with Gaussian distributions. Samples of the kinetic parameters are generated based on the assigned probability distributions, and simulations of the PFC model with the generated samples and the nominal optimal solution are conducted to analyze the influence of the parameter uncertainties on the constraints. Note that only the *soft* constraints, i.e., the inequality constraints in Eqs. (6.26c) and (6.26d), are affected by the parameter uncertainties as explained in Chapter 4. In Fig. 6.9(a), it is shown the evolution profile of the supersaturation and its 99% CI along the PFC axis. When considering parameter uncertainties, the supersaturation exceeds the PNT with a comparatively high probability. In other words, the inequality constraint in Eq. (6.26c) might be violated, and thus, lead to undesired primary nucleation. Fig. 6.9(b) presents the probability distribution of the yield of the PFC. The value of the yield also varies due to the parameter uncertainties, but the corresponding inequality constraint is still satisfied for all realizations. Thus, in the following, it is focused on designing a PFC under the condition of parameter uncertainties, so that the supersaturation does not exceed the PNT.

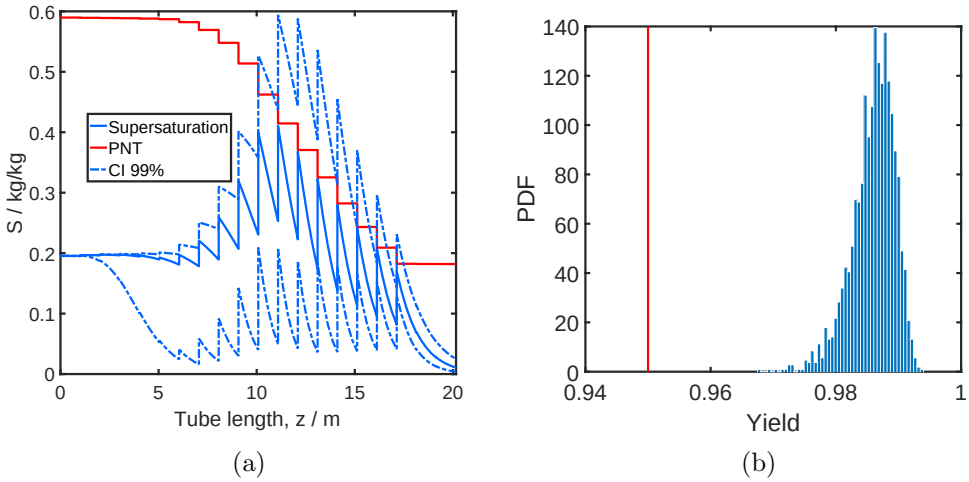


Figure 6.9: (a) Profile of supersaturation (S) and its 99% confidence interval(CI) along the axis of plug-flow crystallizer (PFC). PNT is the primary nucleation threshold. (b) Probability density function (PDF) of the yield.

6.5.3 Robust design with the stochastic back-off strategy

Robust optimization has the same structure as the nominal optimization in Eq. (A.1), except that the back-off terms are added to the inequality constraint in Eq. (6.26c) as:

$$S(z) + b(z) \leq PNT \times C^*(T(z)) \quad \forall z \in [0, z_f], \tag{6.28}$$

in which the value of back-off term $b(z)$ depends on the position in the PFC. In doing so, it is ensured that the supersaturation does not exceed the PNT in the presence of parameter uncertainties. The stochastic back-off strategy introduced in Section 6.3 is then used to solve the robust optimization problem. The maximum iteration number m_{max} is set to 10, the desired violation probability of the inequality constraints is set to 1%, and the convergence criterion ε to calculate back-off terms is set to 0.02. For the PCE model, the polynomial basis is constructed with the Stieltjes procedure integrated in UQLAB [122], as the boundaries on the parameter uncertainties change the structure of the distribution. The full set of the polynomial basis is truncated to the maximum order of 7, and 100 samples are used to estimate the PCE coefficients. For the KDE, a Gaussian kernel is assumed, an optimal bandwidth is determined [119], and 10,000 samples are used, respectively. Note that these 10,000 samples are evaluated with the PCE model, and thus, the computational costs of the KDE are negligible. The optimized

temperature profile is depicted and compared with the result from the nominal design in Fig. 6.10. As it can be seen, the temperature profile from the robust design is lower than that from the nominal design for the first half of the PFC and higher for the second half of the PFC to compensate the effect of the parameter uncertainties on the inequality constraint for supersaturation.

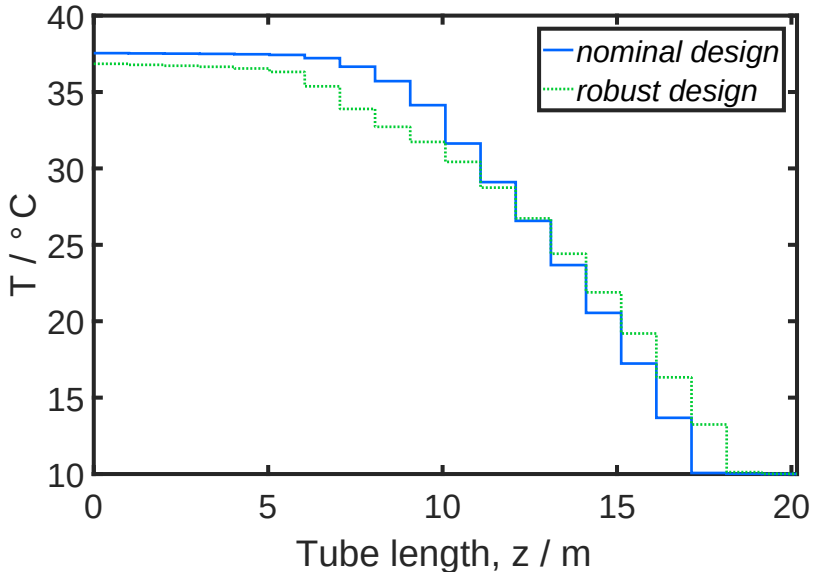
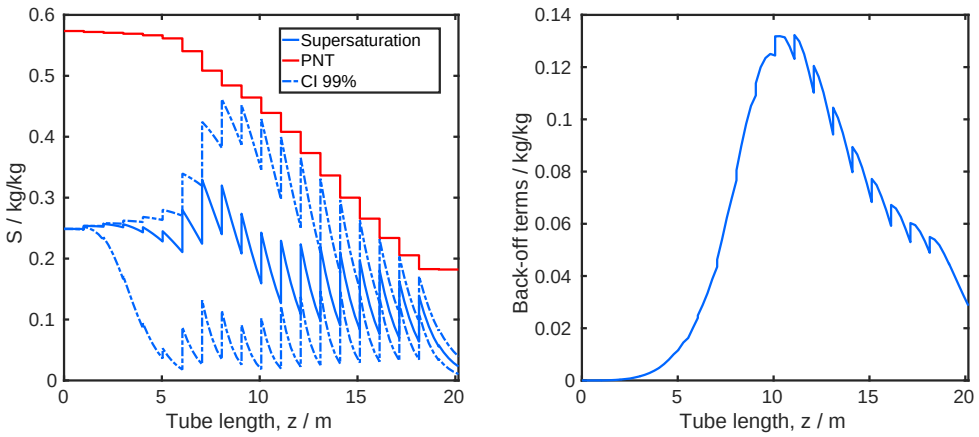


Figure 6.10: The operation temperature profiles from the nominal and robust designs.

In Fig. 6.11, the results of the inequality constraints from the robust design are further analyzed. By comparing the supersaturation profiles given in Fig. 6.11(a) and 6.9(a), it can be seen that the mean value of the supersaturation in the first half of the PFC is a bit higher than that from the nominal case. The robust design attempts to consume more ibuprofen solute and to generate more nuclei in the first half which lowers the supersaturation in the second half of the crystallizer. By doing so, the 99% CI of the supersaturation from the robust design is perfectly below the PNT. Moreover, the corresponding back-off terms are illustrated in Fig. 6.11(b). The magnitude of the back-off terms varies considerably along the PFC axis. Thus, the "time-varying" back-off term is more preferable than the constant back-off terms as the time-varying term provides more flexibility in the robust design.



(a) (b)

Figure 6.11: (a) Profile of supersaturation (S) and its 99% confidence interval(CI) and (b) profile of back-off terms along the axis of plug-flow crystallizer (PFC). *PNT* is the primary nucleation threshold. Results are from robust design of PFC with the stochastic back-off strategy based on polynomial chaos expansion.

6.5.4 Comparison of the different back-off strategies

In what follows, three different back-off strategies are compared in terms of their performance and efficiency. For the sake of readability, the double-loop back-off strategy is labeled *dlboPEM* as the PEM is used to calculate the back-off terms. The proposed stochastic back-off strategy is labeled *sboPCE*. The stochastic back-off strategy, in turn, where Monte Carlo simulations are used for calculating the back-off terms is labeled *sboMCs* and serves as the reference. The general setting of the optimization problem has not been changed; i.e., m_{max}, the desired violation probability, and the convergence criteria are the same as in the previous section.

First, the convergence results are depicted and compared in Fig. 6.12. The convergence of the internal and external loops for *dlboPEM* is shown in Fig. 6.12(a). The back-off terms converge after the first external iteration, but the desired probability of the constraint violations is not satisfied. Thus a second external iteration is required. Within the second external iteration, Eq. (6.8) is used to update the η value which leads to the specified performance of the robust PFC design. The proposed update concept (Eq. (6.8)) ensures a target-oriented and systematic correction of the η value, which is much more efficient than in those studies where different values for η are tested heuristi-

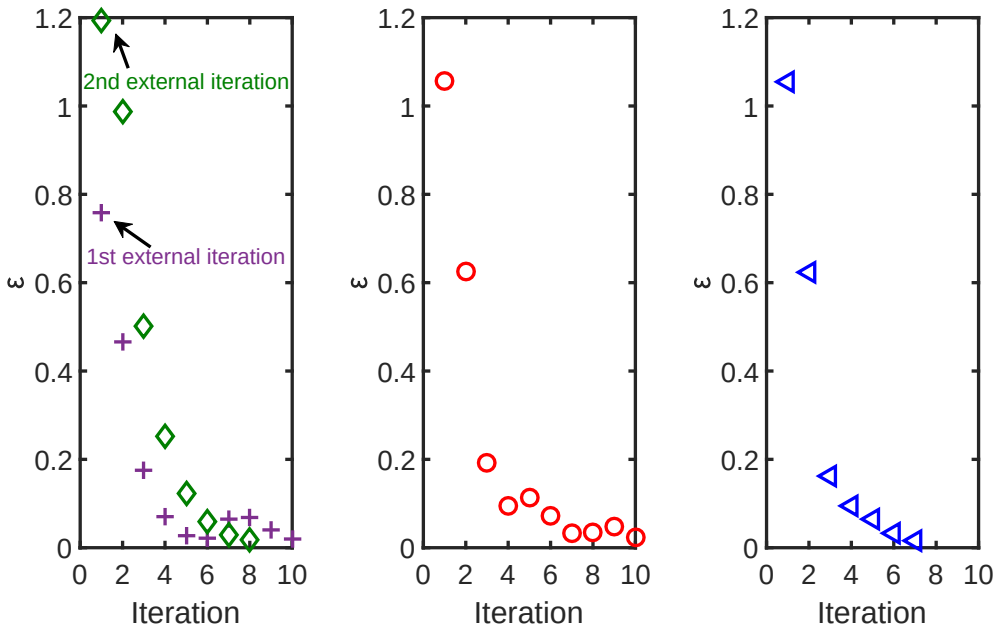


Figure 6.12: The convergence rates of the back-off values for (a) *dlboPEM*, (b) *sboMCs*, and (c) *sboPCE*.

cally [170, 181]. For stochastic back-off strategies, there is only one iteration loop. The convergence results for the single loop of *sboMCs* and *sboPCE* are shown in Fig. 6.12(b) and 6.12(c), respectively. *sboMCs* and *sboPCE* converge with a similar trend as they differ only in the detail of uncertainty propagation, i.e., the use Monte Carlo simulations or the PCE. *sboMCs* requires more iteration steps than *sboPCE*, which might be due to the randomness of the samples. The convergence plots in Fig. 6.12 reveal that all three approaches can ultimately converge to a robust solution. The operation temperature profiles obtained from the three back-off strategies are compared in Fig. 6.13. The temperature profiles are almost identical. Thus, all three robustification concepts that make use of the back-off strategy to converge to the same robust solution. Table 6.3

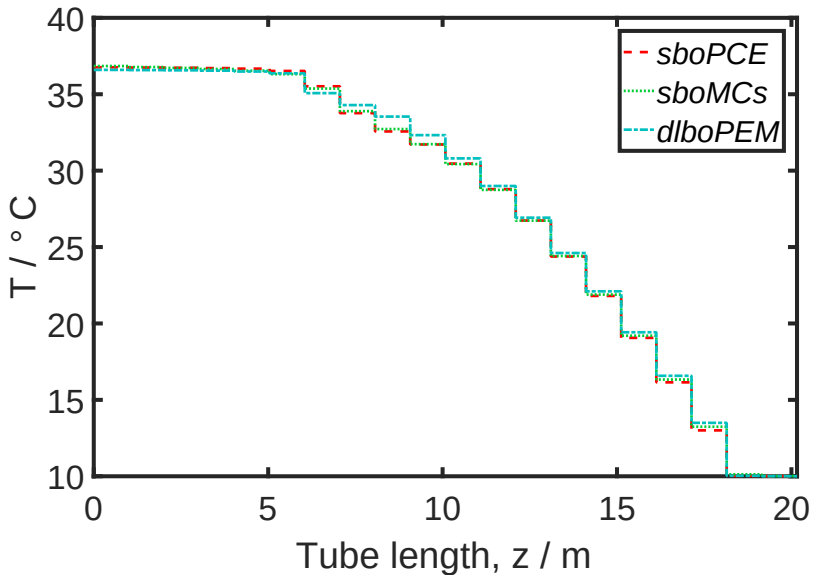


Figure 6.13: The operation temperature profiles from the nominal and robust designs.

lists more details about the PFC performance, i.e., the $\mathbb{E}(d_{43})$), the violation probability, and the computational costs for the different back-off strategies. The same quantities for the nominal design are also listed for the sake of completeness. Please note that the violation probability in Table 6.3 is only for the inequality constraint in Eq. (6.26c), and the computational costs is reflected by the number of evaluations of the original crystallizer model to calculate the back-off terms. The data in Table 6.3 are derived based on 10,000 realizations with the original PFC model. Results from the nominal design have the maximum $\mathbb{E}(d_{43})$ value, but the violation probability, i.e., 48%, is much higher than the given specification of 1%. In contrast, the results from the robust designs have much

Table 6.3: Results of the mean value of mass-based mean crystal size ($\mathbb{E}(d_{43})$), violation probability from 10000 realizations and computational cost with respect to the number of model evaluations.

Approaches	$\mathbb{E}(d_{43})$	Violation probability	Needed reference simulations
Nominal	109.3	48%	0
dlboPEM(1^{st} iter)	108.8	3%	10190
dlboPEM(2^{nd} iter)	108.6	1.1%	20342
sboMCs	108.7	1.2%	100000
sboPCE	108.7	0.9%	700

lower violation probabilities, which are close to the desired value of 1%, while there is only a slight performance loss of $\mathbb{E}(d_{43})$. For the *dlboPEM* implementation, the violation probability is almost three times higher than the desired value after the first external iteration and reduces to almost 1% after the second external iteration. The violation probabilities from the robust design with *sboMCs* and *sboPCE* are close to 1% after the "first external iteration". This indicates that the simpler structure of the proposed stochastic back-off strategy does not need additional CPU-intensive loop iterations to ensure the desired robustness in the inequality constraints. Moreover, *sboPCE* has the highest efficiency, which is more than 20 times faster than *dlboPEM* and *sboMCs*. In summary, all three back-off strategies can guarantee the robustness of the inequality constraints in the design of a PFC under parameter uncertainties. However, the PCE-based stochastic back-off strategy (*sboPCE*) has the best computational efficiency, and thus, it shows the perfect balance of process performance, robustness, and computational demand.

6.6 Chapter Summary

To guarantee reliable results in model-based process design, effective robustification concepts must be applied. For instance, the iterative back-off strategy proposed by [154] has received keen interests in academia and industry and has been implemented in various studies in robust optimization. In this chapter, the original procedure is improved and, especially, a novel stochastic back-off strategy is proposed with two key benefits: 1) a simpler structure and 2) higher efficiency. The stochastic back-off strategy calculates back-off terms with a distance function based on the probability distribution of the constraints. The performance of the adapted conventional back-off approach and

the stochastic back-off approach are studied for the crystallization of ibuprofen within a plug-flow crystallizer. To this end, the nominal process design is implemented, with which the adverse effects of the parameter uncertainties is illustrated. If the parameter uncertainties are ignored in the design phase aiming for the highest mass-based mean crystal size, an optimized temperature profile is derived which, most likely, causes extreme supersaturation constraint violations. Alternatively, it is successfully demonstrated that the stochastic back-off approach results in a temperature profile that shows the perfect balance of process performance and robustness, i.e., a high mass-based mean crystal size and constraint violations within the given specification. As the novel stochastic back-off concept takes the full information of the density function into account, the optimization needs only a single iteration loop to converge. Thus, the simpler structure of the stochastic back-off approach results in lower computational demands than the adapted conventional back-off approach. Moreover, it is also demonstrated that polynomial chaos expansion in combination with the kernel density estimator is essential for deriving meaningful probability density functions at low computational costs. Compared with standard Monte Carlo simulations, the overall need for CPU-intensive reference simulations is reduced considerably, i.e., the stochastic back-off strategy is at least 20 times faster. In general, this strategy also scales well with large-scale process design problems with many uncertain model parameters based on the recent progress in highly efficient PCE routines.

Chapter 7

Conclusions and Future Works

This thesis reveals the significance of parameter uncertainties in the process development of pharmaceutical manufacturing. Thereby, efficient and effective approaches have been proposed for analysis and design of the pharmaceutical processes with incorporating the information from parameter uncertainties. Moreover, the application studies considered in this thesis consists of (bio)chemical reactors, crystallization process and drying process, which cover the entire process chain of pharmaceutical manufacturing. Presented results show that ...

1) ... heterogeneity in the information from parameter uncertainties and process model can be quantitatively analyzed with global sensitivity analysis. The sensitivity results could provide decision-makers an in-depth understanding of interaction between parameters and process outputs to make proper decisions.

2) ... interferences from parameter uncertainties on process performance and critical constraints could be properly handled with robust process design to meet the requirements on product quality and process safety in pharmaceutical manufacturing.

3) ... the novel approaches proposed in this thesis provide efficient solution strategies for both process analysis and design. Moreover, the proposed approaches provide the possibility of incorporating more statistical information of parameter uncertainties, e.g., mutual dependencies and arbitrary probability distributions, which in turn also result in more reliable solutions for process analysis and design.

7.1 Summary

Performance of designs from model-based tools is hindered with the existence of parameter uncertainties in industrial applications. This necessitates the further development

on the methods in process engineering field, which are implemented for analysis and design of process units, in the presence of parameter uncertainties. Moreover, fierce market competition in pharmaceutical manufacturing is the main driving-force for them to modernize their current production process and follow the QbD initiative proposed by regulatory agencies. Thereby, the target of this thesis is to investigate and develop proper methods for sensitivity analysis and robust process design, and implement them in the development of pharmaceutical processes, as described in chapter 1. The structure of the thesis is outlined in Fig. 7.1 and the details are summarized as follows.

In order to be able to incorporate the information of parameter uncertainties in the process analysis and design, efficient methods for uncertainty propagation and quantification (UQ) are needed. In chapter 2, methods for UQ are presented and compared, in which the Point Estimate Method (PEM) and Polynomial Chaos Expansion (PCE) manifest their superiority in terms of computational load and are used in this thesis. The basics of PEM and PCE are addressed. Transformation technique is proposed to generalize the PEM to non-Gaussian distributions and parameter dependencies. Moreover, techniques for generalization of PCE and efficient algorithm for estimating the coefficients were also presented. The computational load for PEM is the lowest among these methods and scales quadratically with the problem size. It is implemented in chapter 4, 5 and 6 for UQ. PCE has moderate efficiency but could approximate the exact shape of probability distributions of process outputs, and is used in chapter 3 and 6.

Global sensitivity analysis (GSA) aims to quantitatively measure the influence of parameter uncertainties on process outputs. Two GSA techniques with desired features were introduced in chapter 3, i.e., (co)variance-based and moment-independent approaches. The impact of parameter dependencies is also investigated, which are described by the Gaussian copula. PCE is used here for UQ, as information about the entire probability distributions of model outputs are required for moment-independent approach. The approaches are demonstrated with a continuous synthesis of API-Scaffold. The results show that both approaches are able to identify the heterogeneity in the information about the influence of parameter uncertainties on process outputs. Additionally, parameter dependencies are crucial for identifying the most relevant parameter uncertainties and also affect the variation in the process outputs. The obtained sensitivity measures enhance process understanding and could be used to assist process design.

After evaluating the influence of parameter uncertainties, the ones with appreciable sensitivities have to be considered in robust process design. In chapter 4, a comprehensive and efficient framework for robust process design is introduced. Robust formulation for both inequality and equality constraints are discussed and approximated with statistical

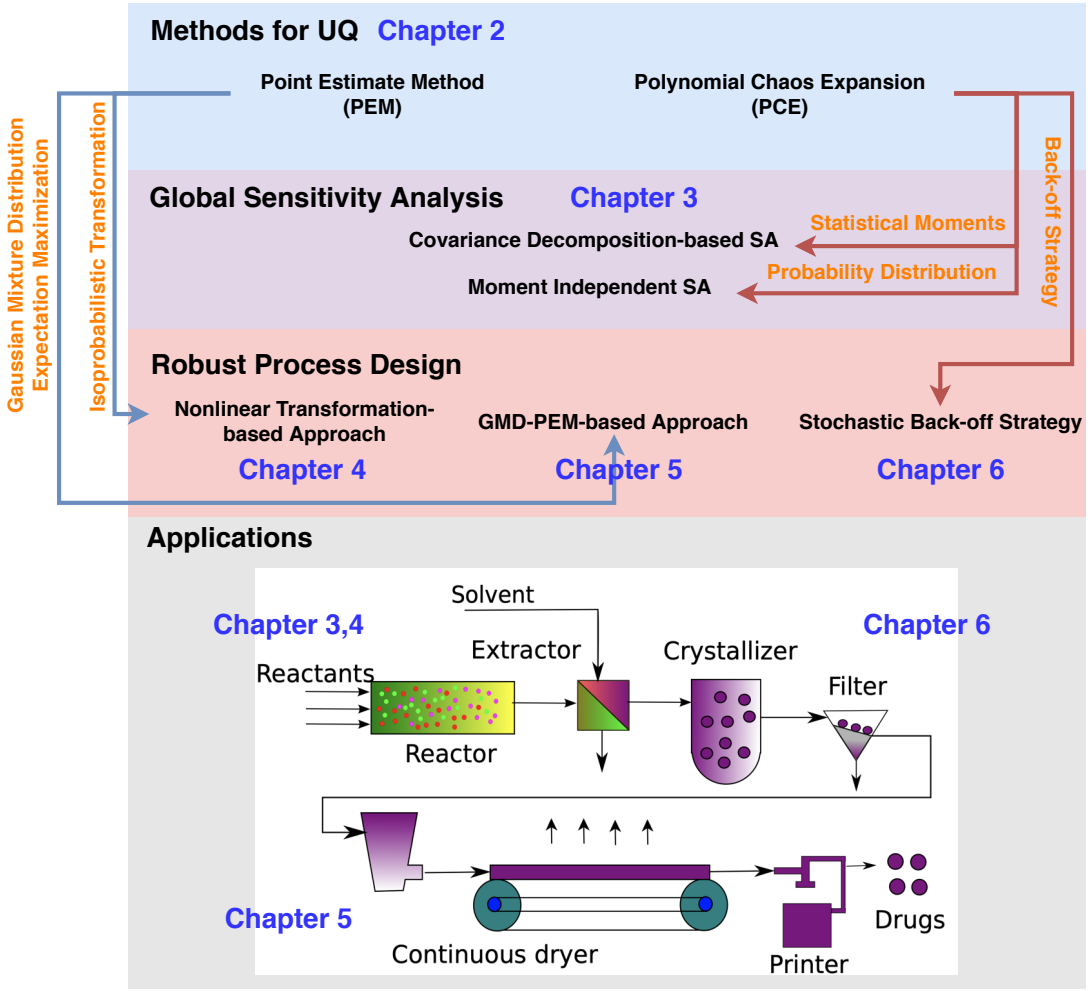


Figure 7.1: Summary of the thesis structure

moments. Especially, the robust equality constraints, which are ill-posed in nature, are reformulated. Here, the statistical moments are approximated with PEM given in chapter 2 in the absence and presented of parameter dependencies. Two application studies, which are related to the synthesis block in pharmaceutical manufacturing, are investigated. The proposed framework provides solutions with the right trade-off between the process performance and robustness, especially the constraints on reactor temperature, biomass and substrate concentration are satisfied in most realizations. The results show that second-order moment method is more preferable for approximating the robust constraints, as the accuracy of higher-order moments is not guaranteed with PEM. However, the violation probability of inequality constraints in both studies are still higher than the expected values. This is because the variation in model outputs are approximated only with the first two moments and assumed to distributed symmetrically. One straightforward solution for this problem is adapting the tolerance factors. More rigorous solution strategies are introduced in chapter 5 and chapter 6. The impact of parameter dependencies manifest also significance on the results from robust process design. Moreover, in the first case study, the proposed formulation for robust inequality constraints is used for reducing the variation in outlet temperature. Dramatical decrease in the final product concentration is observed with narrowing down the relaxed boundaries. In the second study, GSA is implemented and reduces the computational demand of robust design about 80%, which reflects the importance of GSA introduced in chapter 3.

In chapter 5, the proposed framework is further extended to the design of a downstream process, i.e., the freeze-drying process (lyophilization), in which non-Gaussian parameter uncertainties are involved. The isoprobabilistic transformation based PEM presented in chapter 2 is used to generate associated PEM samples for the non-Gaussian parameter uncertainties. However, the additional complexity introduced by the nonlinear transformation step as well as the Gaussian approximation of model output distribution, which is also mentioned in the preview paragraph, deteriorate its approximation accuracy of the variation in process outputs. Alternatively, another approach, in which the Gaussian mixture distribution (GMD) is combined with PEM for UQ, is proposed for the first time for robust process design with non-Gaussian distribution. The non-Gaussian distribution could be decomposed into several Gaussian distributions, for which the PEM samples can be generated with linear transformation. The component distributions are then propagated through the process model individually and the model outputs are gathered and used for robust process design. Both approaches are implemented on the case study and compared with a reference approach, in which the the parameter uncertainties are simplified and approximated with single multivariate Gaussian distribution. The results

reveal that GMD-PEM has the best accuracy among the three approaches in approximating the variation in model outputs for robust process design. It provides the solution, which is neither too conservative nor too risky in this case study. With the proposed GMD-PEM approach, robust profiles for the shelf temperature and chamber pressure are obtained. With the profiles, the efficiency of the drying process is maximized, while the temperature T_i and its 99% confidence intervals are constrained in its appropriate region to avoid irreversible product damage.

In chapter 6, a major limitation on the robust design framework used in chapter 4 and 5 is addressed. That is conducting UQ at each optimization iteration could be still computationally burdensome if the optimization problem is inappropriately formulated. Therefore, the iterative back-off strategy proposed in [154] is used to develop novel robust design framework, which is independent of the formulation of the optimization problem. The double-loop back-off strategy is proposed firstly, in which the PEM is used in the internal loop to calculate the value of back-off terms, and the tolerance factor is adjusted systematically in the external loop to meet the required robustness of inequality constraints. However, the double loops could be troublesome for the convergence of back-off terms and updating the tolerance factor is also non-trivial. Therefore, a novel stochastic back-off strategy is proposed for the first time for robust process design. It has only one iteration loop, in which the PCE is used to calculate the back-off terms with low computational expense. The advantage of using PCE is that the exact shape of probability distributions of process outputs can be approximated and used for determination of back-off terms. These two back-off strategies are demonstrated on the robust design of continuous crystallization process of Ibuprofen, which is an essential separation process for drug production. The results show that both approaches are able to determine the robust profile of cooling temperature for the continuous crystallization process. With this, the mass-based mean crystal size is maximized, while the supersaturation and its 99% confidence interval are ensured to be lower than the primary nucleation threshold. But, the computational efficient of the novel stochastic approach is at least 20 times higher than others.

7.2 Further Works

In future work, the following aspects could be further investigated.

1) The model-based process development is classified to 3 stages, as shown in Fig. 1.3. The focus of the thesis is about stage 2 and 3. That is analysis and design

of various unit operations in the presence of parameter uncertainties. It will be also interesting and important to extend the GSA techniques and robust design frameworks to stage 1, in which the structure of the holistic process is designed and mutual interactions between the unit operations are incorporated.

2) The parameter dependencies considered in this work are described with Gaussian copula, which is available for the dependence structure without tail dependence. For more complex dependence structure between parameter uncertainties, other types of copula might be interesting, such as Archimedean copulas [112], and could be straightforwardly incorporated with the frameworks proposed in this thesis.

3) In this thesis, the parameter uncertainties are described with well-defined probability distributions. However, there could exist ambiguities on the probability distribution itself due to some reasons, such as the degradation of equipment. Therefore, it is interesting to include the information heterogeneity about the uncertainty itself in process analysis and design. Moreover, the concept of parametric probability-boxes (p-boxes) introduced in [188] for describing imprecise uncertainties might be interesting and could be incorporated with the proposed frameworks.

4) The concepts and ideas implemented to GSA techniques and robust design frameworks could also be extended to other process engineer techniques, such as control strategy design, optimal experimental design [6] and active fault detection and diagnose [5]. Moreover, they could also be straightforwardly extended to other industrial sectors.

Appendix A

A.1 Additional Results for Sensitivity Analysis with Different Correlation Coefficients [7]

The results of additional four cases with fictitious correlation coefficients are presented here, where the case 1,2,3, and 4 listed in Table A.1 are the same with the cases depicted in Fig. 3.10 and Fig. A.1.

Table A.1: Sensitivity results for cases with different correlation coefficients. The correlation coefficients for $\theta_1, \theta_2, \theta_3, \theta_4$, θ_7, θ_8 are assigned with 0.5 and 0.9 for case 1 and 2, respectively. The correlation coefficients for $\theta_1, \theta_2, \theta_3, \theta_4$ and θ_7, θ_8 are assigned with 0.5 and 0.9 for the case 3, and with 0.9 and 0.2 for the case 4.

		θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8
S^{cov}	1	0.258	0.184	0.260	0.182	0	0	0.030	0.062
	2	0.242	0.201	0.280	0.168	0	0	-0.026	0.124
	3	0.276	0.197	0.278	0.195	0	0	-0.005	0.026
	4	0.143	0.121	0.168	0.099	0	0	0.175	0.279
δ^{cor}	1	0.126	0.105	0.127	0.104	0.015	0.015	0.036	0.061
	2	0.054	0.048	0.058	0.045	0.016	0.014	0.021	0.057
	3	0.132	0.110	0.133	0.108	0.015	0.014	0.016	0.029
	4	0.045	0.042	0.047	0.038	0.016	0.017	0.117	0.161

As we mentioned in Section 3.5.5, the output variances depend considerably on the parameter correlation values. Although our primary focus was on correlation coefficients derived with experimental data, it might be still interesting to demonstrate the effect of parameter correlations which are less dominating; i.e., not that close to one. Here, we consider two additional test cases with fictitious correlation values: 0.5 and 0.9. CoDSA

[7] Part of this appendix has been published in (Xie et al., Reliab. Eng. Syst. Safe., 187, 159–173, 2019 [9])

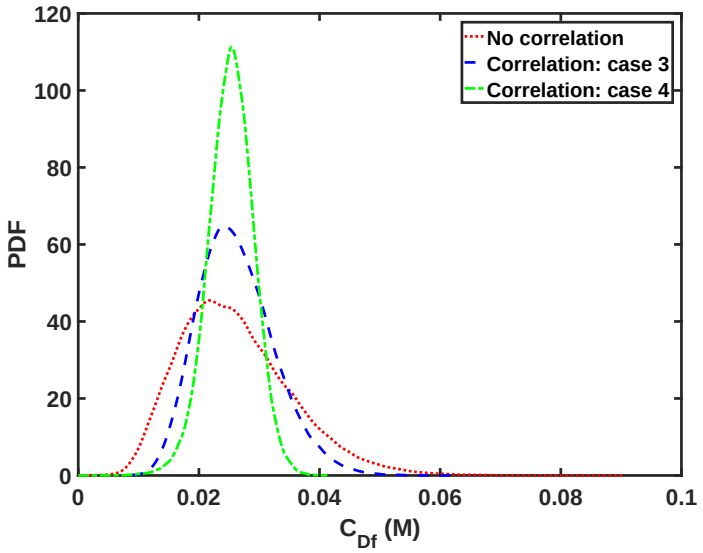


Figure A.1: Comparison of the resulting probability density functions of C_{Df} in the absence and presence of parameter correlations. The correlation coefficients for $\theta_1, \theta_2, \theta_3, \theta_4$ and θ_7, θ_8 are assigned with 0.5 and 0.9 for case 4, and with 0.9 and 0.2 for case 5.

and MISA results are listed in Table A.1. Obviously, for both sensitivity measures the importance of parameters θ_7 and θ_8 increases gradually with higher correlation coefficients. The overall effect of θ_1,θ_2,θ_3, and θ_4 still dominates the output variation but recedes if the correlation coefficients increase further as the case with experimentally derived correlations. Furthermore, we include two additional cases 3 and 4, where the correlation coefficients for $\theta_1, \theta_2, \theta_3, \theta_4$ and θ_7, θ_8 are allocated with different values. The shapes of the output distribution are similar, see Fig. A.1. The sensitivity results, however, are completely different, especially for the most significant parameter. The sensitivity measure from CoDSA and MISA are analogue with minor differences. For instance, the sensitivity of θ_3 and θ_7 for case 5 are similar for CoDSA, while the sensitivity of θ_7 is twice as large compared to θ_3 for MISA. The reason is that MISA takes into account not only the output variance but also higher statistical moments, i.e., the entire output distribution. In Fig. A.2, we see that the shift of the conditional distributions is mainly due to the change of other moments, e.g., kurtosis, but not the variance which supports our conclusions considering MISA as a valuable tool in sensitivity analysis.

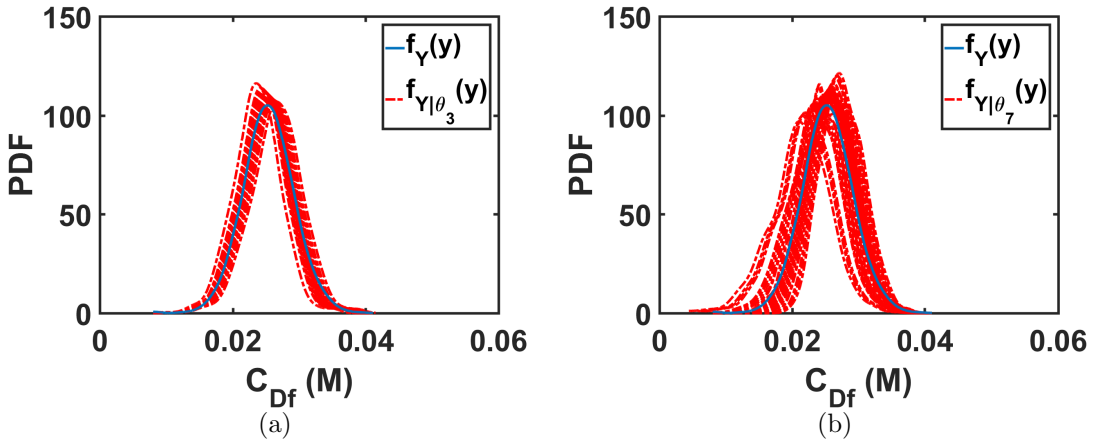


Figure A.2: Comparison of unconditioned distributions (blue line) and conditioned distributions (red lines) of concentrations of component D considering parameters θ_3 and θ_7 in case 5.

A.2 Structure of Nominal Optimization of the Primary Drying Process [8]

The structure of the nominal optimization of the primary drying process is given below. Chamber pressure P_c and shelf temperature T_s are manipulated to minimize final drying time t_f (A.1a), with which the frozen product can be completely dried (A.1g), and satisfy the constraints for the CQA (A.1e) and the technical limitation (A.1f):

$$\min_{\mathbf{T_i}(\cdot),\mathbf{m_{sub}}(\cdot),\mathbf{T_s}(\cdot),\mathbf{P_c}(\cdot)} \quad t_f, \tag{A.1a}$$

subject to:

Mathemical model: Eqs.(5.22) to (5.28) (A.1b)

Bounds: $P_c^L \leq P_c(t) \leq P_c^U$ (A.1c)

$T_s^L \leq T_s(t) \leq T_s^U$ (A.1d)

Inequality constraints: $T_i(t) \leq T_c$ (A.1e)

$\dot{m}_{sub}(t) \leq \dot{m}_{sub,choke,vial}$ (A.1f)

Equality constraints: $m_{sub}(t_f) = \rho_I \epsilon A_p L_{total}$ (A.1g)

[8] Part of this appendix has been published in (Xie et al., Chem. Eng. Sci., 207, 805-819, 2019 [10])

Bibliography

[1] Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Robust design of chemical processes based on a one–shot sparse polynomial chaos expansion concept. *Computer Aided Chemical Engineering*, 40(1):613–618, 2017.

[2] Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Toward a comprehensive and efficient robust optimization framework for (bio)chemical processes. *Processes*, 6(10):183, 2018.

[3] Xiangzhong Xie, Ulrike Krewer, and René Schenkendorf. Robust optimization of dynamical systems with correlated random variables using the point estimate method. *IFAC-PapersOnLine*, 51(2):427–432, 2018.

[4] Xiangzhong Xie, Rüdiger Ohs, Antje Spieß, Ulrike Krewer, and René Schenkendorf. Moment-independent sensitivity analysis of enzyme-catalyzed reactions with correlated model parameters. *IFAC-PapersOnLine*, 51(2):753–758, 2018.

[5] René Schenkendorf, Xiangzhong Xie, and Ulrike Krewer. An efficient polynomial chaos expansion strategy for active fault identification of chemical processes. *Computers & Chemical Engineering*, 122:228–237, 2018.

[6] René Schenkendorf, Xiangzhong Xie, Moritz Rehbein, Stephan Scholl, and Ulrike Krewer. The impact of global sensitivities and design measures in model-based optimal experimental design. *Processes*, 6(4):27, 2018.

[7] Nan Lin, Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Efficient global sensitivity analysis of 3d multiphysics model for li-ion batteries. *Journal of The Electrochemical Society*, 165(7):A1169–A1183, 2018.

[8] Victor N. Emenike, Xiangzhong Xie, René Schenkendorf, Antje C. Spiess, and Ulrike Krewer. Robust dynamic optimization of enzyme-catalyzed carboligation: A point estimate-based back-off approach. *Computers & Chemical Engineering*, 121:232–247, 2019.

[9] Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Efficient sensitivity analysis and interpretation of parameter correlations in chemical engineering. *Reliability Engineering & System Safety*, 187:159–173, 2019.

[10] Xiangzhong Xie and René Schenkendorf. Robust optimization of a pharmaceutical freeze-drying process under non-gaussian parameter uncertainties. *Chemical Engineering Science*, 207:805–819, 2019.

[11] Xiangzhong Xie and René Schenkendorf. Stochastic back-off robust optimization for continuous pharmaceutical processes using polynomial chaos expansion. *Computers & Chemical Engineering*, 124:80–92, 2019.

[12] Xiangzhong Xie and René Schenkendorf. Robust process design in pharmaceutical manufacturing under batch-to-batch variation. *Processes*, 7(8):509, 2019.

[13] Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Robust design of chemical processes based on a one-shot sparse polynomial chaos expansion concept. In *10th World Congress of Chemical Engineering (WCCE10) & 27th European Symposium on Computer Aided Process Engineering (ESCAPE 27)*, Barcelona, Spain, 1 - 5 October, 2017.

[14] René Schenkendorf, Xiangzhong Xie, and Ulrike Krewer. An efficient polynomial chaos expansion strategy for active fault identification of chemical processes. In *10th World Congress of Chemical Engineering (WCCE10) & 27th European Symposium on Computer Aided Process Engineering (ESCAPE 27)*, Barcelona, Spain, 1 - 5 October, 2017.

[15] Xiangzhong Xie, Ulrike Krewer, and René Schenkendorf. Robust optimization of dynamical systems with correlated random variables using the point estimate method. In *9th Vienna International Conference on Mathematical Modelling (MATHMOD)*, Vienna, Austria, 21 - 23 February, 2018.

[16] Xiangzhong Xie, Rüdiger Ohs, Antje Spieß, Ulrike Krewer, and René Schenkendorf. Moment-independent sensitivity analysis of enzyme-catalyzed reactions with correlated model parameters. In *9th Vienna International Conference on Mathematical Modelling (MATHMOD)*, Vienna, Austria, 21 - 23 February, 2018.

[17] Xiangzhong Xie and René Schenkendorf. The effect of correlated kinetic parameters on (bio)chemical reaction networks. In *DECHEMA - Jahrestreffen Reaktionstechnik*, Würzburg, Germany, 7 - 9 May, 2018.

[18] Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Robustes Prozessdesign in der Pharmatechnik mittels performanter Ersatzfunktionen. In *ProcessNet-Jahrestagung und 33. DECHEMA-Jahrestagung der Biotechnologen*, Aachen, Germany, 10 - 13 September, 2018.

[19] René Schenkendorf, Xiangzhong Xie, and Ulrike Krewer. Robustifizierung und Informationsmetriken der modellgestützten Versuchsplanung. In *ProcessNet Jahrestagung und 33. DECHEMA Jahrestagung der Biotechnologen*, Aachen, Germany, 10 - 13 September, 2018.

[20] Yongkui Liu and Xun Xu. Industry 4.0 and cloud manufacturing: A comparative analysis. *Journal of Manufacturing Science and Engineering*, 139(3):034701, 2017.

[21] Thomas F Edgar, David Mautner Himmelblau, Leon S Lasdon, et al. *Optimization of chemical processes*. McGraw-Hill, New York, NY, USA, 2001.

[22] Lorenz T Biegler. *Nonlinear programming: concepts, algorithms, and applications to chemical processes*. SIAM, 2010.

[23] William L Luyben. *Process modeling, simulation and control for chemical engineers*. McGraw-Hill, New York, NY, USA, 1989.

[24] Brian Roffel and Ben Betlem. *Process dynamics and control: modeling for control and prediction*. John Wiley & Sons, Chichester (UK), 2007.

[25] Brahim Benyahia, Richard Lakerveld, and Paul I Barton. A plant-wide dynamic model of a continuous pharmaceutical process. *Industrial & engineering chemistry research*, 51(47):15393–15412, 2012.

[26] Ashley F Emery and Aleksey V Nenarokomov. Optimal experiment design. *Measurement Science and Technology*, 9(6):864, 1998.

[27] René Schenkendorf. *Optimal experimental design for paramter identification and model selection*. PhD thesis, ,Otto-von-Guericke-Universität, Magdeburg, 2014.

[28] Ignacio E Grossmann and RWH Sargent. Optimum design of chemical plants with uncertain parameters. *AIChE Journal*, 24(6):1021–1028, 1978.

[29] Nikolaos V Sahinidis. Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering*, 28(6):971–983, 2004.

[30] Dries Telen, Mattia Vallerio, Lorenzo Cabianca, Boris Houska, Jan Van Impe, and Filip Logist. Approximate robust optimization of nonlinear systems under parametric uncertainty and process noise. *Journal of Process Control*, 33:140–154, 2015.

[31] Salvatore Mascia, Patrick L Heider, Haitao Zhang, Richard Lakerveld, Brahim Benyahia, Paul I Barton, Richard D Braatz, Charles L Cooney, James MB Evans, Timothy F Jamison, et al. End-to-end continuous manufacturing of pharmaceuticals: integrated synthesis, purification, and final dosage formation. *Angewandte Chemie*, 125(47):12585–12589, 2013.

[32] Albert E Cervera-Padrell, Tommy Skovby, Søren Kiil, Rafiqul Gani, and Krist V Gernaey. Active pharmaceutical ingredient (api) production involving continuous processes–a process system engineering (pse)-assisted design framework. *European Journal of Pharmaceutics and Biopharmaceutics*, 82(2):437–456, 2012.

[33] Krist V Gernaey, Albert E Cervera-Padrell, and John M Woodley. A perspective on pse in pharmaceutical process development and innovation. *Computers & Chemical Engineering*, 42:15–29, 2012.

[34] Richard Lakerveld, Brahim Benyahia, Richard D Braatz, and Paul I Barton. Model-based design of a plant-wide control strategy for a continuous pharmaceutical plant. *AIChE Journal*, 59(10):3671–3685, 2013.

[35] Victor N Emenike, René Schenkendorf, and Ulrike Krewer. A systematic reactor design approach for the synthesis of active pharmaceutical ingredients. *European Journal of Pharmaceutics and Biopharmaceutics*, 126:75–88, 2018.

[36] Qinglin Su, Zoltan K Nagy, and Chris D Rielly. Pharmaceutical crystallisation processes from batch to continuous operation using msmpr stages: Modelling, design, and control. *Chemical Engineering and Processing: Process Intensification*, 89:41–53, 2015.

[37] Fani Boukouvala, Vasilios Niotis, Rohit Ramachandran, Fernando J Muzzio, and Marianthi G Ierapetritou. An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process. *Computers & Chemical Engineering*, 42:30–47, 2012.

[38] Terrence F Yee and Ignacio E Grossmann. Simultaneous optimization models for heat integration—ii. heat exchanger network synthesis. *Computers & Chemical Engineering*, 14(10):1165–1184, 1990.

[39] Paul C Collins. Chemical engineering and the culmination of quality by design in pharmaceuticals. *AIChE Journal*, 64(5):1502–1510, 2018.

[40] Barbara K Immel. A brief history of the gmps for pharmaceuticals. *Pharmaceutical technology*, 25(7):44–53, 2001.

[41] John P Swann. The 1941 sulfathiazole disaster and the birth of good manufacturing practices. *Pharmacy in history*, 41(1):16–25, 1999.

[42] Emanuele Tomba, Pierantonio Facco, Fabrizio Bezzo, and Massimiliano Barolo. Latent variable modeling to assist the implementation of quality-by-design paradigms in pharmaceutical development and manufacturing: a review. *International journal of pharmaceutics*, 457(1):283–297, 2013.

[43] FDA. Pharmaceutical cgmps for the 21st century—a risk-based approach. *Final report. U.S. Food and Drug Administration, Rockville, MD*, 2004.

[44] FDA. Guidance for industry, pat-a framework for innovative pharmaceutical development, manufacturing and quality assurance. *U.S. Food and Drug Administration, Rockville, MD*, 2004.

[45] Janet Woodcock. The concept of pharmaceutical quality. *Am Pharm Rev*, 7(6):10–15, 2004.

[46] Levente L Simon, Hajnalka Pataki, György Marosi, Fabian Meemken, Konrad Hungerbühler, Alfons Baiker, Srinivas Tummala, Brian Glennon, Martin Kuentz, Gerry Steele, et al. Assessment of recent process analytical technology (pat) trends: a multiauthor review. *Organic Process Research & Development*, 19(1):3–62, 2015.

[47] ICH. Quality risk management q9. *Guideline, ICH Harmonised Tripartite*, 4:408, 2005.

[48] ICH. Pharmaceutical development q8. *Guideline, ICH Harmonised Tripartite*, 4:11, 2005.

[49] ICH. Pharmaceutical quality system q10. *Guideline, ICH Harmonised Tripartite*, 4, 2008.

[50] ICH. Pharmaceutical development q8(r2). *Guideline, ICH Harmonised Tripartite*, 2009.

[51] Andrea Saltelli, Karen Chan, E Marian Scott, et al. *Sensitivity analysis*, volume 1. Wiley New York, 2000.

[52] Andrea Saltelli, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, and Qiongli Wu. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental modelling & software*, 114:29–39, 2019.

[53] Yunfei Chu and Juergen Hahn. Necessary condition for applying experimental design criteria to global sensitivity analysis results. *Computers & Chemical Engineering*, 48:280–292, 2013.

[54] René Schenkendorf. A General Framework for Uncertainty Propagation Based on Point Estimate Methods. In *Second European Conference of the Prognostics and Health Management Society, PHME14, Nantes, France*, 2014.

[55] Emanuele Borgonovo and Elmar Plischke. Sensitivity analysis: a review of recent advances. *European Journal of Operational Research*, 248(3):869–887, 2016.

[56] Tamás Turányi. Sensitivity analysis of complex kinetic systems. tools and applications. *Journal of Mathematical Chemistry*, 5(3):203–248, 1990.

[57] Andrea Saltelli, Marco Ratto, Stefano Tarantola, and Francesca Campolongo. Sensitivity analysis for chemical models. *Chemical reviews*, 105(7):2811–2828, 2005.

[58] RW Atherton, RB Schainker, and ER Ducot. On the statistical sensitivity analysis of models for chemical kinetics. *AIChE Journal*, 21(3):441–448, 1975.

[59] MPR Haaker and PJT Verheijen. Local and global sensitivity analysis for a reactor design with parameter uncertainty. *Chemical Engineering Research and Design*, 82(5):591–598, 2004.

[60] Judit Zádor, I Gy Zsely, and Tamás Turányi. Local and global uncertainty analysis of complex chemical kinetic systems. *Reliability Engineering & System Safety*, 91(10):1232–1240, 2006.

[61] Géraud Blatman. *Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis*. PhD thesis, Clermont-Ferrand 2, 2009.

[62] Emanuele Borgonovo. A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784, 2007.

[63] Brandon J Reizman and Klavs F Jensen. An automated continuous-flow platform for the estimation of multistep reaction kinetics. *Organic Process Research & Development*, 16(11):1770–1782, 2012.

[64] Alfredo López-Benito and Ricardo Bolado-Lavín. A case study on global sensitivity analysis with dependent inputs: The natural gas transmission model. *Reliability Engineering & System Safety*, 165:11–21, 2017.

[65] Éva Valkó, Tamás Varga, Alison Tomlin, and Tamas Turányi. Investigation of the effect of correlated uncertain rate parameters via the calculation of global and local sensitivity indices. *Journal of Mathematical Chemistry*, pages 1–26, 2017.

[66] Genichi Taguchi, Don Clausing, et al. Robust quality. *Harvard Business Review*, 68(1):65–75, 1990.

[67] Mattia Vallerio, Dries Telen, Lorenzo Cabianca, Flavio Manenti, Jan Van Impe, and Filip Logist. Robust multi-objective dynamic optimization of chemical processes using the sigma point method. *Chemical Engineering Science*, 140:201–216, 2016.

[68] Xiangzhong Xie, René Schenkendorf, and Ulrike Krewer. Robust design of chemical processes based on a one-shot sparse polynomial chaos expansion concept. *Computer Aided Chemical Engineering*, 40:613–618, 2017.

[69] Zoltán K Nagy and Richard D Braatz. Worst-case and distributional robustness analysis of finite-time control trajectories for nonlinear distributed parameter systems. *IEEE Transactions on Control Systems Technology*, 11(5):694–704, 2003.

[70] Séverine Thérèse FC Mortier, Pieter-Jan Van Bockstal, Jos Corver, Ingmar Nopens, Krist V Gernaey, and Thomas De Beer. Uncertainty analysis as essential step in the establishment of the dynamic design space of primary drying during freeze-drying. *European Journal of Pharmaceutics and Biopharmaceutics*, 103:71–83, 2016.

[71] Zoltan K Nagy and Richard D Braatz. Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis. *Journal of process control*, 14(4):411–422, 2004.

[72] Amith Singhee and Rob A Rutenbar. Why quasi-monte carlo is better than monte carlo or latin hypercube sampling for statistical circuit analysis. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 29(11):1763–1776, 2010.

[73] Jun Shi, Lorenz T Biegler, Intan Hamdan, and John Wassick. Optimization of grade transitions in polyethylene solution polymerization process under uncertainty. *Computers & Chemical Engineering*, 95:260–279, 2016.

[74] René Schenkendorf. A general framework for uncertainty propagation based on point estimate methods. In *Second european conference of the prognostics and health management society, phme14. Nantes, France*. Citeseer, 2014.

[75] Thierry Crestaux, Olivier Le Maıˆtre, and Jean-Marc Martinez. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7):1161–1172, 2009.

[76] Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical Science*, pages 273–304, 1995.

[77] Dongbin Xiu. *Numerical methods for stochastic computations: a spectral method approach*. Princeton University Press, 2010.

[78] Christopher Z Mooney. *Monte carlo simulation*, volume 116. Sage Publications, Thousand Oaks, CA, 1997.

[79] Gene H Golub and John H Welsch. Calculation of gauss quadrature rules. *Mathematics of computation*, 23(106):221–230, 1969.

[80] Uri N Lerner. *Hybrid Bayesian networks for reasoning about complex systems*. PhD thesis, stanford university Stanford, CA, 2002.

[81] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.

[82] Dongbin Xiu and George Em Karniadakis. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, 24(2):619–644, 2002.

[83] Géraud Blatman and Bruno Sudret. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliability Engineering & System Safety*, 95(11):1216–1229, 2010.

[84] Johannes Maußner and Hannsjörg Freund. Optimization under uncertainty in chemical engineering: Comparative evaluation of unscented transformation methods and cubature rules. *Chemical Engineering Science*, 183:329–345, 2018.

[85] Simon J Julier and Jeffrey K Uhlmann. A general method for approximating nonlinear transformations of probability distributions. *Robotics Research Group, Department of Engineering Science, University of Oxford, Oxford, OC1 3PJ United Kingdom, Tech. Rep*, 1996.

[86] René Schenkendorf and Jörn C Groos. Global sensitivity analysis applied to model inversion problems: a contribution to rail condition monitoring. *International Journal of Prognostics and Health Management*, 6(019):1–14, 2015.

[87] René Schenkendorf and Michael Mangold. Qualitative and quantitative optimal experimental design for parameter identification of a map kinase model. *IFAC Proceedings Volumes*, 44(1):11666–11671, 2011.

[88] Dries Telen, Filip Logist, Eva Van Derlinden, and Jan F Van Impe. Robust optimal experiment design: A multi-objective approach. *IFAC Proceedings Volumes*, 45(2):689–694, 2012.

[89] Régis Lebrun and Anne Dutfoy. A generalization of the nataf transformation to distributions with elliptical copula. *Probabilistic Engineering Mechanics*, 24(2):172–178, 2009.

[90] Roger G Ghanem and Pol D Spanos. *Stochastic finite elements: a spectral approach*. Courier Corporation, 2003.

[91] Xiaoliang Wan and George Em Karniadakis. Beyond wiener–askey expansions: handling arbitrary pdfs. *Journal of Scientific Computing*, 27(1-3):455–464, 2006.

[92] Jeroen AS Witteveen, Sunetra Sarkar, and Hester Bijl. Modeling physical uncertainties in dynamic stall induced fluid–structure interaction of turbine blades using arbitrary polynomial chaos. *Computers & structures*, 85(11):866–878, 2007.

[93] Sergey Oladyshkin and Wolfgang Nowak. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliability Engineering & System Safety*, 106:179–190, 2012.

[94] Maria Navarro, Jeroen Witteveen, and Joke Blom. Polynomial chaos expansion for general multivariate distributions with correlated variables. *arXiv preprint arXiv:1406.5483*, 2014.

[95] Aleksandr F Timan. *Theory of approximation of functions of a real variable*, volume 34. Elsevier, 2014.

[96] Douglas C Montgomery. *Design and analysis of experiments*. John Wiley & Sons, 2008.

[97] Bradley Efron, Trevor Hastie, Iain Johnstone, Robert Tibshirani, et al. Least angle regression. *The Annals of statistics*, 32(2):407–499, 2004.

[98] Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, 22(10):1509–1518, 2007.

[99] Ilya M Sobol and Sergei Kucherenko. Derivative based global sensitivity measures and their link with global sensitivity indices. *Mathematics and Computers in Simulation*, 79(10):3009–3017, 2009.

[100] Jon C Helton and Freddie Joe Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1):23–69, 2003.

[101] Marcus Degerman, Karin Westerberg, and Bernt Nilsson. Determining critical process parameters and process robustness in preparative chromatography–a model-based approach. *Chemical engineering & technology*, 32(6):903–911, 2009.

[102] Andrea Saltelli and Ilya M Sobol. About the use of rank transformation in sensitivity analysis of model output. *Reliability Engineering & System Safety*, 50(3):225–239, 1995.

[103] Andrea Saltelli, Stefano Tarantola, and KP-S Chan. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1):39–56, 1999.

[104] Il'ya M Sobol'. On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe Modelirovanie*, 2(1):112–118, 1990.

[105] Genyuan Li, Herschel Rabitz, Paul E Yelvington, Oluwayemisi O Oluwole, Fred Bacon, Charles E Kolb, and Jacqueline Schoendorf. Global sensitivity analysis for systems with independent and/or correlated inputs. *The Journal of Physical Chemistry A*, 114(19):6022–6032, 2010.

[106] Chonggang Xu and George Zdzislaw Gertner. Uncertainty and sensitivity analysis for models with correlated parameters. *Reliability Engineering & System Safety*, 93(10):1563–1573, 2008.

[107] Thierry A Mara and Stefano Tarantola. Variance-based sensitivity indices for models with dependent inputs. *Reliability Engineering & System Safety*, 107:115–121, 2012.

[108] Thierry A Mara, Stefano Tarantola, and Paola Annoni. Non-parametric methods for global sensitivity analysis of model output with dependent inputs. *Environmental Modelling & Software*, 72:173–183, 2015.

[109] Éva Valkó, Tamás Varga, Alison Tomlin, and Tamas Turányi. Investigation of the effect of correlated uncertain rate parameters on a model of hydrogen combustion using a generalized hdmr method. *Proceedings of the Combustion Institute*, 36(1):681–689, 2017.

[110] Bruno Sudret and Y Caniou. Analysis of covariance (ancova) using polynomial chaos expansions. *11th International Conference on Structural Safety & Reliability*, 2013.

[111] Ilya M Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation*, 55(1):271–280, 2001.

[112] Roger B Nelsen. *An introduction to copulas*. Springer-Verlag, Berlin, Heidelberg, 2007.

[113] M Sklar. *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8, 1959.

[114] Jeremy E Oakley and Anthony O'Hagan. Probabilistic sensitivity analysis of complex models: a bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):751–769, 2004.

[115] Murray Rosenblatt. Remarks on a multivariate transformation. *The annals of mathematical statistics*, 23(3):470–472, 1952.

[116] Genyuan Li, Carey Rosenthal, and Herschel Rabitz. High dimensional model representations. *The Journal of Physical Chemistry A*, 105(33):7765–7777, 2001.

[117] Genyuan Li and Herschel Rabitz. General formulation of hdmr component functions with independent and correlated variables. *Journal of Mathematical Chemistry*, 50(1):99–130, 2012.

[118] Genyuan Li and Herschel Rabitz. Relationship between sensitivity indices defined by variance-and covariance-based methods. *Reliability Engineering & System Safety*, 167:136–157, 2017.

[119] Zdravko I Botev, Joseph F Grotowski, Dirk P Kroese, et al. Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5):2916–2957, 2010.

[120] Vassiliy A Epanechnikov. Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications*, 14(1):153–158, 1969.

[121] Pengfei Wei, Zhenzhou Lu, and Xiukai Yuan. Monte carlo simulation for moment-independent sensitivity analysis. *Reliability Engineering & System Safety*, 110:60–67, 2013.

[122] Stefano Marelli and Bruno Sudret. Uqlab: A framework for uncertainty quantification in matlab. In *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management*, pages 2554–2563. 2014.

[123] Ronald L Iman and WJ Conover. A measure of top–down correlation. *Technometrics*, 29(3):351–357, 1987.

[124] Filip Logist, IY Smets, and JF Van Impe. Derivation of generic optimal reference temperature profiles for steady-state exothermic jacketed tubular reactors. *Journal of Process Control*, 18(1):92–104, 2008.

[125] Yan-Gang Zhao and Tetsuro Ono. Moment methods for structural reliability. *Structural safety*, 23(1):47–75, 2001.

[126] Sirisha Rangavajhala, Anoop Mullur, and Achille Messac. The challenge of equality constraints in robust design optimization: examination and new approach. *Structural and Multidisciplinary Optimization*, 34(5):381–401, 2007.

[127] Gennady Ostrovsky, Nadir Zyatdinov, and Tatyana Lapteva. Optimal design of chemical processes with chance constraints. *Computers & Chemical Engineering*, 59:74–88, 2013.

[128] Andrei N Kolmogorov. *Foundations of the Theory of Probability: Second English Edition*. Courier Dover Publications, 2018.

[129] Alex J Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222, 2004.

[130] Lorenz T Biegler. An overview of simultaneous strategies for dynamic optimization. *Chemical Engineering and Processing: Process Intensification*, 46(11):1043–1053, 2007.

[131] Joel Andersson, Johan Åkesson, and Moritz Diehl. Casadi: A symbolic package for automatic differentiation and optimal control. In *Recent advances in algorithmic differentiation*, pages 297–307. Springer, 2012.

[132] Andreas Wächter and Lorenz T Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106(1):25–57, 2006.

[133] Iain S Duff. Ma57—a code for the solution of sparse symmetric definite and indefinite systems. *ACM Transactions on Mathematical Software (TOMS)*, 30(2):118–144, 2004.

[134] RK Bajpai and M Reuss. A mechanistic model for penicillin production. *Journal of Chemical Technology and Biotechnology*, 30(1):332–344, 1980.

[135] Ka-Yiu San and Gregory Stephanopoulos. Optimization of fed-batch penicillin fermentation: A case of singular optimal control with state constraints. *Biotechnology and bioengineering*, 34(1):72–78, 1989.

[136] Hans-Georg Beyer and Bernhard Sendhoff. Robust optimization–a comprehensive survey. *Computer methods in applied mechanics and engineering*, 196(33):3190–3218, 2007.

[137] Ali Mesbah, Stefan Streif, Rolf Findeisen, and Richard D Braatz. Stochastic nonlinear model predictive control with probabilistic constraints. *2014 American Control Conference (ACC)*, 1:2413–2419, 2014.

[138] Ralph C Smith. *Uncertainty quantification: theory, implementation, and applications*, volume 12. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2013.

[139] Philippe Nimmegeers, Dries Telen, Filip Logist, and Jan Van Impe. Dynamic optimization of biological networks under parametric uncertainty. *BMC systems biology*, 10(1):86, 2016.

[140] Xiaoqing Shi, Ming Ye, Gary P Curtis, Geoffery L Miller, Philip D Meyer, Matthias Kohler, Steve Yabusaki, and Jichun Wu. Assessment of parametric uncertainty for groundwater reactive transport modeling. *Water Resources Research*, 50(5):4416–4439, 2014.

[141] Paul Sheehan and Athanasios I Liapis. Modeling of the primary and secondary drying stages of the freeze drying of pharmaceutical products in vials: Numerical results obtained from the solution of a dynamic and spatially multi-dimensional lyophilization model for different operational policies. *Biotechnology and Bioengineering*, 60(6):712–728, 1998.

[142] Alessio Carullo and Alberto Vallan. Measurement uncertainty issues in freeze-drying processes. *Measurement*, 45(7):1706–1712, 2012.

[143] Jayashree Kalyanaraman, Yanfang Fan, Ying Labreche, Ryan P Lively, Yoshiaki Kawajiri, and Matthew J Realff. Bayesian estimation of parametric uncertainties, quantification and reduction using optimal design of experiments for co 2 adsorption on amine sorbents. *Computers & Chemical Engineering*, 81:376–388, 2015.

[144] Milind Joshi, Andreas Seidel-Morgenstern, and Andreas Kremling. Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. *Metabolic Engineering*, 8(5):447–455, 2006.

[145] Niko Rossner. *Robuste modellgestützte Prozessführung auf Basis von Gauß'schen Mischdichten am Beispiel der Bray-Liebhafsky-Reaktion und der autotrophen Kultivierung von Ralstonia eutropha H16*. PhD thesis, ,TU Berlin, Berlin, 2014.

[146] Teng Xu and Jaime J Gómez-Hernández. Characterization of non-gaussian conductivities and porosities with hydraulic heads, solute concentrations, and water temperatures. *Water Resources Research*, 52(8):6111–6136, 2016.

[147] Christopher Bishop. *Pattern recognition and machine learning*. Springer-Verlag, Berlin, Heidelberg, 2006.

[148] Kay N Shu, Thriyambakam Krishnan, and McLachlan J Geoffrey. *The EM algorithm*. Springer-Verlag, Berlin, Heidelberg, 2012.

[149] Jean P Amorij, J Meulenaar, Wouter Hinrichs, Toon Stegmann, Anke Huckriede, F Coenen, and HW Frijlink. Rational design of an influenza subunit vaccine powder with sugar glass technology: preventing conformational changes of haemagglutinin during freezing and freeze-drying. *Vaccine*, 25(35):6447–6457, 2007.

[150] Julia C Kasper, Gerhard Winter, and Wolfgang Friess. Recent advances and further challenges in lyophilization. *European Journal of Pharmaceutics and Biopharmaceutics*, 85(2):162–169, 2013.

[151] Davide Fissore, Roberto Pisano, and Antonello A Barresi. On the methods based on the pressure rise test for monitoring a freeze-drying process. *Drying Technology*, 29(1):73–90, 2010.

[152] Davide Fissore, Roberto Pisano, and Antonello A Barresi. Advanced approach to build the design space for the primary drying of a pharmaceutical freeze-drying process. *Journal of Pharmaceutical Sciences*, 100(11):4922–4933, 2011.

[153] Xiaolin Charlie Tang and Michael J Pikal. Design of freeze-drying processes for pharmaceuticals: practical advice. *Pharmaceutical Research*, 21(2):191–200, 2004.

[154] Bala Srinivasan, Dominique Bonvin, Erik Visser, and Srinivas Palanki. Dynamic optimization of batch processes: Ii. role of measurements in handling uncertainty. *Computers & chemical engineering*, 27(1):27–44, 2003.

[155] Nicolas M Kaiser, Robert J Flassig, and Kai Sundmacher. Probabilistic reactor design in the framework of elementary process functions. *Computers & Chemical Engineering*, 94:45–59, 2016.

[156] David G Kleinbaum and Mitche Klein. *Maximum likelihood techniques: An overview*. Springer-Verlag, Berlin, Heidelberg, 2010.

[157] Todd K Moon. The expectation-maximization algorithm. *IEEE Signal Processing Magazine*, 13(6):47–60, 1996.

[158] JM Bernardo, MJ Bayarri, JO Berger, AP Dawid, D Heckerman, AFM Smith, M West, et al. The variational bayesian em algorithm for incomplete data: with application to scoring graphical model structures. *Bayesian Statistics*, 7:453–464, 2003.

[159] Ujjwal D Gupta, Vinay Menon, and Uday Babbar. Parameter selection for em clustering using information criterion and pddp. *International Journal of Engineering and Technology*, 2(4):340, 2010.

[160] Daniel M Murphy and Thomas Koop. Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Quarterly Journal of the Royal Meteorological Society*, 131(608):1539–1565, 2005.

[161] Roberto Pisano, Davide Fissore, Antonello A Barresi, Philippe Brayard, Pierre Chouvenc, and Bertrand Woinet. Quality by design: optimization of a freeze-drying cycle via design space in case of heterogeneous drying behavior and influence of the freezing protocol. *Pharmaceutical Development and Technology*, 18(1):280–295, 2013.

[162] Jim Searles. Observation and implications of sonic water vapor flow during freeze-drying. *American Pharmaceutical Review*, 7:58–69, 2004.

[163] K Krishna and M Narasimha Murty. Genetic k-means algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 29(3):433–439, 1999.

[164] Ronald A Fisher. *Statistical methods and scientific inference*. Oliver & Boyd, Edinburgh, 2nd ed. edition, 1959.

[165] Dongying E Shen and Richard D Braatz. Polynomial chaos-based robust design of systems with probabilistic uncertainties. *AIChE Journal*, 62(9):3310–3318, 2016.

[166] Mina Rafiei-Shishavan, Siddharth Mehta, and Luis A Ricardez-Sandoval. Simultaneous design and control under uncertainty: A back-off approach using power series expansions. *Computers & Chemical Engineering*, 99:66–81, 2017.

[167] Mina Rafiei and Luis A Ricardez-Sandoval. Stochastic back-off approach for integration of design and control under uncertainty. *Industrial & Engineering Chemistry Research*, 57(12):4351–4365, 2018.

[168] Erdal Aydin, Dominique Bonvin, and Kai Sundmacher. Nmpc using pontryagin's minimum principle-application to a two-phase semi-batch hydroformylation reactor under uncertainty. *Computers & Chemical Engineering*, 108:47–56, 2018.

[169] Federico Galvanin, Massimiliano Barolo, Fabrizio Bezzo, and Sandro Macchietto. A backoff strategy for model-based experiment design under parametric uncertainty. *AIChE journal*, 56(8):2088–2102, 2010.

[170] Johannes Maußner and Hannsjörg Freund. Efficient calculation of constraint backoffs for optimization under uncertainty: A case study on maleic anhydride synthesis. *Chemical Engineering Science*, 192:306–317, 2018.

[171] Stefano Marelli and Bruno Sudret. Uqlab user manual polynomial chaos expansions. *Chair of Risk, Safety & Uncertainty Quantification*, 1:9–104, 2015.

[172] Md A Rashid. *Crystallization engineering of ibuprofen for pharmaceutical formulation.* PhD dissertation, Queensland University of Technology, 2011.

[173] JD Perkins, C Gannavarapu, and GW Barton. Choosing control structures based on economics. *IEE Colloquium on Successful Industrial Applications of Multivariable Analysis*, 1:1–3, 1990.

[174] Lawrence T Narraway and John D Perkins. Selection of process control structure based on linear dynamic economics. *Industrial & Engineering Chemistry Research*, 32(11):2681–2692, 1993.

[175] Ioannis K Kookos and John D Perkins. Control structure selection based on economics: Generalization of the back-off methodology. *AIChE Journal*, 62(9):3056–3064, 2016.

[176] Parisa A Bahri, Jose A Bandoni, and Jose A Romagnoli. Back-off calculations in optimising control: a dynamic approach. *Computers & chemical engineering*, 19:699–708, 1995.

[177] Parisa A Bahri, Jose A Bandoni, and Jose A Romagnoli. Effect of disturbances in optimizing control: Steady-state open-loop backoff problem. *AIChE Journal*, 42(4):983–994, 1996.

[178] Jose L Figueroa, Parisa A Bahri, Jose A Bandoni, and Jose A Romagnoli. Economic impact of disturbances and uncertain parameters in chemical processes—a

dynamic back-off analysis. *Computers & chemical engineering*, 20(4):453–461, 1996.

[179] Erik Visser, Bala Srinivasan, Srinivas Palanki, and Dominique Bonvin. A feedback-based implementation scheme for batch process optimization. *Journal of Process Control*, 10(5):399–410, 2000.

[180] Moritz Diehl, Hans G Bock, and Ekaterina Kostina. An approximation technique for robust nonlinear optimization. *Mathematical Programming*, 107(1-2):213–230, 2006.

[181] Robert W Koller, Luis A Ricardez-Sandoval, and Lorenz T Biegler. Stochastic back-off algorithm for simultaneous design, control, and scheduling of multiproduct systems under uncertainty. *AIChE Journal*, 2018.

[182] Niko Rossner, Th Heine, and Rudibert King. Quality-by-design using a gaussian mixture density approximation of biological uncertainties. *IFAC Proceedings Volumes*, 43(6):7–12, 2010.

[183] Géraud Blatman and Bruno Sudret. Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6):2345–2367, 2011.

[184] Walter Gautschi. *Orthogonal polynomials: computation and approximation*. Oxford University Press on Demand, 2004.

[185] Rafael JP Eder, Stefan Radl, Elisabeth Schmitt, Sabine Innerhofer, Markus Maier, Heidrun Gruber-Woelfler, and Johannes G Khinast. Continuously seeded, continuously operated tubular crystallizer for the production of active pharmaceutical ingredients. *Crystal growth & design*, 10(5):2247–2257, 2010.

[186] Qinglin Su, Brahim Benyahia, Zoltan K Nagy, and Chris D Rielly. Mathematical modeling, design, and optimization of a multisegment multiaddition plug-flow crystallizer for antisolvent crystallizations. *Organic Process Research & Development*, 19(12):1859–1870, 2015.

[187] Shamsul Qamar, Martin P Elsner, Ivan A Angelov, Gerald Warnecke, and Andreas Seidel-Morgenstern. A comparative study of high resolution schemes for solving population balances in crystallization. *Computers & Chemical Engineering*, 30(6-7):1119–1131, 2006.

[188] Roland Schöbi and Bruno Sudret. Structural reliability analysis for p-boxes using multi-level meta-models. *Probabilistic Engineering Mechanics*, 48:27–38, 2017.