Thesis for the degree of doctor of philosophy

Advances in optimal design and retrofit of chemical processes with uncertain parameters

- Applications in design of heat exchanger networks

CHRISTIAN LANGNER

Department of Space, Earth and Environment
Chalmers University of Technology
Gothenburg, Sweden 2023

Advances in optimal design and retrofit of chemical processes with uncertain parameters
- Applications in design of heat exchanger networks

CHRISTIAN LANGNER
ISBN 978-91-7905-864-7

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5330
ISSN 0346-718X

Department of Space, Earth and Environment
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone +46 31 772 1000

Cover: Illustration of the uncertainty that the industrial sector is facing now and in the (near) future. As one of the main contributors to climate change, the industrial sector needs to take immediate action to mitigate greenhouse gas emissions. This transition to a more sustainable industrial sector is challenging and comes with many uncertainties.

Printed by Chalmers Digitaltryck
Gothenburg, Sweden 2023

Advances in optimal design and retrofit of chemical processes with uncertain parameters - Applications in design of heat exchanger networks

CHRISTIAN LANGNER
Department of Space, Earth and Environment
Chalmers University of Technology

Abstract

There is widespread consensus that the omnipresent climate crisis demands humanity to rapidly reduce global greenhouse gas (GHG) emissions. To allow for such a rapid reduction, the industrial sector as a main contributor to GHG emissions needs to take immediate actions. To mitigate GHG emissions from the industrial sector, increasing energy efficiency as well as fuel and feedstock switching, such as increased use of biomass and (green) electricity, are the options which can have most impact in the short- and medium-term. Such mitigation options usually create a need for design of new or redesign of existing processes such as the plant energy systems. The design and operation of industrial plants and processes are usually subject to uncertainty, especially in the process industry. This uncertainty can have different origins, e.g., process parameters such as flow rates or transfer coefficients may vary (uncontrolled) or may not be known exactly.

This thesis proposes theoretical and methodological developments for designing and/or redesigning chemical processes which are subject to uncertain operating conditions, with a special focus on heat recovery systems such as heat exchanger networks. In this context, this thesis contributes with theoretical development in the field of deterministic flexibility analysis. More specifically, new approaches are presented to enhance the modelling of the expected uncertainty space, i.e., the space in which the uncertain parameters are expected to vary. Additionally, an approach is presented to perform (deterministic) flexibility analysis in situations when uncertain long-term development such as a switch in feedstocks interferes with operational short-term disturbances. In this context, the thesis presents an industrial case study to *i)* show the need for such a theoretical development, and *ii)* illustrate the applicability.

Aside of advances in deterministic flexibility analysis, this thesis also explores the possibility to combine valuable designer input (e.g., non-quantifiable knowledge) with the efficiency of mathematical programming when addressing a design under uncertainty problem. More specifically, this thesis proposes to divide the design under uncertainty problem into a design synthesis step which allows direct input from the designer, and several subsequent steps which are summarized in a framework presented in this thesis. The proposed framework combines different approaches from the literature with the theoretical development presented in this thesis, and aims to identify the optimal design specifications which also guarantee that the final design can operate at all expected operating conditions. The design synthesis step and the framework are decoupled from each other which allows the approach to be applied to large and complex industrial case studies with acceptable computational effort. Usage of the proposed framework is illustrated by means of an industrial case study which presents a design under uncertainty problem.

Keywords: Design under Uncertainty, Flexibility Analysis, Uncertainty Space, Process Design, Managing Variations, Process Integration.

List of publications

The following papers are included in this thesis:

Paper I: Langner, C., Svensson, E., & Harvey, S. (2021). A computational tool for guiding retrofit projects of industrial heat recovery systems subject to variation in operating conditions. Applied Thermal Engineering, 182. https://doi.org/10.1016/j.applthermaleng.2020.115648

Paper II: Langner, C., Svensson, E., & Harvey, S. (2020). A framework for flexible and cost-efficient retrofit measures of heat exchanger networks. Energies, 13(6). https://doi.org/10.3390/en13061472

Paper III: Marton, S., Langner, C., Svensson, E., & Harvey, S. (2021). Costs vs. flexibility of process heat recovery solutions considering short-term process variability and uncertain long-term development. Frontiers in Chemical Engineering. https://doi.org/10.3389/fceng.2021.679454

Paper IV: Langner, C., Svensson, E., Papadokonstantakis, S., & Harvey, S. (2022). Flexibility analysis of chemical processes considering overlaying uncertainty sources. Computer Aided Chemical Engineering, 49, (Proceedings of 14th International Symposium on Process Systems Engineering held in Kyoto, 2022). https://doi.org/10.1016/B978-0-323-85159-6.50128-7

Paper V: Langner, C., Svensson, E., Papadokonstantakis, S., & Harvey, S. (2023). Flexibility analysis using boundary functions for considering dependencies in uncertain parameters. Computers & Chemical Engineering, 108231. https://doi.org/10.1016/J.COMPCHEMENG.2023.108231

Paper VI: Langner, C., Svensson, E., Papadokonstantakis, S., & Harvey, S. Novel reformulations for modelling uncertainty and variations in a framework for chemical process design. submitted to Industrial & Engineering Chemistry Research

Co-authorship statement

Christian Langner (CL) is the principal author of **Papers I and II and IV - VI**. As second author, he contributed to **Paper III** with planning, developing and discussing the studied approach together with the principal author Sofie Marton (SM). CL also conducted the optimization and calculations, and wrote the methodology section of **Paper III**. Elin Svensson (ES) contributed with the conceptualization, discussions and editing to **Papers I - VI**. Stavros Papadokonstantakis (SP) contributed to **Papers IV - VI** with the conceptualization and discussions. Simon Harvey (SH) contributed with discussions and editing to **Papers I - VI**. All manuscripts were reviewed and discussed by all authors. The tool appended to **Paper I** as supporting material was developed by Pontus Bokinge (PB) and David Erlandsson (DE) who also collected process data used for the case studies presented in **Papers I & VI**.

Other publications

Other publications by the author, not included in the thesis because they are either outside the scope of the thesis or overlap with the appended papers.

- Langner, C., Svensson, E., Bokinge, P., Erlandsson, D., & Harvey, S. (2019). A computational tool for analysing the response of complex heat exchanger networks to disturbances. Proceedings of the 29th International Conference on Efficiency, Cost, Optimisation, Simulation and Environmental Impact of Energy Systems (pp. 621–633). http://www.s-conferences.eu/ecos2019
- Langner, C., Svensson, E., & Harvey, S. (2019). Combined Flexibility and Energy Analysis of Retrofit Actions for Heat Exchanger Networks. Chemical Engineering Transactions, 76(2014), pp. 307–312. https://doi.org/10.3303/CET1976052
- Svensson, E., Edland, R., Langner, C. & Harvey, S. (2020). Assessing the value of a diversified by-product portfolio to allow for increased production flexibility in pulp mills. Nordic Pulp & Paper Research Journal, 35(4), pp. 533-558. https://doi.org/10.1515/npprj-2020-0034
- Langner, C., Svensson, E., & Harvey, S. (2021). Flexibility analysis of chemical processes considering dependencies between uncertain parameters. Computer Aided Chemical Engineering, 50, pp. 1105–1110. https://doi.org/10.1016/B978-0-323-88506-5.50170-4
- Biermann, M., Langner, C., Eliasson, Å., Normann, F., Harvey, S., & Johnsson, F. (2021). Partial capture from refineries through utilization of existing site energy systems. Proceedings of the 15th Greenhouse Gas Control Technologies Conference 15-18 March 2021. https://doi.org/10.2139/ssrn.3820101
- Kruse, J., & Langner, C. (2021). Covid-19 in court: A quantitative evaluation of administrative law cases (published in German under the original title: Covid-19 vor Gericht: Eine quantitative Auswertung der verwaltungsgerichtlichen Judikatur). Neue Juristische Wochenschrift, 74, pp. 3707-3712.
- Biermann, M., Langner, C., Roussanaly, S., Normann, F., & Harvey, S. (2022). The role of energy supply in abatement cost curves for CO2 capture from process industry – A case study of a Swedish refinery. Applied Energy, 319, 119273. https://doi.org/10.1016/J.APENERGY.2022.119273

Acknowledgements

I am immensely thankful for for all that I have experienced on my PhD journey. The entire adventure would have been less-fruitful, or perhaps even impossible, without the unwavering support that I received from so many people, all of whom I wish to thank.

First and foremost, I thank my supervisors Dr. Elin Svensson and Prof. Simon Harvey for their support and for giving me the opportunity to pursue my PhD education. In particular, I thank Elin for her continuous support of my PhD studies as well as related research, and for her patience, motivation and the ability to always ask the right questions. I am very grateful for Simon's guidance and critical feedback which helped me in all the time of research and writing of this thesis. I also want to thank Prof. Stavros Papadokonstantakis for joining my supervision team after the licentiate, and the interesting discussions and new perspectives on flexibility analysis. I could not have imagined having better supervisors and mentors for my PhD studies.

I am indebted to my co-authors and the partners in the research projects in which I have been involved. To Sofie - thank you for your commitment and the great time during our collaboration. Thank you, Max, for the new perspectives on CCS also resulting in our popular science podcast. My sincere thanks to the collaborating project partners working at Södra. Special thanks to Örjan Johansson, Magnus Tyrberg and Karin Dernegård who hosted me and my Master thesis students during our study trips to Södra Cell Mönsterås. Many thanks also to Dr. Johan Isaksson for the interesting discussions and for being accessible for further advice. I also want to thank Prof. Fredrik Normann for supporting my studies after the licentiate and the great discussions we shared on applications of AI. I am thankful for the financial support received from the Swedish Energy Agency (grant nr P42326.1), the Södra Foundation for Research, Development and Education as well as the ÅForsk foundation (project nr 20-304).

I would like to thank all my current and former colleges at Energy Technology, especially my roommate Johanna for being the best companion one could wish for during this journey through academia. Special thanks to my (former) colleagues in the Industrial Energy Systems Analysis research group as well as the Industrial Process and Systems research group. A very special thank you also to my Master thesis students Pontus Bokinge, David Erlandsson and Arslan Farooqi for their excellent work.

Thank you to all my friends inside and outside of work for giving me different perspectives on life. Unfortunately, I cannot list you all here due to space limitation. I am especially thankful to my best friend Johannes who is always available to me with advice despite the kilometers separating us.

I would like to thank my parents and brothers for encouraging and supporting me. Thank you for being there for me, believing in me and supporting me; without your support I would not have been able to reach this point.
Finally, I want to thank my wife Lisa for being with me through all these years, supporting me but also keeping me grounded. You and our two wonderful children Ella and Anton remind me more than anything else in the world that there is a life outside of studies and work. I could not have wished for a greater partner and family accompanying me through the game of life.

Christian Langner, Gothenburg, May 2023

Contents

List of Figures

List of Tables

1

Introduction

Global emissions of greenhouse gases (GHG) have increased steadily since pre-industrial times, reaching the highest levels in human history over the last decade [1]. The main effect of the increasing GHG emissions has been the rise of the global average temperature causing severe danger to the earth's ecosystems. By signing the Paris Agreement in 2015, countries acknowledged the necessity of keeping the most severe climate change risks in check by limiting the raise of the global average temperature to well below 2 °C and pursuing efforts to limit warming to 1.5 °C [3]. Both the sixth assessment report (working group III) of the Intergovernmental Panel on Climate Change (IPCC) [1] as well as the most recent World Energy Outlook of the International Energy Agency (IEA) [4] point out that this requires rapid and sustained reduction of GHG emissions towards net zero carbon dioxide (CO_2) emissions well within the 21st century along with deep reductions of non-CO_2 (GHG) emissions. More specifically, based on the dataset provided by Friedlingstein et al. [5], the IPCC report [1] concluded that the remaining CO_2 budget to limit warming to 1.5 °C (with a chance of 67%) is of similar magnitude as the total amount of CO_2 emitted during the last decade (see Figure 1.1). The challenges related to achieving the necessary reduction of GHG emissions become even more obvious when considering that 50% of the total GHG emissions shown in Figure 1.1 were emitted after 1970 which highlights the current high levels of GHG emissions [1].

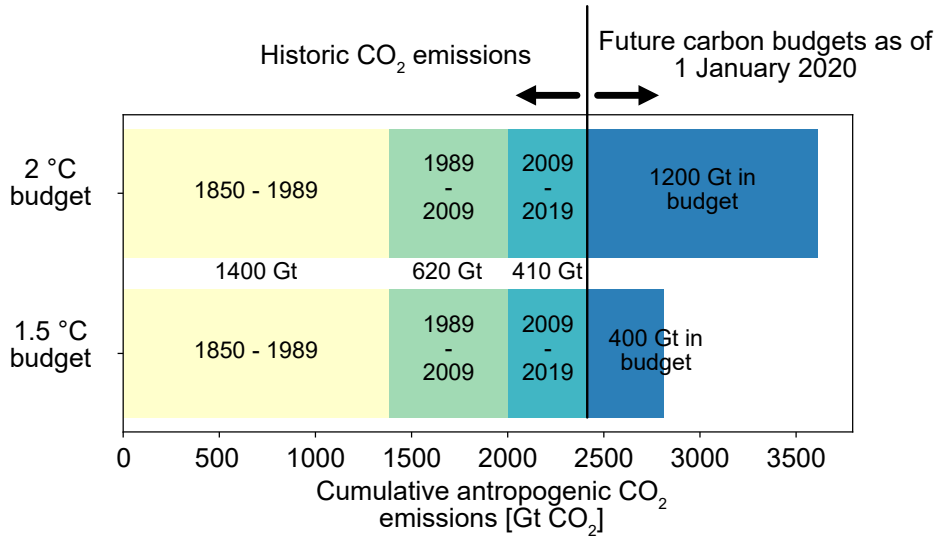


Figure 1.1: Historic cumulative anthropogenic CO_2 emissions for the periods 1850–1989, 1990–2009, and 2010–2019 as well as remaining future carbon budgets as of 1 January 2020 to limit warming to 1.5 °C and to 2 °C at the 67th percentile of the transient climate response to cumulative CO_2 emissions. Data was taken from the sixth assessment report (working group III) of the Intergovernmental Panel on Climate Change (IPCC) [1].

To gain a better understanding of the sources of current GHG emissions, Minx et al. [6] allocated emissions to different sectors. The authors reported that in 2019 the industrial sector was responsible for the second biggest share of all direct anthropogenic GHG emissions, namely 24%, which originated from fuel combustion, process emissions, product use and waste. Note that when including indirect emissions from power and heat generation, the share of emissions from the industrial sector increased to 34%, and was thereby the biggest emitter of GHG in 2019 [1]. To mitigate GHG emissions from the industrial sector, the IPCC report [1] points out several options. Additionally, the report presents estimations regarding the mitigation potential as well as the approximate costs connected with implementation of these options for the year 2030. An overview of these options in shown in Figure 1.2.

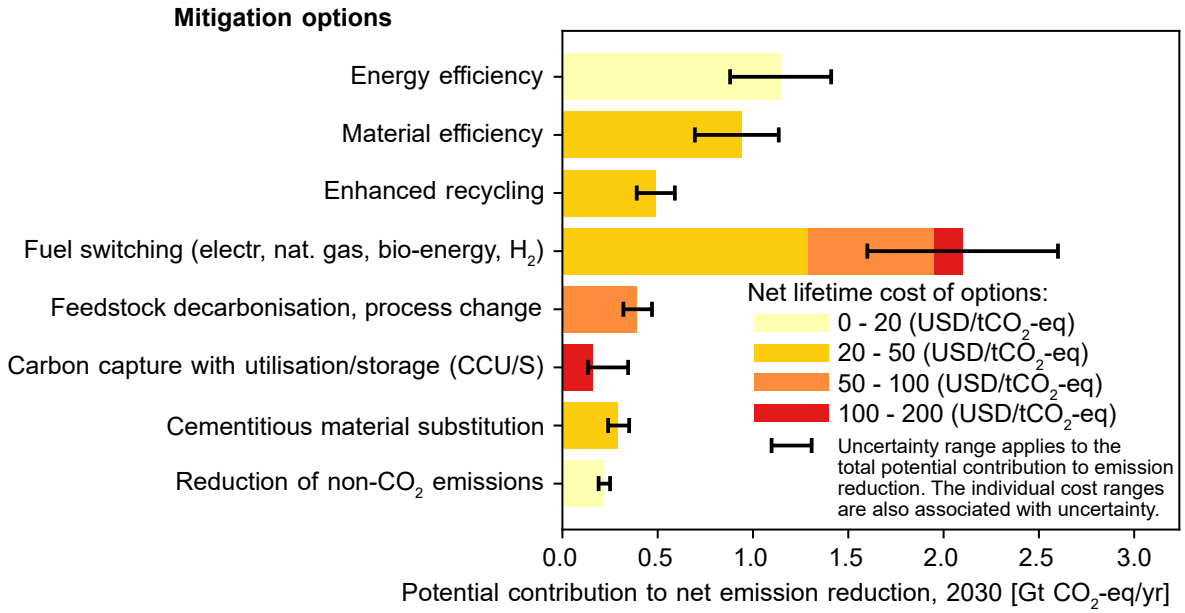


Figure 1.2: Overview of GHG mitigation options and their estimated ranges of costs and potentials in 2030 for the industrial sector. Data was taken from the sixth assessment report (working group III) of the Intergovernmental Panel on Climate Change (IPCC) [1].

Figure 1.2 shows that increasing energy efficiency by appropriate measures is an important mitigation option as it offers the second largest mitigation potential with lowest expected costs. On the other hand, there is indication that energy efficiency improvements have been slowing down over the recent decades since the technological improvement has come closer to the thermodynamic optimum [1, 7]. Nevertheless, studies such as Papapetrou et al. [8] identified a waste heat potential of about 300 TWh/yr within the EU which corresponds to about 10% of total energy use in the industrial sector. Additional options shown in Figure 1.2 correspond to reducing primary production by increased material efficiency or circular economy solutions boosted by enhanced recycling. Since primary production will remain necessary, especially in the short-and medium-term, further GHG mitigation options include switching to new processes and feedstocks such as biomass, increased use of (green) electricity and hydrogen (highest mitigation potential until 2030 according to the IPCC report [1]) as well as carbon capture and utilization/storage (CCU/S).

The different options for the mitigation of GHG emissions from industrial plants and value chains typically create a need for design of new processes as well as redesign of existing

processes. For example, the energy systems of industrial plants, such as heat recovery systems, can be affected as a result of measures to ensure the energy efficient integration of new technologies and systems. Additionally, the integration of new technologies and systems as well as the adaption or redesign of the existing system are likely to influence the operating conditions of the industrial plants. For example, a switch of feedstocks may change process flow characteristics, such as flow rates and compositions, in unit processes that are otherwise unchanged. Another example is the integration of an absorption-based carbon capture unit which would introduce a significant new energy demand, e.g., in the form of heat for absorbent regeneration. On the other hand, the quality and in the best case also the quantity of the core product(s) of the production process should not be affected by such measures, which sets restrictions and constraints when pursuing GHG mitigation potentials in the industrial sector. For example, a carbon capture absorption process could be driven by excess process heat which could also be delivered to a district heating system.

Besides the restrictions resulting from product quality and quantity constraints, an additional barrier for the implementation of new processes as well as the enhancement of energy efficiency measures can be that industrial plants and processes are commonly subject to uncertain operating conditions. Such uncertain operating conditions usually result from uncertainty related to external process parameters such as feed stream characteristics, utility streams, ambient conditions, or economic cost data. In addition, the uncertain parameters can also be internal process parameters such as heat transfer coefficients, reaction constants, equipment efficiencies, or physical properties. Consequently, when designing new processes or redesigning existing plants to enhance energy efficiency, such uncertainties should be accounted for. Hereafter, two examples are presented to illustrate the challenges connected to enhancing energy efficiency as well as switching feedstocks in industrial processes.

1.1 Motivational examples

1.1.1 Example 1 - Energy efficiency measure in a Swedish pulp mill

Figure 1.3 shows the flowsheet of a subsystem of the secondary heating system of a Swedish pulp mill. In this system, three cold process streams (combustion air, feed water and district heating water) are heated to specified target temperatures by heat exchanging with different heat sources, namely hot secondary heating water, excess hot water and hot diluted process water. Although a substantial amount of heat can be recovered by heat exchangers (HEX), steam is needed to ensure that the target temperatures of the process streams are met. Figure 1.3 shows a possible retrofit of the subsystem that reduces the demand for low and medium pressure steam by increased heat recovery. More precisely, it is suggested to recover heat from a diluted black liquor stream. This case study is presented in more detail in **Paper VI**. Note that in the presented case study, it was assumed that the existing steam heaters must be replaced.

Figure 1.3 shows nine locations (marked in yellow) within the investigated system where process conditions, more specifically heat capacity flow rates and/or (supply) temper-

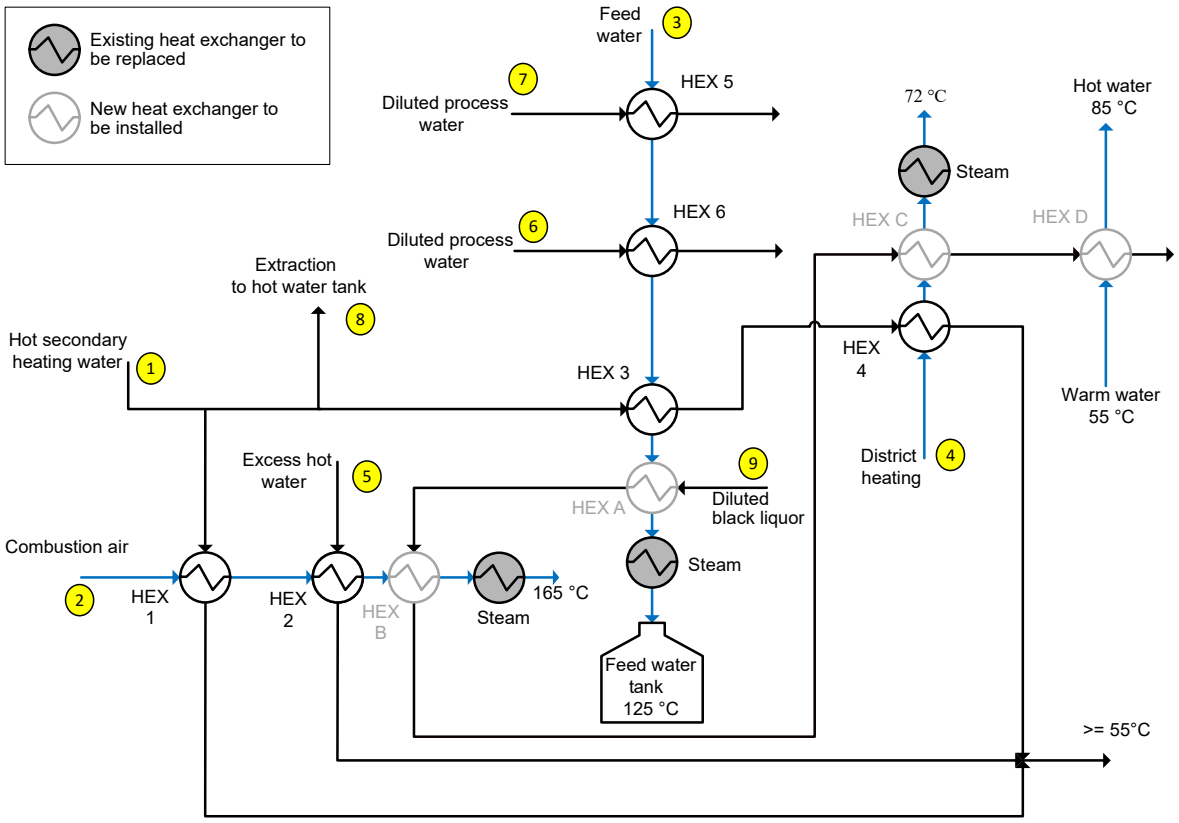


Figure 1.3: Process flow diagram of a subsystem of the secondary heating system of a Swedish pulp mill including a suggested retrofit for increased heat recovery. Locations where process conditions (heat capacity flow rates and/or supply temperatures) are determined outside the system's boundary are highlighted in yellow.

atures, are determined outside of the system's boundary. Available measurement data indicates that these process conditions do not remain constant during regular operation. However, the operational target of the given subsystem which is to heat the three process streams to their respective target temperatures needs to be met at all times during regular operation, regardless of the variation in the uncertain parameters. Consequently, the designer needs to identify the optimal heat transfer areas of the different new HEX units in order to ensure that the operational targets are always met while also considering both investment cost as well as operational cost/revenue, i.e., the designer needs to identify the optimal solution for a design under uncertainty problem.

1.1.2 Example 2 - Implications of switching to biomass feedstock in a Swedish oil refinery

The second example explores possible consequences of the implementation of deep decarbonization strategies such as switching feedstocks in an oil refinery. The case study involves three process streams in two different process units of a large Swedish oil refinery (crude oil capacity of 11.5 million tons/year). Figure 1.4 presents the corresponding flowsheet. Streams 1 and 2 are located in the Naphtha Hydro Treatment Unit (NHTU) and stream 3 is located in the Catalytic Reforming Unit (CRU). As in example 1 presented in Section 1.1.1, certain process conditions (locations highlighted in yellow) are uncertain since these conditions are determined outside of the system's boundary. In addition, it is

expected that the temperature at location 3 (highlighted in orange) needs to be adjusted during operation to handle variations in the operating conditions of the subsequent reactors. Consequently, the temperature at location 3 is also uncertain since it is determined by operating conditions which cannot be influenced by the system. In **Papers III & IV**, a scenario is discussed that assumes a future implementation of biomass co-processing that affects the operation of the NHTU. The scenario is based on the possibility of the NHTU being one potential feed-in point for pre-processed biomass feed in existing refineries (see e.g., van Dyk et al. [9]). It is reasonable to assume that switching feedstocks would affect the process conditions, and in the appended papers it was assumed that the nominal flow rates of the streams in the NHTU increase by $50 - 100\%$. It can therefore be important to assess whether currently installed process equipment is able to operate at these expected conditions taking into account that the operation requires handling of short-term uncertainty such as variations in process conditions. Additionally, if retrofitting of process equipment is considered before switching feedstocks, a designer should be able to consider such planned modifications to avoid the additional need for retrofitting at a later point in time.

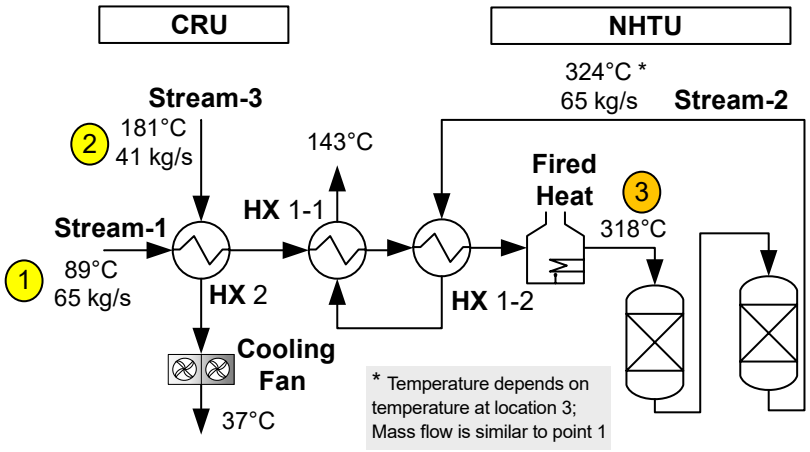


Figure 1.4: Process flow diagram showing the heat integration between the Naphtha Hydro Treatment Unit (NHTU) and the Catalytic Reforming Unit (CRU) in a Swedish oil refinery. The process conditions highlighted in yellow and orange are expected to vary during (normal) operation, and the shown values represent nominal operating conditions.

Both examples presented in Section 1.1 illustrate that implementing GHG mitigation options in existing industrial plants is complex due to the presence of uncertainties. The following section presents a comprehensive literature review about design under uncertainty for chemical processes/plants with a focus on heat recovery solutions.

1.2 Literature review

Traditionally, as reported in the literature [10–12], the approach to handle uncertainty in process parameters when designing or redesigning chemical processes/plants is to consider (only) nominal conditions, and use overdesign to compensate for the potential impact of the uncertainty. Some of this overdesign may be unnecessary, resulting in design solutions that are more expensive than necessary. To identify the optimal solution to such a problem, Grossmann and Sargent [13] defined the objective of the optimal design problem with uncertain parameters to design processes "that are always able to meet the specifications

for any feasible [or expected] values of the [uncertain] parameters and that, at the same time, are optimum with respect to a [...] cost function". Note that the authors explicitly considered the uncertain parameters to be continuous (and bounded), i.e., the parameters can take any value between a lower and an upper bound, and that the optimal process design allows for feasible operation at any of these values. The authors concluded that design under uncertainty is an infinite problem.

There exist two categories of approaches to solve such a problem, as pointed out by Steimel and Engell [14], namely sampling-based approaches and parametric approaches. In parametric approaches, the sampling of the uncertainty space is not required, meaning that the expected value of the objective function is obtained without approximation (see e.g., [15–17]). Difficulties may arise when the number of uncertain parameters increases due to increased computational burden. Sampling-based approaches are commonly based on the two-stage stochastic programming with recourse formulation and its transformation into a discretized deterministic equivalent problem. Compared to parametric approaches, in sampling-based approaches the expected value of the objective function is approximated for a discrete set of realizations of the uncertain parameters. The interested reader is referred to the work of Sahinidis [18] and Birge and Louveaux [19] for further information.

A major challenge with sampling-based approaches is that the number of samples necessary for accurate estimation of the expected objective function value as well as guaranteed feasibility strongly increases with the number of uncertain parameters. Several works in the literature suggest strategies to reduce the number of samples while maintaining a fair approximation of the expected objective value (see, e.g., [20–22]). In this context, several recently published works present approaches based on machine learning and data clustering to identify representative operating points/periods from a given distribution of operating points (see e.g., [23–28]). In addition to the expected value of the objective function, the discretization needs to detect the minimum set of samples required to guarantee steady-state flexible operation, i.e., feasible operation at any combination of uncertain parameter values within the expected uncertainty space[1]. In this context, flexibility analysis has evolved as a useful tool.

Flexibility analysis was introduced in the 1980s and aims to quantify a system's ability to cope with uncertainties. In this context, Halemane and Grossmann [29] formulated a flexibility test problem for fixed design specifications which evaluates whether steady-state flexible operation of a process is possible given an expected uncertainty space. The authors further defined critical parameter values as those combinations of the uncertain parameter values for which the feasible region of operation is the smallest. In this context, they postulated that a design which meets its operational targets at these critical operating points will also allow for steady-state flexible operation. To identify these critical operating points for processes which can be described by convex constraints, the authors adopted a procedure suggested by Grossmann and Sargent [13] which builds upon individual maximization of each constraint assuming monotonicity. Based on the work by Halemane and Grossmann, Pintarič and Kravanja [30] defined critical operating points as those combinations of the uncertain parameter values that require the largest overdesign

[1]The term expected uncertainty space refers to the space in which the uncertain parameters are expected to vary, i.e., the expected uncertainty space is defined by the expected disturbance range of each uncertain parameter.

of process units for the expected uncertainty space. Later, the same authors suggested in two works [31, 32] several algorithms to identify these combinations of the uncertain parameter values. Based on the suggested algorithms to identify critical operating points, the authors developed frameworks for synthesis of chemical process design [33] and for designing heat exchanger networks [34, 35] which are subject to uncertain operating data.

In order to quantify a system's ability to cope with uncertainties, several quantitative measures were suggested in the 1980s, including the resilience index suggested by Saboo et al. [36] and the (deterministic) flexibility index developed by Swaney and Grossmann [2]. The latter index has gained much attention in the literature, and is still subject to research (see e.g., [37–39]). The flexibility index is a scalar value which expresses the share of the expected uncertainty space which is feasible. More specifically, for each uncertain parameter the flexibility index indicates the share of the expected disturbance range in which the respective parameter may vary while still achieving feasible operation. For an extensive review on flexibility analysis, the interested reader is referred to the reviews by Grossmann et al. [11] and Zhang et al. [40]. Hereafter, an overview of some selected literature is provided.

To determine the flexibility index, Swaney and Grossmann [2] proposed search procedures for the special case of exclusively convex constraint functions. To overcome this limitation, Grossmann and Floudas [41] reformulated the flexibility index problem following the idea that the solution must be on the boundary of the feasible region. Detailed information on the study by Grossmann and Floudas [41] and their Mixed-Integer (Non-)Linear Program (MI(N)LP) formulation for the flexibility index is provided in Appendix A. Raspanti et al. [42] used constraint aggregation functions (Kreisselmeier–Steinhaus a.k.a. KS functions) and smoothing functions to reformulate the MI(N)LP formulation for the flexibility index by Grossmann and Floudas [41] to a single non-linear programming problem (NLP). Li et al. [43] suggested a framework to calculate an upper bound of the flexibility index by means of an alternating direction matrix embedded in a Simulated Annealing algorithm. Furthermore, the flexibility index was utilized in several (step-wise) frameworks to design or redesign chemical processes which are subject to uncertain parameters (see e.g., [34, 35, 44, 45]).

In the aforementioned frameworks on flexibility analysis, it is assumed that control variables can be manipulated to counteract variation in uncertain parameters which implies that the uncertain parameters are measurable. In addition to these works, Rooney and Biegler [17] as well as Ostrovsky et al. [46] also accounted for variation of unmeasured uncertain parameters which cannot be counteracted by means of recourse actions (i.e., manipulation of control variables). Recently, Ochoa and Grossmann [12] reformulated the MI(N)LP formulation by Grossmann and Floudas [41] to analyse the flexibility also when unmeasured uncertain parameters are present. Based on their work on flexibility analysis considering unmeasured uncertain parameters, Ochoa et al. [47] suggested novel formulations to identify the operating range of controllable process parameters within a design space.

An alternative approach for flexibility analysis was suggested by Pistikopoulos and Mazzuchi [48] as well as Straub and Grossmann [49]. Both works suggest determining a stochastic flexibility index which measures the probability that a given design (defined

by linear constraint functions) remains feasible given the joint probability density function, p, of the uncertain parameters. The stochastic flexibility index can be obtained directly by integrating p over the feasible region projected in the space of the uncertain parameters [11]. This integration can be done using Monte Carlo Sampling and requires a feasibility check of every sampled realization. Straub and Grossmann [50] later extended the framework of the stochastic flexibility index also for systems described by non-linear constraint functions. Since Monte Carlo Sampling can be computationally expensive due to the large number of samples required to cover the expected uncertainty space, Pulsipher and Zavala [51] suggested a mixed-integer conic program which can be used to compute a lower bound for the stochastic flexibility index.

In addition to the deterministic and the stochastic flexibility indexes, a third flexibility index has been proposed by Lai and Hui [52]: the volumetric flexibility index. The volumetric flexibility index quantifies the percentage of the expected uncertainty space that can be feasibly handled. In a geometric sense, the volumetric flexibility index describes the volumetric fraction of the hypervolume of the feasible uncertainty space, compared to the hypervolume of expected uncertainty space. The authors concluded that the approximation of hypervolume of the feasible space is challenging, and, recently, Zheng et al. [53] presented a novel approach to approximate the volumetric flexibility index based on the symbolic computation method.

A common application addressed in the field of design under uncertainty is the design and retrofit of heat recovery systems and heat exchanger networks (HENs). Typical examples of uncertain parameters affecting HENs are the variation of input conditions, i.e., temperatures and flow rates, but also uncertainty in design characteristics, such as heat transfer coefficients of heat exchangers (HEXs). Additionally, HENs are (usually) interconnected systems transferring heat between different process streams. Therefore, uncertainty easily propagates through the entire system leading to difficult challenges when designing HENs in such a way that target temperatures are met despite these uncertainties. The following section presents a comprehensive overview of methodologies to specifically design and retrofit HENs.

1.2.1 Design and retrofit of heat recovery systems

The recovery of excess process heat is an essential tool for increasing energy efficiency in the industrial sector. Generally, approaches to design and retrofit HENs can be grouped into graphical approaches based on Pinch Analysis, approaches based on mathematical programming, and hybrid approaches combining Pinch Analysis and mathematical programming [54]. However, many of the published approaches neglect the presence of uncertain parameters and instead assume only average operating conditions during the design stage. A good overview on approaches which allow for considering uncertainty is provided in the review paper on the synthesis of flexible HENs by Kang and Liu [55]. Some selected literature as well as recent development is discussed hereafter. For graphical approaches, Kotjabasakis and Linnhoff [56] developed an approach to mitigate the unwanted response of a HEN to variations based on sensitivity tables and systematic utilization of downstream paths. Additionally, Hafizan et al. [57] utilized the plus-minus principle, introduced by Linnhoff and Vredeveld [58], to visualize the impact of process modifications on the minimum utility target, to derive heuristics for HEN design synthe-

sis based on graphical methods if the respective HEN is subject to disturbances in inlet temperatures.

Mathematical programming has great potential when dealing with variations and uncertainty in operating data. In the 1980s, Floudas and Grossmann [59] developed an automated approach to generate HEN configurations which can operate at multiple operating points (multi-period operation). The authors suggested to solve the discretized design under uncertainty problem for a multi-period superstructure to obtain the HEN configuration which can operate feasibly at lowest cost at several discrete operating points. In a later publication [44], the authors extended their approach by including flexibility analysis sub-steps to ensure that the final design can operate not only at discrete operating points, but also at all operating points within the expected uncertainty space (i.e., steady-state flexible operation). In this context, the previously mentioned works by Pintarič and Kravanja [34] and Zirngast et al. [35] also enable the synthesis of HENs which allow for steady-state flexible operation. Several other authors have suggested approaches to synthesize HEN configurations which can operate feasibly at multiple periods while steady-state flexible operation is not (rigorously) addressed. More specifically, Aaltola [60] extended the (single-period) SYNHEAT model of Yee and Grossmann [61] to multiple periods. Additionally, Verheyen and Zhang [62] as well as Short et al. [63] improved the model of Aaltola [60] by considering more specific HEX design data. The multi-period HEN synthesis strategy was further applied by Tveit et al. [64] to industrial applications.

In addition to the approaches focusing on the synthesis of new HEN configurations, approaches for retrofitting existing configurations are also presented in the literature. An overview of the different approaches is presented in the review papers by Sreepathi and Rangaiah [54] and Čuček et al. [65]. Hereafter, retrofit approaches which allow for considering uncertainty are discussed. A common strategy to deal with variations in operating conditions from a Pinch Analysis perspective is to develop different retrofit proposals for a number of selected sets of operating conditions, e.g., annual, seasonal or monthly average values, as done by Persson and Berntsson [66]. The different design proposals are then evaluated and may be combined to achieve an operable and energy-efficient retrofit proposal, accounting for all considered operating points. Another strategy is to develop different retrofit proposals by applying a graphical retrofitting methodology (e.g., methodologies based on temperature driving force curves [67] or identifying retrofit bridges [68]) for a specific nominal point and analyse the network's response to variations in a separate analysis step. Recently, Lal et al. [69] applied this approach to obtain insights and to identify the best performing retrofit design proposal by means of Monte Carlo simulation.

Papalexandri and Pistikopoulos [45] suggested to obtain the cost-optimal retrofit of a given HEN configuration by solving the discretized design under uncertainty problem for a multi-period superstructure comprising all structural alternatives. The authors included a flexibility analysis sub-problem to check if the achieved design solutions allow for steady-state flexible operation, and they also suggested a strategy for handling cases where steady-state flexible operation cannot be achieved. Since the quantification of investment costs for all possible structural alternatives of a given HEN configuration (including re-sequencing of existing HEX units) can be challenging in real-world applications, Kang and Liu [70] suggested an alternative approach for retrofitting HENs operating in multiple pe-

riods which explicitly allows for re-sequencing of existing units. The authors suggested to solve the multi-period HEN synthesis problem (e.g., by following the approach by Aaltola [60]) to obtain a target for the retrofit followed by a strategy for re-sequencing existing HEX units to satisfy heat transfer area demand revealed during the targeting step. They further extended this method by suggesting different strategies to perform this matching of heat transfer area, which was shown to be especially beneficial for large-scale HEN retrofit problems [71]. Additionally, the authors presented a systematic strategy to incorporate multiple practical restrictions during the retrofitting process [72]. Recently, Stampfli et al. [73] presented an evolutionary two-level algorithm for multi-period HEN retrofit. The authors utilized a genetic algorithm to first optimize the structural layout (topology) followed by the application of a differential evolution algorithm to identify the heat loads of the different HEXs. Note that the approaches by Kang and Liu and Stampfli et al. lead to retrofit options of HENs which allow for cost-optimal and feasible operation at discrete operating points while steady-state flexible operation is not addressed.

1.2.2 Knowledge gap

As outlined in the previous sections, a variety of approaches can be found in the literature to address chemical processes design subject to uncertainty. In this context, several approaches have been specifically formulated to design or redesign HENs which operate during multiple periods. However, steady-state flexible operation is not consistently addressed, and, e.g., Aaltola [60] concluded that "the major difficulty in the proposed framework is to guarantee the network feasibility". In some of the aforementioned works (e.g., [31–35, 44, 45]), concepts relating to deterministic flexibility analysis, e.g., the calculation of the flexibility index or the identification of critical operating points are utilized to ensure that the final design allows for steady-state flexible operation.

In deterministic flexibility analysis, a mathematical model is needed to express the expected uncertainty space, and a hyperbox representation based on expected upper and lower bound values is commonly utilized. The hyperbox model has been proven to be satisfactorily accurate if the expected distribution of the uncertain parameter values is independent, i.e., no correlating trends between different uncertain parameters can be observed. However, for real applications, this assumption is not always valid. For industrial case studies particularly, it is likely that when defining the system boundaries and thereby also the input parameters, which often are subject to some uncertainty, upstream dependencies between (some of) these input parameters can be missed (due to size and complexity of the underlying system). Several studies [41, 51, 74] have investigated the impact of considering dependencies between uncertain parameters when modelling the expected uncertainty space for flexibility analysis. All these studies conclude that ignoring dependencies may lead to the flexibility analysis metric underestimating the actual flexibility of the process, sometimes even significantly, as a consequence of bad resemblance between the modelled and the actual expected uncertainty space. Another consequence of inexact modelling of the expected uncertainty space can be that identified critical operating points represent combinations of uncertain parameter values which are not expected to occur, which in turn can result in unnecessary overdesign of the equipment. In this context, Rooney and Biegler [75] reported differences in the obtained design parameter values, i.e., equipment sizes, when solving the discretized design under uncertainty problem using realizations of uncertain parameters drawn from an independent distribution

and from distributions which show a (positive or negative) correlation.

Consequently, further research is needed to specify situations in which the traditional hyperbox modelling approach is inaccurate to model the expected uncertainty space as well as suggestions for alternative modelling strategies. Additionally, to the author's knowledge, no approaches have been published in the literature to address a situation such as that described in Section 1.1.2, where operating conditions are expected to change due to planned long-term development such as a switch of feedstock.

In terms of design and retrofit of HENs which are subject to uncertainty, the development of graphical-insight based approaches has been limited when dealing with uncertain parameters. Such graphical approaches commonly rely to a large extent on time-consuming procedures since different design proposals must be evaluated and combined manually. Additionally, thorough strategies to ensure steady-state flexible operation have not been incorporated in graphical approaches. On the other hand, graphical approaches offer several benefits, especially when applied to large and complex case studies. For example, Stampfli et al. [73] concluded that "one of the key advantages of such methods is the efficient visualization of the problem [based on graphs], which allows stakeholder internal communication and development of practical solutions based on the engineers inputs". Consequently, further research is needed to fully harness the benefits of graphical-insight based approaches in combination with approaches based on mathematical programming and flexibility analysis in hybrid frameworks.

1.3 Aim and scope

The aim of this thesis is to propose theoretical and methodological development for designing and/or redesigning chemical processes which are subject to uncertain operating conditions with a special focus on heat recovery systems such as HENs. In order to reach this aim, specific research objectives were formulated which address some of the shortcomings of existing approaches:

O1: Develop a strategy to identify situations where the traditional approach using upper and lower bound values to approximate the space in which uncertain parameters are expected to vary can be inexact, and enhance the modelling of this expected uncertainty space in the aforementioned situations that can be applied to concepts relating to deterministic flexibility analysis, i.e., calculation of the deterministic flexibility index and identification of critical operating points.

O2: Develop an approach to perform (deterministic) flexibility analysis in situations when uncertain long-term development such as a switch of feedstock interferes with operational short-term disturbances (i.e., overlaying uncertainty sources) as exemplified in Section 1.1.2.

O3: Facilitate the modelling of (complex) HENs commonly found in industry and characterized by stream splitting and mixing, as well as recirculating streams and closed circulation loops.

O4: Develop approaches whereby valuable designer input (e.g., non-quantifiable knowledge, experience-based heuristics, etc.) can be combined with the efficiency of mathematical programming when addressing a design under uncertainty problem.

The research objectives $O1$ and $O2$ refer to theoretical development in the field of deterministic flexibility analysis whereas objectives $O3$ and $O4$ aim at methodological development in the field of design under uncertainty. The theoretical and methodological development resulting from these research objectives were utilized to develop a hybrid framework which allows for combining graphical-insight based design approaches with mathematical programming. The proposed framework can be utilized for designing and/or redesigning chemical processes/plants which are subject to uncertainty, aiming to identify the design/retrofit option which optimises a given objective (such as costs) while simultaneously guaranteeing steady-state flexible operation.

1.4 Appended papers

This thesis summarizes the work presented in six appended papers:

Paper I: Langner, C., Svensson, E., & Harvey, S. (2021). A computational tool for guiding retrofit projects of industrial heat recovery systems subject to variation in operating conditions. Applied Thermal Engineering, 182.

Paper II: Langner, C., Svensson, E., & Harvey, S. (2020). A framework for flexible and cost-efficient retrofit measures of heat exchanger networks. Energies, 13(6).

Paper III: Marton, S., Langner, C., Svensson, E., & Harvey, S. (2021). Costs vs. flexibility of process heat recovery solutions considering short-term process variability and uncertain long-term development. Frontiers in Chemical Engineering.

Paper IV: Langner, C., Svensson, E., Papadokonstantakis, S., & Harvey, S. (2022). Flexibility analysis of chemical processes considering overlaying uncertainty sources. Computer Aided Chemical Engineering, 49, (Proceedings of 14th International Symposium on Process Systems Engineering held in Kyoto, 2022).

Paper V: Langner, C., Svensson, E., Papadokonstantakis, S., & Harvey, S. (2023). Flexibility analysis using boundary functions for considering dependencies in uncertain parameters. Computers & Chemical Engineering, 108231.

Paper VI: Langner, C., Svensson, E., Papadokonstantakis, S., & Harvey, S. Novel reformulations for modelling uncertainty and variations in a framework for chemical process design. submitted to Industrial & Engineering Chemistry Research

Figure 1.5 illustrates how the research objectives $O1$ to $O4$ are addressed in the different papers, as well as the links between the different papers. In all appended papers, different HEN applications were utilized as case studies to illustrate the theoretical findings.

As shown in Figure 1.5, two different principles were investigated to analyse the effect of uncertain operating conditions on a suggested design or retrofit of a chemical process, and to further identify measures to ensure feasible operation when operating conditions vary. The two principles are sensitivity analysis and flexibility analysis. In **Paper I**, an approach is presented which utilizes automated sensitivity analyses. A computational analysis tool was developed which enables fast evaluation of the response of a HEN, i.e., temperatures and heat loads, when operating conditions change, design specifications are modified and/or operational settings are manipulated. A systematic methodology is presented for applying this type of sensitivity analyses in HEN retrofitting processes

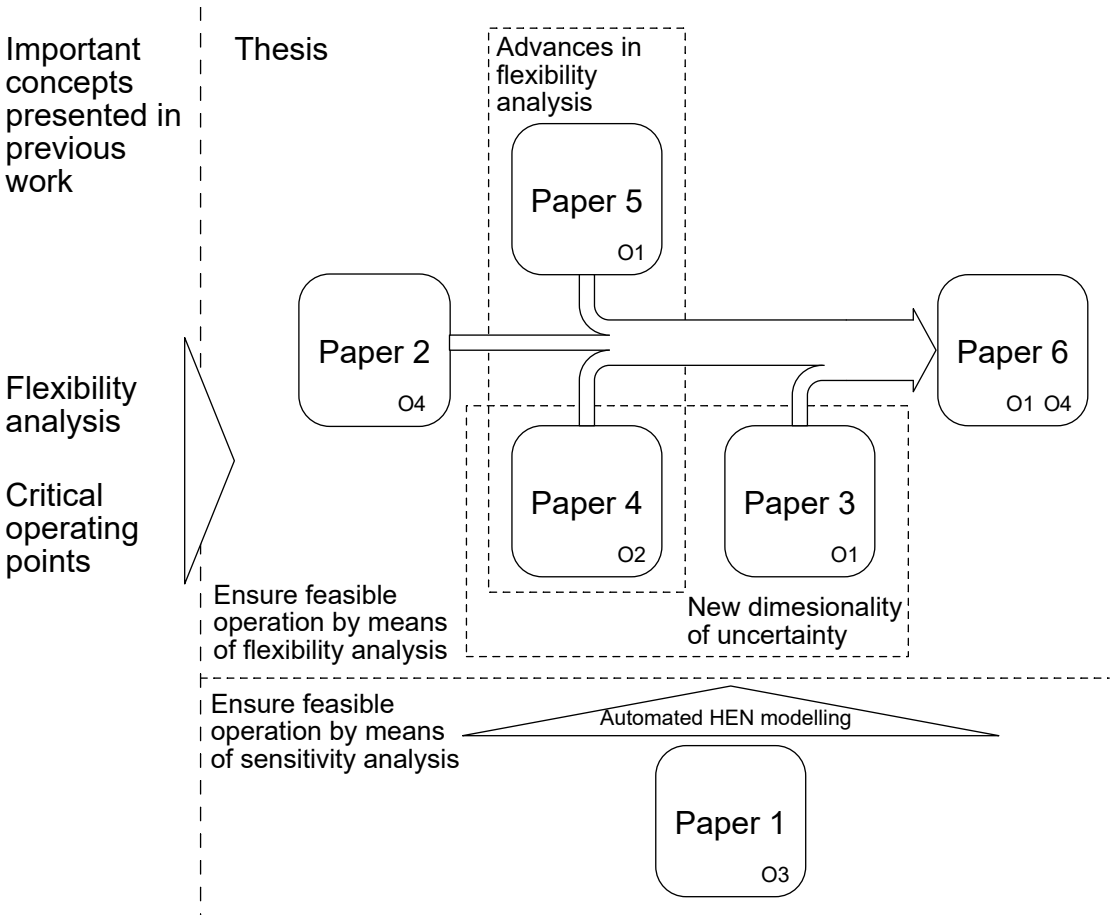


Figure 1.5: Overview of the links between the different papers appended to this thesis. Additionally, the research objectives $O1$ to $O4$ of the thesis are allocated to the different papers.

to evaluate the operability and energy efficiency of different retrofit proposals. However, since sensitivity analysis is based on the evaluation of discrete operating conditions, steady-state flexible operation cannot be rigorously guaranteed. Therefore, concepts based on sensitivity analysis were not pursued further, and this is further discussed in Chapter 6. Note that in **Paper I**, a strategy was outlined to automatically derive the complete set of equations necessary to describe the heat and mass balances as well as temperature constraints of a given HEN. This approach enables automatic modelling of HENs of any size, also taking into account the presence of common structural complexities such as stream splitting and mixing as well as recirculating streams and closed circulation loops. This strategy was adopted in the other papers to facilitate the mathematical modelling, especially of large-scale HENs, and avoid error-prone manual definition of mathematical constraints.

In **Paper II**, a step-wise framework was developed to achieve flexible and cost-efficient retrofit measures of (industrial) HENs. The proposed framework combines different concepts presented in the literature. Initially, different structural design proposals[2] are collected in a superstructure. These design proposals can be derived using graphical approaches enabling the incorporation of experience-based heuristics. Flexibility analysis is then utilized to reduce the superstructure by discarding design proposals which do not

[2]A structural design proposal describes the structural layout of the process including the placement of equipment while the design specifications, such as the size of the equipment, remain unspecified.

allow for feasible operation when operating conditions vary, i.e., structurally infeasible design proposals are discarded. Additionally, those combinations of the uncertain parameter values which require the largest equipment size and are thus critical for process operation are identified. In the final step, the discretized design under uncertainty problem is solved considering representative as well as the previously identified critical operating points to identify the most cost-efficient design proposal within the reduced superstructure which also allows for steady-state flexible operation.

In **Paper II**, it is assumed that the uncertain parameters vary around a mean value or nominal operating point, which is a common assumption in the literature. In **Paper III**, a new dimension of uncertainty was investigated, namely that the aforementioned uncertain parameters are influenced by an uncertain singular/rare event. Examples for such a singular/rare event are a permanent switch of feedstock (as presented in the motivational example in Section 1.1.2), a change of operational parameters required to comply with new emission legislation and/or a change in the production rate, among others. All these events may have a lasting effect on the operation of the process in question and a possible consequence is that the nominal operating conditions change temporarily or even permanently. In addition to the consequences of such planned/expected long-term development, operational short-term disturbances which affect the process during operation (at a distinct nominal operating point) are commonly present. Such short-term operational disturbances are comparable to the traditional interpretation of uncertainty meaning that they occur around a given nominal point. In **Paper III**, iterative evaluation of the deterministic flexibility index is proposed to identify if a process subject to both short-term disturbances as well as long-term development can operate feasibly at all expected operating points (before and after the change). Additionally, the paper proposes a strategy to enable consideration of planned long-term development in retrofitting projects. The outlined strategy ensures that equipment size is sufficient to allow for steady-state flexible operation at (different) nominal operating points.

Paper III describes situations where long-term development interferes with short-term operational disturbances. In this context, the strategies presented in **Paper III** require that the consequences of the long-term development need to be approximated with satisfactory accuracy (i.e., the nominal operating conditions after the singular/rare event need to be quantified) to perform flexibility analysis at the different nominal operating points in an iterative fashion. In **Paper IV**, a new approach to deterministic flexibility analysis is presented which allows for considering uncertainty related to the change of the nominal operating conditions. The aim of this new approach is to identify the maximum feasible change of the nominal operating conditions so that the process can cope with the expected short-term operational disturbances after the nominal operating conditions have changed. The new approach involves reformulations of the original flexibility index problem presented by Swaney and Grossmann [2]. In this context, the presented reformulations enable to identify if the installed process equipment in the second motivational example (see Section 1.1.2) would allow for feasible operation also after the implementation of biomass co-processing.

In **Papers II - IV**, it is assumed that the variation of the uncertain parameters is independent, i.e., no correlating trends can be observed between different uncertain parameters in the operating data. Therefore, the expected uncertainty space was modelled using

upper and lower values (traditional hyperbox approach) which reflects this assumption of independent uncertain parameters. However, as mentioned in Section 1.2.2, this assumption is not always valid. In **Paper V**, the literature was reviewed to identify approaches which consider parameter dependencies when modelling the expected uncertainty space to be utilized in deterministic flexibility analysis. One alternative approach compared to the hyperbox modelling approach was identified. However, **Paper V** shows that the flexibility index based on this alternative approach can overestimate the feasible variation range of the uncertain parameters, i.e., variations which are considered to be feasible according to the flexibility index are in fact not feasible. To avoid this overestimation while allowing for a better representation compared to the hyperbox model, a novel approach to model the expected uncertainty space is presented in **Paper V**. Additionally, a generic mixed-integer (non-)linear program (MINLP) is presented to calculate the deterministic flexibility index based on the presented approach.

Finally, in **Paper VI**, the framework presented in **Paper II**, was reworked to include the development presented in **Papers III - V**. In this context, the proposed framework includes strategies for considering parameter dependencies as well as the presence of multiple independent operating periods (a phenomenon which is connected to the findings presented in **Paper III**) when modelling the expected uncertainty space. Additionally, a novel algorithm is presented to identify critical operating points which is also applicable when the uncertainty space is modelled using the previously presented strategies. The framework presented in **Paper VI** is applicable to general chemical process design, and further allows to be applied in both greenfield and retrofit cases.

2
Theoretical background

2.1 Flexibility analysis

Flexibility analysis originates from the aspiration to avoid unnecessary overdesign of the equipment while guaranteeing steady state flexible operation [11]. Flexibility analysis denotes different concepts to evaluate the capability of a physical system to react towards uncertainty (e.g., disturbances) in order to maintain feasible operation [11]. In this context, feasible operation is achieved if the physical system with given equipment (sizes) reaches pre-defined target values which can be formulated as mathematical constraints. A well-established concept to quantify a system's ability to cope with uncertainties (i.e., for performing flexibility analysis of chemical processes) is the deterministic flexibility index which was introduced by Swaney and Grossmann [2]. The identification of critical operating points which require the largest overdesign is also related to flexibility analysis. Hereafter, the theoretical background of the deterministic flexibility index as well as the identification of critical operating points are presented.

2.1.1 Deterministic flexibility index

The core idea of the flexibility index is to provide a scalar value ≥ 0 which scales a geometric shape which is used to model the expected uncertainty space in the (hyper-)space of the uncertain parameters. More specifically, the returned value of the flexibility index problem is the scaling factor for which the scaled geometric shape intersects with the boundary of the feasible region meaning that for any higher value of the scaling factor parts of the shape would be outside the feasible region. The feasible region defines the operation and the functional purpose of the (chemical) process, and can be expressed as equality (2.1a) and inequality constraints (2.1b):

$$h_i(d, x, z, \theta) = 0;\ i \in I, \tag{2.1a}$$

$$g_j(d, x, z, \theta) \leq 0;\ j \in J. \tag{2.1b}$$

In Eq. (2.1a) and Eq. (2.1b) the parameters which are uncertain are denoted by θ. Additionally, d is the vector of design parameters, x corresponds to the state variables and z is used the express the degrees of freedom or control variables of the process. In the mathematical sense, control variables or degrees of freedom are present if the equation system formed by the equality constraints, $i \in I$, as well as the union set of control and state variables is under-determined, i.e., not all variables in the union set of control and state variables are explicitly determined by the equation system.

Swaney and Grossmann [2] modelled the expected uncertainty space using the expected extreme values, which they expressed as expected deviations, $\Delta\theta^-$ and $\Delta\theta^+$, from a

nominal or mean operating point, θ_N. Consequently, the modelled uncertainty space can be imagined as a hyperbox, i.e., a multi-dimensional rectangle such as a cuboid for three dimensions. This means that the flexibility index can be imagined as the ratio between the largest scaled hyperbox within the feasible region and the hyperbox defined by the expected extreme values. The mathematical formulation of the scaled hyperbox is given in Eq. (2.2). Note that the scaling factor is denoted by δ.

$$T_{box}(\delta) = \left\{\theta_i \mid \theta_{i,N} - \delta\Delta\theta_i^- \leq \theta_i \leq \theta_{i,N} + \delta\Delta\theta_i^+\right\} \forall\ \theta_i \in \theta \tag{2.2}$$

To obtain the flexibility index, Swaney and Grossmann [2] defined the flexibility index problem. The aim of the problem is to identify the maximum value of δ for which none of the constraint functions describing the process (i.e., Eq. (2.1a) and Eq. (2.1b)) is violated, i.e., the maximum value of δ for which the operational targets of the process are met given the (fixed) design specifications of the process. Consequently, if the index is greater or equal to 1, the process can be operated for all expected variation. The mathematical formulation of the flexibility index problem can be found in Appendix A. For two uncertain parameters, this search for the largest scaled rectangle within the feasible region (i.e. feasible uncertainty space) is illustrated in Figure 2.1. In Figure 2.1, the feasible uncertainty space is smaller than the expected uncertainty space, i.e., the rectangle corresponding to the expected variations. Thus, feasible operation with respect to all expected variations cannot be guaranteed, i.e., the flexibility index is < 1. Note that in Figure 2.1, a point has been marked as critical point. This is discussed further in Section 2.1.3.

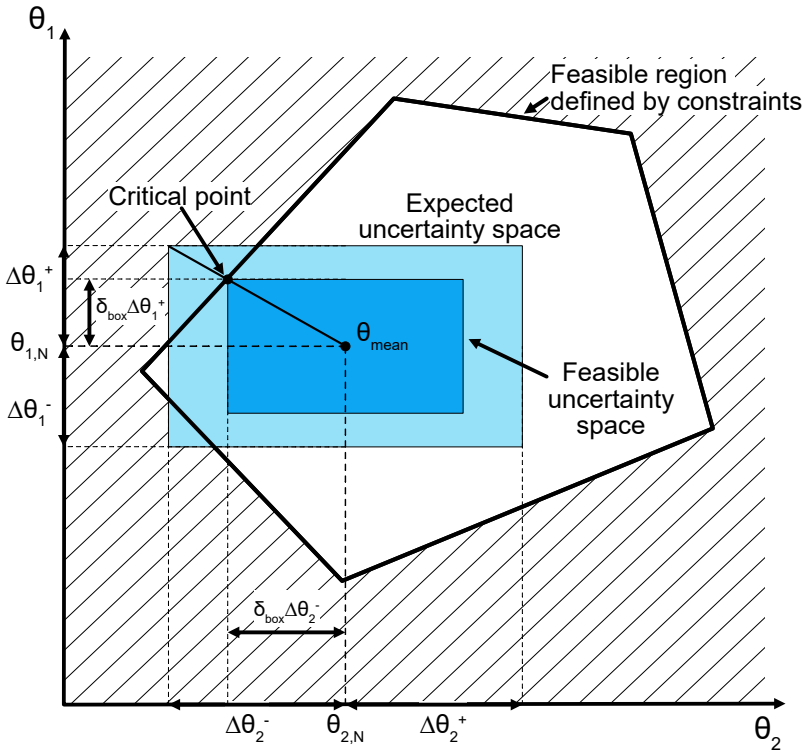


Figure 2.1: Visualization of the flexibility index by Swaney and Grossmann [2] for two uncertain parameters. The expected uncertainty space was modelled using a hyperbox model considering the expected lower and upper bound values of the uncertain parameters which can be visualised as a rectangle. The expected uncertainty space is scaled to identify the largest rectangle (with the same aspect ratio) which can be inscribed in the feasible region.

As mentioned in Section 1.2, different algorithms have been suggested to identify the flexibility index for a given process subject to uncertain parameters, including the active set approach based on an MI(N)LP formulation for the flexibility index problem proposed by Grossmann and Floudas [41]. The advantage of the active set approach is that, unlike other proposed approaches, an iterative evaluation of the vertices of the expected uncertainty space can be avoided. More information of the active set approach as well as its mathematical formulation is provided in Appendix A.

2.1.2 Flexibility analysis considering dependencies in the uncertain parameters

As mentioned in Section 1.2.2, several studies in the literature [41, 51, 74] point out that the deterministic flexibility index based on the hyperbox model potentially underestimates the flexibility of a process if dependencies in the uncertain parameters are present. In this context, the term "underestimation" refers to the phenomenon that the flexibility index indicates only a small feasible interval $[\delta_{box}\Delta\theta_i^-, \delta_{box}\Delta\theta_i^+]$ while a larger share of the expected operating points or expected realizations of the uncertain parameters are within the feasible region. In **Paper V**, the literature was analysed to identify suitable approaches for modelling the expected uncertainty space when parameter dependencies are present which are further applicable for deterministic flexibility analysis. One approach [41] was identified from the literature review, and it is summarized hereafter. More detailed information can be found in **Paper V**.

Grossmann and Floudas [41] proposed to express the dependent uncertain parameters through algebraic equations, $f(\theta) = 0$, which can be included as additional constraints in the flexibility index problem (compare Problem (A.2)). Although not explicitly formulated by the authors, the approach can be generalized by reformulating the mathematical formulation for the hyperbox model, $T_{box}(\delta)$. For this purpose, the uncertain parameters are grouped into independent uncertain parameters, θ_{ind}, and dependent uncertain parameters, θ_{dep}. The dependent uncertain parameters are then expressed by single equation models. The generalized reformulation of the hyperbox model based on Grossmann and Floudas [41] is given in Eq. (2.3).

$$T_{corr,box}(\delta) = \begin{cases} \left\{\theta_i \mid \theta_{i,N} - \delta\Delta\theta_i^- \leq \theta_i \leq \theta_{i,N} + \delta\Delta\theta_i^+\right\} \ \forall\ \theta_i \text{ in } \theta_{ind} \\ \left\{\theta_j \mid f_j(\theta_{ind}) = 0\right\} \ \forall\ \theta_j \text{ in } \theta_{dep} \end{cases} \tag{2.3}$$

An intuitive approach to express the dependent uncertain parameters is to utilize (linear) regression models based on the independent uncertain parameters. A conceptual illustration of the flexibility index for a linear single equation model in comparison to the hyperbox model is shown in Figure 2.2.

In the two-dimensional case shown in Figure 2.2, the result of the flexibility index (the scaling parameter δ) is different when the expected uncertainty space is modelled using a single equation model compared to the hyperbox model. More specifically, δ_{corr} is larger than δ_{box} which implies that the underestimation of the flexibility by means of the flexibility index based on the hyperbox model could possibly be decreased. On the other hand, it is important to ensure that the larger value of δ_{corr} compared to δ_{box} is not the result of an overestimation of the flexibility.

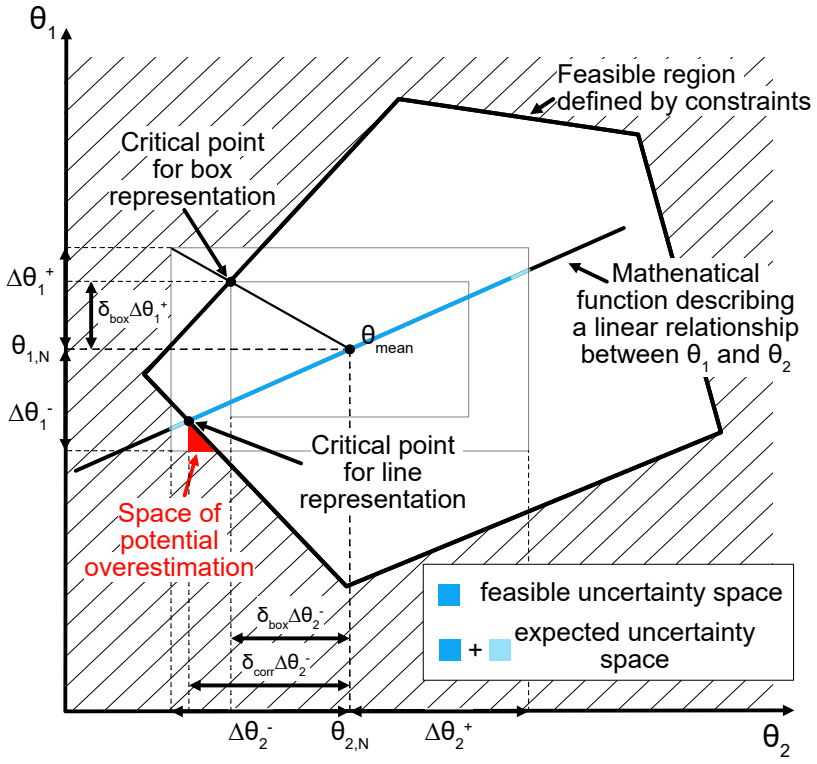


Figure 2.2: Conceptual illustration of the deterministic flexibility index using a single equation (regression) model to capture a dependency between the uncertain parameters. Note, the presence of expected operating points in the red coloured area would indicate that the feasible interval indicated by the flexibility index $[\delta_{corr}\Delta T_2^-,\ \delta_{corr}\Delta T_2^+]$ overestimates the flexibility since these operating points would be within the feasible interval but are de facto infeasible.

When using single equation (regression) models, there is a high likelihood that expected realizations of the uncertain parameters are not included in the modelled uncertainty space, i.e., the modelled uncertainty space underestimates the actual uncertainty space. This underestimation of the actual uncertainty space leads to the possibility that the flexibility index based on single equation models overestimates the flexibility of a process. In this context, overestimation of flexibility describes the observation that operating points which should be feasible according to the analysis (i.e., the values of the independent uncertain parameters are within the interval $[\delta_{corr}\Delta\theta_{ind,i}^-,\ \delta_{corr}\Delta\theta_{ind,i}^+]$) are indeed outside of the feasible region and thereby not feasible. For the two-dimensional example shown in Figure 2.2, this space of potential overestimation is shown in Figure 2.2. A numerical example where the deterministic flexibility index based on a single equation model overestimates the flexibility of a given process is presented in **Paper V**.

The observations described in the previous paragraph relate to the main drawback of single equation regression models, which is that they are only exact if the strongest possible agreement exists between the correlated uncertain parameters. Commonly, correlated uncertain parameters in chemical processes agree only to some extent, which means that single equation regression models are able to capture the trend between these uncertain parameters well while neglecting operating points which deviate from this trend (i.e., operating points caused by "other" sources of uncertainty). Note that overestimating the flexibility of a process can have severe consequences since the infeasibility of certain operating conditions may not be identified before actual operation. Consequently, (very)

costly retrofits may be required which are likely to exceed the cost of (unnecessary) overdesign which may occur as a result of underestimating the flexibility of the process.

2.1.3 Critical Operating Points

In Figures 2.1 and 2.2, the points at which the feasible uncertainty space coincides with the boundary of the feasible region are marked as critical points[1]. These points define the sizes of the feasible uncertainty spaces, i.e., the area of the rectangle in Figure 2.1 and the length of the line in Figure 2.2. Note that at the marked critical points, feasible operation is possible but even a small deviation from the critical parameter values may lead to infeasibility. This observation is of special interest if some of the constraints forming the feasible region are functions of some design parameters, d, which, e.g., can represent the size of certain equipment. If the size of the feasible uncertainty space is limited by constraint(s) depending on d, debottlenecking may be possible by manipulating the respective design parameter(s), enabling a larger feasible uncertainty space (if desired, e.g., if the flexibility index is < 1). Note that in general not all constraint functions which form the feasible region (i.e., Eq. (2.1a) and Eq. (2.1b)) are dependent on d. Those constraints, which are dependent on d and therefore can be manipulated by increasing (or decreasing) the equipment size, are hereafter denoted as design constraints.

When conducting so-called structural flexibility analysis, design constraints are discarded, and only structural constraints are considered, i.e., unlimited equipment size is assumed to be available (see, e.g., [76, 77]). This implies that the value of the structural flexibility index is an upper bound for the general flexibility index, which also considers the design constraints. This is illustrated by an example in Figure 2.3. For this example, the feasible uncertainty space which is limited by structural constraints is larger than the expected uncertainty space, i.e., the structural flexibility index is > 1. However, when including the design constraint for the installed or planned equipment size, the feasible uncertainty space is smaller than the expected uncertainty space resulting in a flexibility index which is < 1.

Figure 2.3 shows that if design constraints are included in the flexibility analysis and the resulting flexibility index is smaller than the structural flexibility index, the design constraints and thereby the corresponding equipment size limit the value of the flexibility index. In this context, Pintarič and Kravanja [31, 32] assumed that if, for a given structural design proposal, the structural flexibility index is > 1, it is possible to determine at least one set of values for the design parameters which yields a (general) flexibility index of exactly 1. To identify this set of values for the design parameters, Pintarič and Kravanja proposed to identify those combinations of the uncertain parameter values within the expected uncertainty space that require the largest overdesign of process units in order to allow for steady-state flexible operation. Note that this idea is based on the work by Halemane and Grossmann [29] who observed that if feasible operation at the critical operating points can be guaranteed, the same applies for any other point within the expected uncertainty space.

[1] The term critical operating points was introduced by Halemane and Grossmann [29] who defined critical operating points as those combinations of the uncertain parameter values for which the feasible region of operation is the smallest.

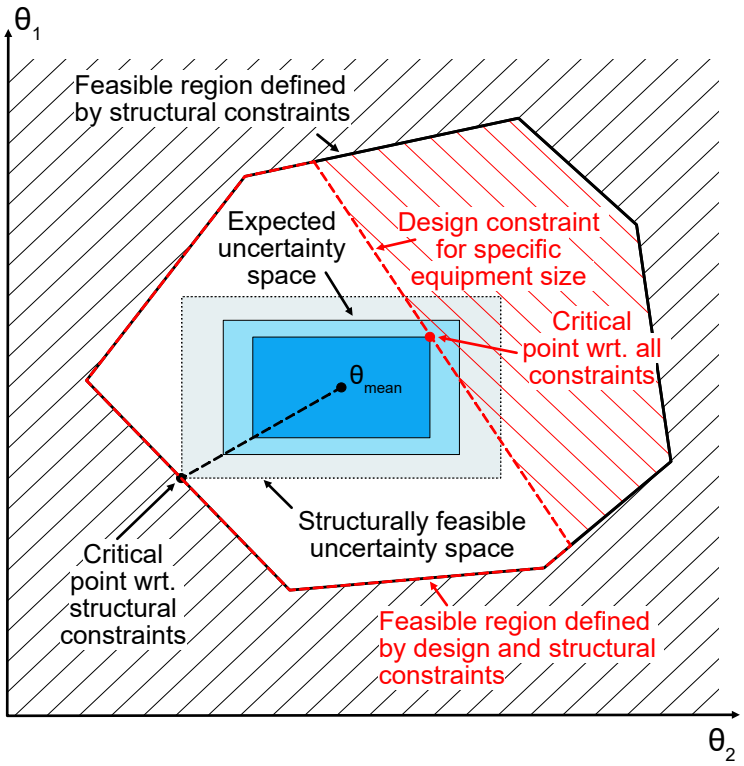


Figure 2.3: A theoretical example to illustrate the difference between the structural flexibility index where design constraints are discarded and the (general) flexibility index where design constraints are included.

In the example shown in Figure 2.3, the necessary change of the equipment size can be identified graphically by parallel shifting the design constraint to the right, so that the design constraint intersects with the upper right corner point of the expected uncertainty space. This is possible since in this two-dimensional example, the identification of the critical operating point (upper right corner point) is trivial. However, as Pintarič and Kravanja [31, 32] also pointed out, in more complex cases, i.e., if more than two uncertain parameters and/or more design parameters are present, several combinations for the design parameter values may yield a flexibility index of 1 (when design constraints are included). In such cases, the aim should be to identify the unique combination which simultaneously yields smallest cost and a flexibility index of 1. To achieve this, the authors suggested to solve the bi-level optimization problem in which each design variable is individually maximized while simultaneously minimizing a given cost function. Furthermore, the authors suggested several algorithms to solve this bi-level optimization problem and the interested reader is referred to their work for further information. Note that critical operating points are dependent of the nature and placement of the equipment, i.e., for each structural layout a unique set of critical operating points can be identified.

2.2 Design under Uncertainty

In this section, the generic problem formulation of the design under uncertainty problem is presented. Commonly, a design under uncertainty problem is addressed using the stochastic two-stage with recourse formulation [14, 18]. The challenge connected with a design under uncertainty problem is the numerical integration over the expected un-

certainty space, i.e., the distribution of uncertainty. In sampling-based approaches, this integration is performed by a discretization of the uncertainty space [14, 34]. This discretization enables to solve the discretized equivalent to the stochastic two-stage design under uncertainty (with recourse) formulation which is given in Problem (2.4).

$$\begin{aligned} \min_{z_s,d} \; & G(d) + \sum_{s\in S}[w_s \cdot C_s(d, z_s, \theta_s)] \\ s.t. \; & h_i(x_s, z_s, d, \theta_s) = 0; \; i \in I \\ & g_j(x_s, z_s, d, \theta_s) \leq 0; \; j \in J \\ & d \geq 0 \\ & x_s, z_s, d, \theta_s \in \mathbb{R} \end{aligned} \Bigg\} \; s \in S. \tag{2.4}$$

The objective function in Problem (2.4) consists of two parts. The first term $G(d)$ describes the cost connected to investment in process units to provide the capacity necessary for steady-state flexible operation - this selection of the design parameters d is referred to as the first-stage decisions. The second term is the expected value of the operating costs which is approximated via summation of the costs for discrete scenarios C_s multiplied by the probabilities of the scenarios w_s. The expected value of the operating costs is dependent on the selection of the design parameters, d, and the second-stage decisions which are the selection of the control variables, z_s. The design problem is subject to equality constraints, $h_i \; \forall \; i \in I$, which often result from the models of the process units and the interconnections between them, and inequality constraints, $g_j \; \forall \; j \in J$, which commonly arise from product property constraints and capacity specifications. Note that state variables, x_s, are explicitly included in Problem (2.4).

Due to the discretization of the scenarios, $s \in S$, in Problem (2.4), the numerical integration over the expected uncertainty space to approximate the expected value of the operating costs as well as the selection of the design parameter values can be done simultaneously. In Problem (2.4), the discrete choices of the uncertain parameters to represent the discrete scenarios are explicitly denoted by θ_s. Problem (2.4) can consequently be seen as a multi-period (Non-)Linear Program (NLP) which aims to minimize some objective value (e.g. the total annualized cost, TAC) of the design or retrofit of a chemical process considering its operation during discrete scenarios or operating periods.

When designing or retrofitting a chemical process which is subject to uncertainty, a sufficient number of realizations of the uncertain parameters, i.e., scenarios, needs to be considered to allow for steady-state flexible operation as well as a fair approximation of the objective function value, e.g., TAC. To ensure steady-state flexible operation within the expected uncertainty space, critical operating points (see Section 2.1.3) can be identified and included in the constraints of Problem (2.4) (i.e., $h_i \; \forall \; i \in I$ and $g_j \; \forall \; j \in J$). Note that critical operating points usually represent extreme operating conditions and not typical conditions that occur during normal operation. Therefore, operation at the critical operating points should not be considered in the objective function of Problem (2.4), i.e., critical operating points should not be considered when approximating the expected value of the operating costs, C_s.

To obtain a good approximation of the objective function value, a common approach is to identify the most representative scenarios for operation of the process, i.e., scenarios that are expected to be a good representation of typical conditions that occur during

normal operation. Such representative operating scenarios should represent the operating conditions during specific time periods (e.g., different seasons or different production campaigns), and should allow for adequate approximation of the expected objective value. When denoting the set of representative scenarios with OP and the set of critical operating points with CP, the operating costs, C_s, are calculated for all scenarios in OP while the constraints of Problem (2.4) are evaluated at all scenarios/points in the union, $OP \cup CP$.

2.3 Heat Exchanger Network Modelling

As outlined in Section 1.2, Heat Exchanger Networks (HEN) are usually subject to uncertainty in operating data which needs to be accounted for during the design process. For applying the design under uncertainty approaches discussed in this thesis, mathematical modelling of the respective HEN is necessary, i.e., the definition of the mathematical equations describing the physical constraints of a HEN such as energy and mass balances. Such modelling can be burdensome and error-prone if conducted manually, especially for large and complex industrial HENs. This thesis aims to facilitate the mathematical modelling of HENs by an automated modelling strategy. Hereafter, the thermodynamic fundamentals of HEN modelling are described.

A HEN is usually characterized by a number of fluid streams which interchange heat by means of Heat Exchangers (HEXs). The heat exchange can be direct or indirect (i.e., by means of a heat transfer medium). The network temperatures describe the temperature change over the different streams due to heat exchange in HEXs, stream splits and mixing of streams. This implies that the number of (unknown) network temperatures for a fixed HEN structure can be calculated a priori [78]. For a network with N streams, n_E process-to-process HEXs, n_S stream splits (1 stream splits into 2 streams), n_M stream mixing points and n_U utility HEXs, the number of (unknown) network temperatures N_T is:

$$N_T = N + 2 \cdot n_E + 2 \cdot n_S + n_M + n_U. \tag{2.5}$$

The number of (unknown) temperatures N_T is reduced by N if the supply temperatures for all streams are specified. The temperature change over the different streams can be described by a set of equations which depend on the physical components present in the HEN. Common physical components are:

- Heat Exchangers (HEXs):
  - direct process-to-process HEXs,
  - indirect process-to-(heat transfer medium) HEXs,
  - utility HEXs (e.g., steam heaters or coolers),
- Splitters, and
- Mixers.

Other physical components such as reactors can be found in HENs, but these components were not considered explicitly in this thesis. On the other hand, components such as reactors can be modelled implicitly by means of identity changes of process streams. Switch models were used to model these identity changes of process streams and further explanation can be found in Section 2.3.3.

2.3.1 Heat Exchanger Modelling

A common approach to model HEXs in HENs is formulating energy balances on the hot and cold stream side of each HEX. Additionally, heat transport equations are used to account for the design characteristics of a HEX since the heat exchange in a HEX is constrained by the available surface area. This is formulated in Equations 2.6a to 2.6c:

$$Q_{HEX} = CP_h \cdot (T_{h,in} - T_{h,out}), \tag{2.6a}$$

$$Q_{HEX} = CP_c \cdot (T_{c,out} - T_{c,in}), \tag{2.6b}$$

$$Q_{HEX} = U_{HEX} \cdot A_{HEX} \cdot \Delta T_{LM}. \tag{2.6c}$$

In Equations 2.6a to 2.6c, the index h is used for hot process (or secondary) streams and the index c is used for cold process (or secondary) streams. Consequently, Q_{HEX} is positive if heat is transferred from the hot stream to the cold stream connected by means of the HEX. In several mathematical HEN models, this is an additional constraint (i.e., $Q_{HEX} \geq 0$). In order to model specific HEX types such as shell-and-tube HEXs or plate HEXs, standard methods such as the P-NTU method can be used (see e.g., [79, 80]).

In industrial HENs, bypasses are usually present around HEXs for control purposes. These bypasses can be modelled explicitly (see e.g., Appendix B.1 of [81]). However, as explicit modelling introduces additional non-linearities, alternative modelling approaches have been investigated. One possibility commonly found in the literature (e.g., in [44]) is to relax the equality sign in Equation 2.6c by an inequality ($\leq$) for those HEXs which feature a bypass. This way, it is assumed that the effective area of a HEX can be smaller than the installed area which gives a similar result as a bypass. Additionally, physical operating constraints for the temperatures must be defined. For a counter-current HEX, these operating constraints are given in Equations 2.7a and 2.7b:

$$T_{h,in} - T_{c,out} \geq \Delta T_{min}, \tag{2.7a}$$

$$T_{h,out} - T_{c,in} \geq \Delta T_{min}. \tag{2.7b}$$

In Equations 2.7a and 2.7b, ΔT_{min} must be larger or equal to 0 if heat is transferred from the hot to the cold stream. Additionally, in comparison to available mathematical HEN models in the literature, in this thesis it was considered that Equations 2.7a and 2.7b cannot be evaluated independently of Equations 2.6a to 2.6c. If, for example, a HEX is entirely bypassed (i.e., $Q_{HEX} = 0$), the physical operating constraints 2.7a and 2.7b must not be strictly obeyed as the HEX is out of operation. In order to model this mathematically, Equations 2.7a and 2.7b can be multiplied by the temperature difference given on the right-hand side of either Equation 2.6a or Equation 2.6b[2]. This way, it is ensured that the physical operating constraints only are meaningful if the HEX is transferring heat.

2.3.2 Modelling of Stream Splitting and Mixing

Besides HEXs, splitters and mixers are components commonly present in HENs. To model stream splitting and mixing, energy and mass balances are important to consider. In the

[2]Note that such a strategy leads to (additional) non-linearities in a HEN model which may be undesirable.

case of stream splitting, isothermal splitting can usually be assumed. Therefore, for a split of one stream into n streams, the energy balance can be simplified:

$$T_{out,1} = T_{out,n} = T_{in}. \tag{2.8}$$

The mass balance for a split of one stream into n streams is:

$$\sum_{i=1}^{n} \dot{m}_{out,i} = \dot{m}_{in}. \tag{2.9}$$

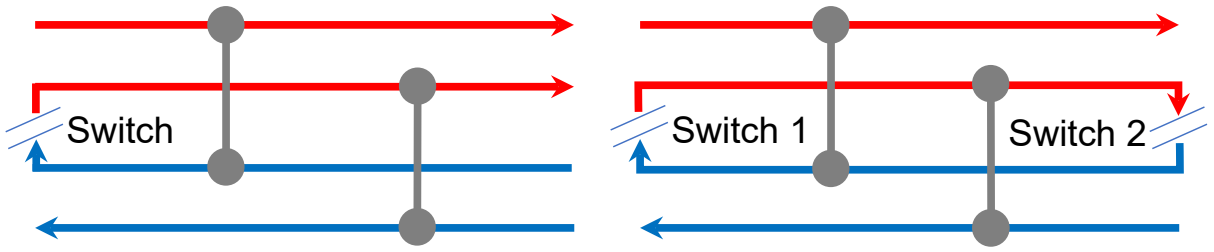
In the case of mixing n input streams to one output stream, mass and energy balances are given in the following (assuming mixing of ideal fluids):

$$\sum_{i=1}^{n} (\dot{m}_{in,i} \cdot cp_{in,i} \cdot T_{in,i}) = \dot{m}_{out} \cdot cp_{out} \cdot T_{out}, \tag{2.10a}$$

$$\sum_{i=1}^{n} \dot{m}_{in,i} = \dot{m}_{out}. \tag{2.10b}$$

2.3.3 Modelling of Identity Changes

In this thesis, the identity of a stream is defined by whether the stream releases heat (hot stream) or receives heat (cold stream) in the HEXs connected to the stream. Consequently, an identity change implies that a stream changes from being identified as cold to hot or vice versa. These identity changes may occur in (industrial) HENs, e.g., when streams are re-circulated or in closed circulation loops. Figure 2.4 visualizes the identity change of a) a re-circulating stream and b) a closed circulation loop by means of switches. In order to model an identity change, one solution is to model a new stream for each identity change. However, inlet conditions to these new streams are defined by the conditions prior to the identity change (see Figure 2.4). Therefore, (trivial) pairs of equations describing the relation between the temperatures, $T_{out,i}$ and $T_{in,i}$, as well as the CP-values, $CP_{out,i}$ and $CP_{in,i}$, over the identity change $i \in IC$ can be derived.



(a) Identity change in a re-circulating stream

(b) Identity changes in a closed circulation loop

Figure 2.4: Visualization of the identity change of a stream (i.e., a hot stream becomes a cold stream or vice versa).

3

Theoretical Development

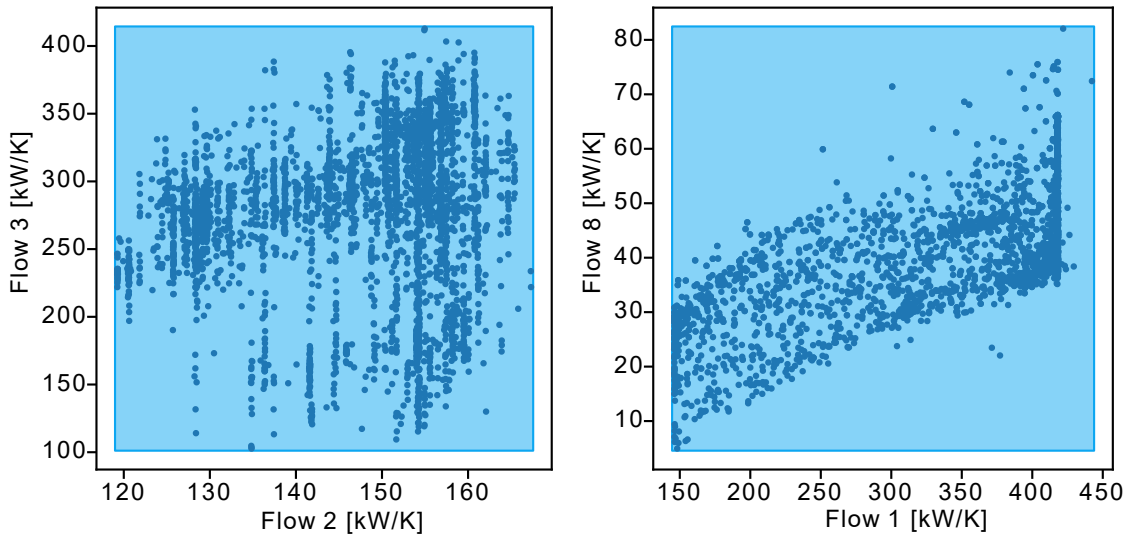
3.1 Enhanced modelling of the expected uncertainty space for deterministic flexibility analysis

For deterministic flexibility analysis, the expected uncertainty space needs to be modelled. This modelled expected uncertainty space should represent the actual expected uncertainty space as accurately as possible. Traditionally, as outlined in Section 2.1, the expected lower and upper bound values of the uncertain parameters are utilized to model the expected uncertainty space as a hyperbox. For independent distributions of uncertainty such a hyperbox shape is satisfactorily accurate. On the other hand, if dependencies are expected in the uncertain parameters, it was shown in the literature that the hyperbox approach to model the expected uncertainty space can be inaccurate due to a bad resemblance with the actual uncertainty space. However, as outlined in Section 2.1.2, the alternative approach to model the expected uncertainty space for deterministic flexibility analysis if parameter dependencies are present (i.e., single equation models) has significant shortcomings. More specifically, the value of the flexibility index may overestimate the flexibility of a process which can have severe consequences. Therefore, in Section 3.1.1, a novel approach to model the expected uncertainty space is presented.

Furthermore, this thesis aims to provide strategies to identify situations where the aforementioned bad resemblance between the hyperbox model and the actual uncertainty space may occur. To identify such strategies, the influence of parameter dependencies on the expected uncertainty space was investigated in more detail. For illustrative purposes, the available measurement values of the uncertain parameters penetrating the pulp mill subsystem described in Section 1.1.1 were analysed for parameter dependencies using the correlation coefficient of Pearson[1]. In Figure 3.1, the measured operating points are shown in the space of the flow rates of streams 3 and 2 (Figure 3.1a) as well as in the space of the flow rates of streams 8 and 1 (Figure 3.1b). Note that Pearson's correlation coefficient was identified to be 0.15 for the measured flow rates of streams 3 and 2, and 0.58 for the measured flow rates of streams 8 and 1.

In addition to the measured operating points, Figures 3.1a and 3.1b show the rectangle defined by the respective minimum and maximum measured values. Figure 3.1a confirms that the hyperbox approach, here resulting in a rectangular shape, reflects the actual uncertainty space with good accuracy if the (linear) relationship between two uncertain

[1]Pearson's correlation coefficient, $\rho(X,Y)$, determines the strength and the direction of the linear relationship between two random variables, X and Y, with $0 \leq |\rho(X,Y)| \leq 1$ (see e.g., [82]). Note that larger absolute values indicate stronger correlations and that the sign indicates the direction of the correlation (correlated/anti-correlated).



(a) **(b)**

Figure 3.1: Operating points measured for the pulp mill subsystem described in Section 1.1.1 in the space of the flow rates of streams 3 and 2 (a) as well as in the space of the flow rates of streams 8 and 1 (b). Additionally, the rectangular shape defined by the hyperbox approach to model the expected uncertainty set is shown for both cases in light blue.

parameters is weak. Additionally, Figure 3.1b confirms the opposite if the (linear) relationship between two uncertain parameters is stronger. In this context, Figures 3.1a and 3.1b show that the hyperbox model overestimates the actual uncertainty space while this overestimation is significantly bigger if the (linear) relationship between two uncertain parameters is stronger. The problem connected with overestimating the actual uncertainty space is that combinations of the uncertain parameter values are included in the modelled uncertainty space which are not expected to occur. Consequently, in such cases operating conditions which are not expected to occur may limit the scaling of the expected uncertainty space resulting in wrong conclusions drawn from the flexibility analysis. More specifically:

- as outlined in Section 2.1, the obtained value for the flexibility index may underestimate the feasible variation range in which uncertain parameters may vary, and
- the identified critical operating points may indeed not be expected to occur.

The findings gained from Figure 3.1 can be generalized resulting in a strategy to identify situations where modelling the expected uncertainty space by means of (overall) expected upper and lower bound values can be inaccurate resulting in the aforelisted consequences. More precisely, this thesis proposes to investigate, prior to modelling the uncertainty space, if there is an indication that certain combinations of the (individually expected) uncertain parameter values are not expected to occur.

Based on this strategy, another situation was identified where a single hyperbox model may be an oversimplified representation of the expected uncertainty space. More specifically, flexibility analysis may be conducted based on time-dependent data, e.g., historic operating data, and such data may reveal that certain combinations of the uncertain pa-

rameters may only occur during certain periods of the time-horizon considered for the flexibility analysis. For example, Figure 3.2 shows the inlet temperature of the air used for drying at a Swedish pulp mill over a three-year period. The operating data shows both irregular fluctuations as well as a regular seasonal trend. To highlight this seasonal trend, the operating data was divided into seasonal periods of three months each[2]. In addition to the overall mean value and the maximum and minimum temperature measured over the three-year period, Figure 3.2 shows the seasonal mean values as well as and the respective maximum and minimum values. Figure 3.2 shows that (during the respective seasonal periods) the deviations from the seasonal mean values are significantly smaller compared to the deviation from the overall mean value to the (overall) maximum and minimum temperature.

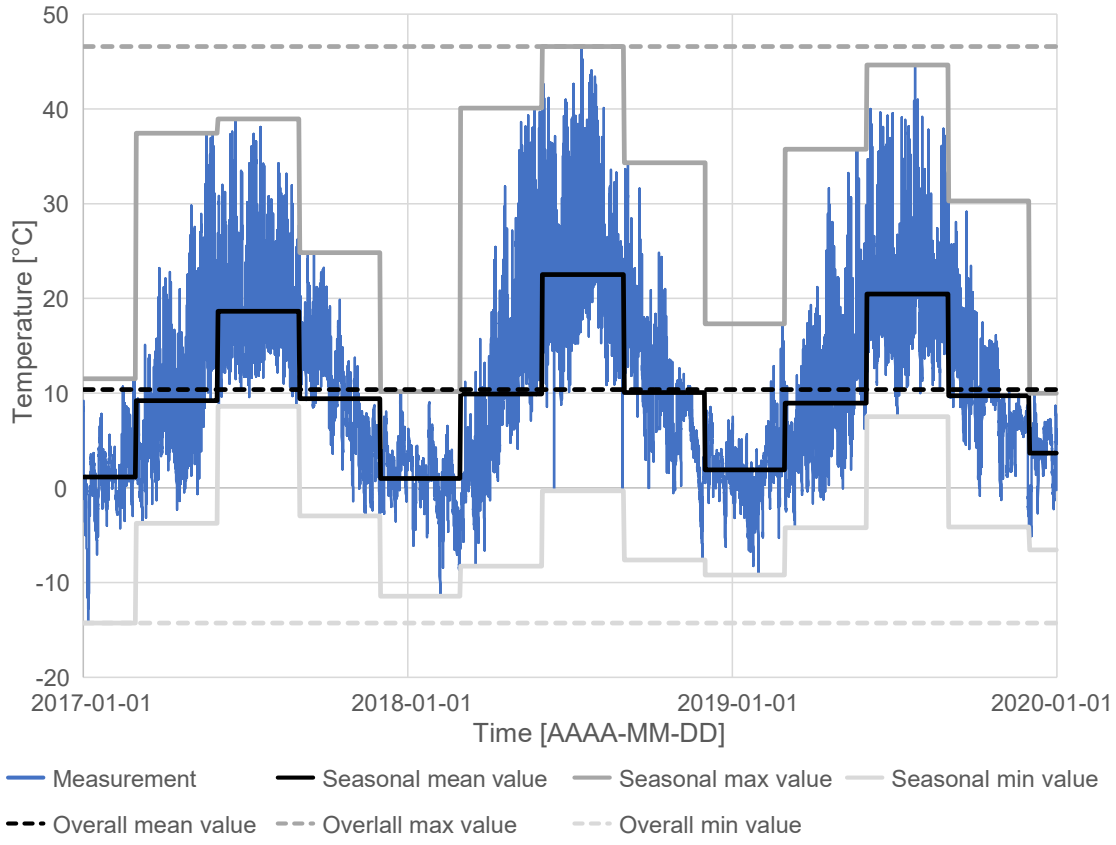


Figure 3.2: Measurement data of the inlet temperature of the drying air flow at a Swedish pulp mill over a period of three years 2017 – 2019.

Figure 3.2 illustrates that operating data may be structured into different independent operating periods, and this thesis proposes to model an individual uncertainty space for each independent operating period. By means of a theoretical example presented in Section 3.1.2, it is demonstrated that, in case of independent operating periods, the actual uncertainty space may be overestimated if the expected uncertainty space is modelled using a single hyperbox.

[2]Note that this division is an illustration and neither a recommendation nor a suggestion to (always) divide operating data according to calendar seasons

3.1.1 Novel approach to consider parameter dependencies for deterministic flexibility analysis

As an alternative to single equation models, this thesis proposes to model the expected uncertainty space by means of upper and lower boundary functions to capture dependencies in the uncertain parameters. Based on the grouping of the uncertain parameters into independent uncertain parameters, θ_{ind}, and dependent uncertain parameters, θ_{dep}, (see Section 2.1.2), a reformulation of the hyperbox model to Eq. (3.1) is proposed.

$$T_{boundary}(\delta) = \begin{cases} \left\{\theta_i \mid \theta_{i,N} - \delta\Delta\theta_i^- \leq \theta_i \leq \theta_{i,N} + \delta\Delta\theta_i^+\right\} \forall\ \theta_i \text{ in } \theta_{ind} \\ \left\{\theta_j \mid f_L(\theta_{ind}) \leq \theta_j \leq f_U(\theta_{ind})\right\} \forall\ \theta_j \text{ in } \theta_{dep} \end{cases} \tag{3.1}$$

A conceptual illustration of the flexibility index based on upper and lower boundary functions in comparison to the hyperbox uncertainty space is shown in Figure 3.3. Figure 3.3 shows that the feasible uncertainty space based on boundary functions is smaller than the expected uncertainty space (a similar result was obtained for the hyperbox model). However, Figure 3.3 shows that the scaling factor of the expected uncertainty space (value of the flexibility index) is closer to 1 for the model based on boundary functions compared to the hyperbox model.

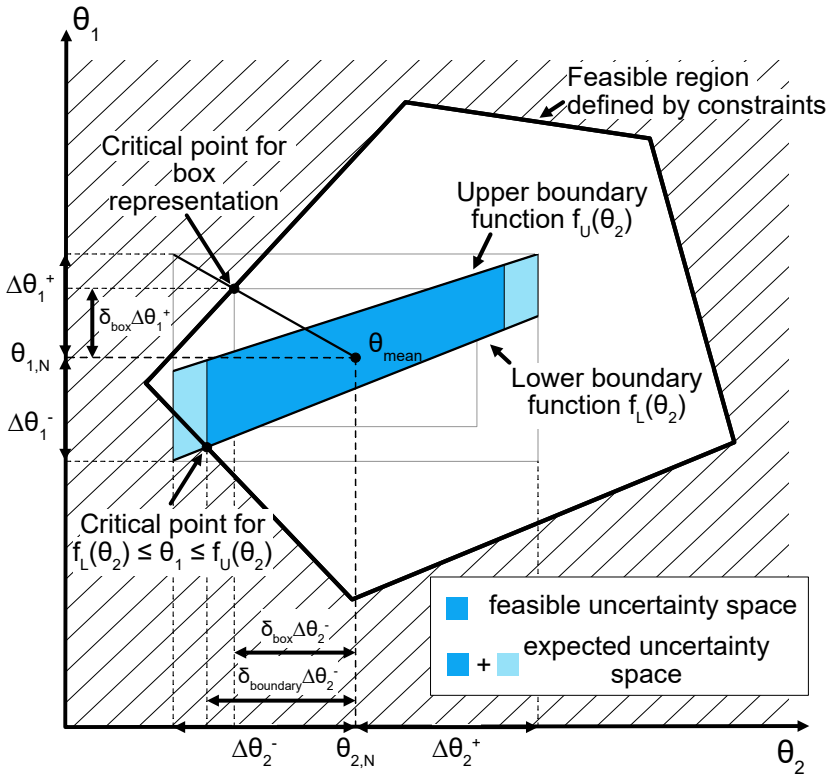


Figure 3.3: Visualization of the flexibility index for the case where two uncertain parameters show a dependency. Reformulated model of the uncertainty space based on boundary functions to capture a dependency between the two uncertain parameters.

In comparison to the hyperbox model, incorporating boundary functions transforms the model of the expected uncertainty space into an irregular square (see Figure 3.3) which can be imagined as hyperpolygon in higher dimensions. This transformation excludes

irrelevant sub-spaces of the hyperbox model, i.e., sub-spaces in which no operation is expected. Note that the geometric shape of the model of the expected uncertainty space shown in Figure 3.3 is defined by the extreme values of the independent parameter, θ_2 and the boundary functions. When assuming that the boundary functions in Figure 3.3 enclose all expected realizations of the uncertain parameters θ_1 and θ_2, Figure 3.3 shows such subsets (not coloured parts of the rectangle). Note that the extreme values (maximum and minimum values of the uncertain parameters, θ_1 and θ_2) are similar for both models of the expected uncertainty space while not all possible combinations are expected to occur in the model based on boundary functions. Consequently, the proposed approach offers additional degrees of freedom when modelling the expected uncertainty space which allow for excluding regions or sub-spaces of the hyperbox model in which no operating points are expected. To summarize this paragraph, boundary functions can allow for a better resemblance between the modelled uncertainty space and the real uncertainty space which helps decreasing the risk for underestimating the feasible variation range when calculating the deterministic flexibility index.

The advantage of boundary functions in comparison to single equation (regression) models is that the dimensionality of the uncertainty set is not reduced. Boundary functions allow uncertainty to be considered even in the dependent parameters since this uncertainty is expressed as the space between the upper and lower boundary function(s) (see coloured irregular square in Figure 3.3). Due to the additional uncertainty in the dependent parameters, the overestimation of the flexibility can be avoided (compare Figures 2.2 and 3.3). More specifically, the boundary functions can be chosen in such a way that all expected realizations of the uncertain parameters are enclosed by the modelled uncertainty space.

In **Paper V**, a mixed-integer formulation for the deterministic flexibility index based on boundary functions is derived. The derived formulation can be solved by means of the active constraint strategy developed by Grossmann and Floudas [41] and allows for the generic application of boundary functions when performing deterministic flexibility analysis. Additionally, **Paper V** presents an algorithm to automate the definition of boundary functions in such a way that the expected uncertainty space is represented as accurately as possible. Such a task can be solved manually but such a manual definition may be burdensome and error-prone, especially for multi-dimensional dependencies. The proposed algorithm is based on the polygon convex hull enabling the effective identification of multiple upper and lower boundary functions. Consequently, the number of degrees of freedom increases even more (in contrast to single upper and lower boundary functions) which means that the expected uncertainty space can be approximated with a high accuracy as long as a tight convex hull representation can be found.

3.1.2 Deterministic flexibility analysis considering independent operating periods

When formulating the flexibility index problem, the uncertainty space should reflect all operating conditions expected during the life-time of the process. Usually, these different realizations of the uncertain parameters are aggregated in a single model/representation, e.g., a hyperbox for independent uncertain parameters or a hyperpolygon when considering parameter dependencies. The modelled uncertainty space is, thus, independent of

time, potentially ignoring that certain realizations of the uncertain parameters may only occur during certain periods of the expected life-time.

However, as shown in Figure 3.2, the seasonal variation of operating conditions over a year represents a situation where certain realizations of the uncertain parameters only occur during certain periods over a year. Another example is differences between day-time and night-time operation. In these two examples (some of) the uncertain parameters show a systematic pattern which is also known as seasonality [83]. Note that seasonality can be observed in available data (retrofit problem) while it can also be anticipated (greenfield problems). Since seasonality occurs with a certain frequency (which may be known or anticipated) it allows the data to be divided into several (independent) intervals as shown in Figure 3.2. The aforementioned phenomenon can also occur if the system of interest is part of a plant which is adjusted to produce different products, e.g., during different production campaigns.

When modelling the expected uncertainty space, such independent operating periods should be considered. In this context, **Papers III & IV** propose to divide the time-horizon of the analysis into several instances representing the different independent operating periods. The expected combinations of the uncertain parameter values can then be allocated to the respective (independent) operating periods defining an individual uncertainty space for each period. Another theoretical example (see Figure 3.4) illustrates how the structuring of data into independent operating periods, e.g., due to seasonality, can impact the result of flexibility analysis, i.e., the flexibility index. In Figure 3.4a, the feasible uncertainty space of the theoretical example is determined using the expected uncertainty space as defined over all operating periods (modelled as a rectangle). On the other hand, in Figure 3.4b, this overall uncertainty space is divided into three independent (or uncoupled) uncertainty spaces resulting from independent operating periods. Note, that the overall extreme values (maximum and minimum values of the uncertain parameters, θ_1 and θ_2) are still expected to occur also for the representation of uncertainty shown in Figure 3.4b, but their occurrences are expected in different periods. It should also be noted that similarly to the boundary function approach for modelling the expected uncertainty space in the presence of parameter dependencies (see Section 3.1.1), the approach shown in Figure 3.4b excludes sub-spaces in the overall uncertainty space in which no operation is expected.

Generally, different approaches can be chosen to model the expected uncertainty spaces of independent operating periods, e.g., hyperbox or hyperpolygon models. For simplicity, in Figure 3.4b each of the three uncertainty spaces is modelled using the hyperbox approach. Note, that operation of the process would be feasible for all expected variations in two of the three periods while for the period around $\theta_{mean,1}$, the feasible uncertainty space is smaller than the expected uncertainty space.

In a case such as that illustrated in Figure 3.4b, the flexibility can be assessed by first determining the flexibility index for each independent operating period and then calculating the overall flexibility index as proposed in **Paper III**. For N independent operating periods, the overall flexibility index is found by:

$$FI = min(FI_n) \; \forall \; n \in 1, 2, ..., N. \tag{3.2}$$

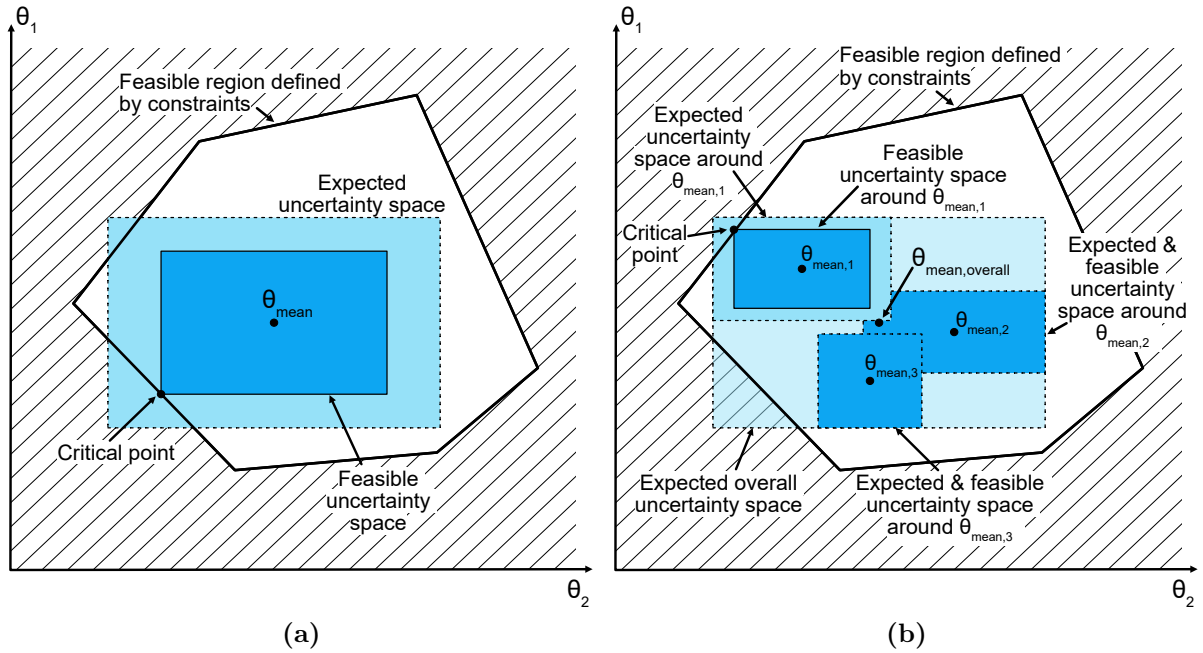


Figure 3.4: Visualization of the flexibility index for two uncertain parameters considering a single overall expected uncertainty space (a) and three independent operating periods observed within the overall expected uncertainty space (b). Note that the (hyper-)rectangle approach was used to model the expected uncertainty space(s). Further note that the exact number and the placement of the different uncertainty spaces in (b) are case-dependent.

To summarize this section, it was established that certain realizations of the uncertain parameters may only occur during certain periods of the time-horizon considered for the flexibility analysis. By allocating realizations of the uncertain parameters to different independent operating periods and modelling those periods as individual uncertainty spaces, some time-dependency can be enabled in the flexibility analysis. However, this time-dependency should not be mistaken for the dynamic flexibility analysis of processes that have an inherently dynamic nature and/or are subject to dynamic uncertainty and feasibility conditions (see e.g., Dimitriadis and Pistikopoulos [84]). Flexibility of dynamic processes, often called resilience, is beyond the scope of this thesis.

3.1.3 Novel developments for identifying critical operating points

For identifying the critical operating point(s) of a structural design proposal, the expected uncertainty space must be modelled, i.e., the search space needs to be specified. Traditionally, the hyperbox approach was utilized (see e.g., Halemane and Grossmann [29] or Pintarič and Kravanja [32]). In the previous sections, situations were outlined where the hyperbox approach can be an inexact representation of the expected uncertainty space. Two novel approaches for modelling the expected uncertainty space were presented, and Figures 3.3 and 3.4 show examples where the combination(s) of uncertain parameter values which are identified to be the critical operating point(s) strongly depend on the chosen model of the expected uncertainty space. Ignoring factors such as parameter dependencies and independent operating periods when identifying critical operating points could imply that, at a later stage, equipment is designed for combinations of uncertain parameter values which are not expected to occur. In particular, conservative and oversimplified modelling of the expected uncertainty space may result in equipment being

oversized. To avoid such problems, this thesis presents a new strategy for identifying critical operating points. The proposed strategy is part of the design under uncertainty framework presented in **Paper VI** and comprises three steps which are outlined hereafter.

Step 1: *Critical operating points in the presence of independent operating periods*
In the first step, one of the independent operating periods is selected for which in the following the critical operating points are identified. Note that the critical points need to be identified for each independent operating period in an iterative procedure. The reason for this can be explained using Figure 3.4b. Figure 3.4b shows three individual uncertainty spaces which represent different independent operating periods, and their consideration allows for excluding sub-spaces of the overall hyperbox uncertainty space in which no operation is expected. This implies that process equipment only needs to allow for steady-state flexible operating within these three uncertainty spaces or independent operating periods to guarantee that the final process design can operate feasibly at all expected operating conditions. Consequently, critical operating points identified for the overall uncertainty space may represent combinations of uncertain parameter values which are not expected to occur, and are thereby not relevant. Since each uncertainty space is independent, it is necessary to identify the critical operating points for each independent operating period/uncertainty space, individually (as further described in Steps 2 and 3). Note that the different sets of critical operating points (identified for the different independent operating periods) must be considered simultaneously when eventually solving the design problem to guarantee steady-state flexible operation for all independent operating periods. Further note that this step can be omitted if the expected operating conditions are captured in (just) one representation of the uncertainty space.

Step 2: *Critical operating points for hyperbox representation*
In the second step, the expected uncertainty space of each independent operating period is modelled as a hyperbox for identifying the critical operating point(s). Thus, at this stage, potential parameter dependencies are ignored. Note that parameter dependencies can be considered by means of Step 3 which is outlined at a later point in this section. This intermediate Step 2 allows for utilizing the existing strategies to identify critical operating points, if these strategies are applicable, e.g., the strategy outlined by Halemane and Grossmann [29] is applicable for design proposals described by convex constraints, only. In this context, **Paper II** reports difficulties when utilizing the algorithms suggested by Pintarič and Kravanja [31, 32]. More specifically, it was identified that the complexity of a structural design proposal can be an essential barrier for the successful application of these algorithms. In this context, different structural design proposals for a HEN were investigated in **Paper II**, and the main difference was the number of HEX units while additional complexities such as stream splitting were not investigated. Since it is expected that such additional complexities are present in relevant case studies (see, e.g., Section 1.1.1), **Paper VI** proposes a new approach to identify the critical operating points within a hyperbox uncertainty space which is outlined hereafter.

The proposed approach is a two-stage iterative algorithm which builds upon the theory of critical operating points described in Section 2.1.3. More specifically, for the example shown in Figure 2.3, the solution of the flexibility index problem, θ^*, considering an initial equipment size (installed or estimated) and the critical operating point, θ_c, (demanding the necessary change in equipment size) are on the same diagonal of the

rectangular expected uncertainty space. Consequently, it would be possible to identify θ_c by projecting θ^* from the feasible uncertainty space (given the initial equipment size) to the expected uncertainty space, i.e., following the aforementioned diagonal or direction from the feasible to the expected uncertainty space. The proposed two-stage algorithm exploits this relationship between θ^* obtained for specific design specifications and θ_c. More precisely, in the first stage, preliminary design specifications are identified solving a simplified design under uncertainty problem considering a list of candidates for the critical operating points[3]. In the second stage, the flexibility index problem is solved considering the identified design specifications. If the obtained design specifications do not allow for steady-state flexible operation within the (entire) expected uncertainty space (flexibility index < 1), it can be concluded that the list of candidates is incomplete, i.e., not all critical operating points have been identified. If the list of candidates is incomplete, θ^* (solution of the flexibility index problem) is projected to the expected uncertainty space to update/extend the list of candidates and the algorithm returns to the first stage.

Multiple critical operating points can be expected in multi-dimensional examples, i.e., examples with more than two uncertain parameters and/or more than one design parameter/constraint. To illustrate such a case, a modification of the aforementioned theoretical example is shown in Figure 3.5. In the modified theoretical example, two design constraints are present and Figure 3.5 shows that the design parameters of both design constraints need to be manipulated to enable steady-state flexible operation. When following the proposed algorithm, after the first iteration, critical point A would have been identified which would ensure that design constraint 1 is shifted. However, the feasible uncertainty space is still smaller than the expected uncertainty space since it is limited also by the initial equipment size implied by design constraint 2. Therefore, a second iteration is necessary which results in the identification of critical point B. Note that the number of design constraints does not necessarily correspond to the number of critical operating points. More detailed information on the mathematical formulations to project θ^* from the feasible to the expected uncertainty space is provided in **Paper VI**.

Step 3: *Critical operating points for boundary function models*
In the third step, the proposed strategy suggests to update the critical operating points identified for the hyperbox representation considering previously identified parameter dependencies. The proposed updating scheme is an iterative algorithm which aims to determine if the previously identified critical operating points are within the hyperpolygon model of the expected uncertainty space (i.e., the model of the expected uncertainty space defined by the extreme values of the independent parameters and the boundary functions). If the previously identified critical operating points represent combinations of uncertain parameter values which are outside of the hyperpolygon model, the critical values of (at least) some uncertain parameters must be adjusted/updated. For this adjustment, it is assumed that the actual critical operating points represent uncertain parameter values on the boundary of the hyperpolygon. Note that in multi-dimensional cases, not all independent uncertain parameters, θ_{ind}, are expected to be utilized when defining boundary function models. Thus, it can be assumed that updating/adjusting of the critical parameter values is only required for those uncertain parameters (dependent and independent) which show correlating trends.

[3]The proposed simplified design under uncertainty problem searches for the set of optimal design parameters with respect to the investment cost while operating cost are not included in the problem.

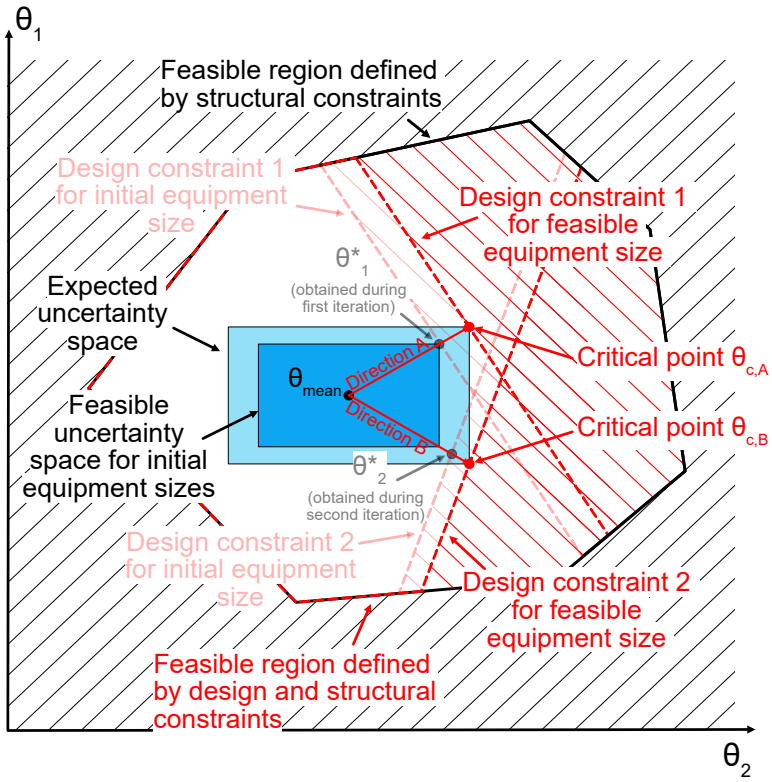
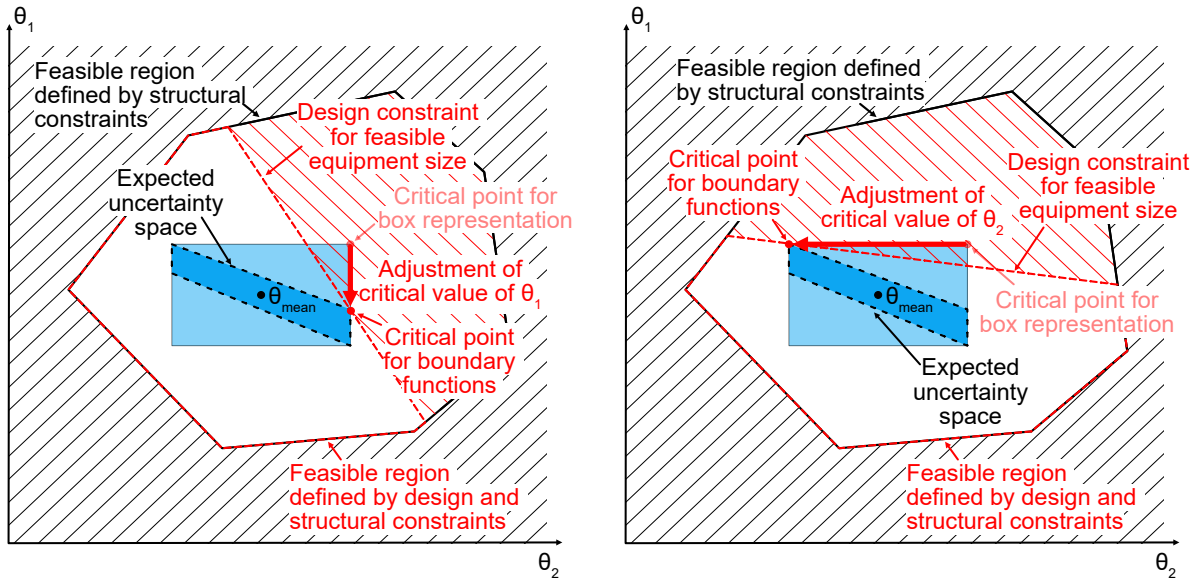


Figure 3.5: A theoretical example with two design constraints to illustrate the identification of two critical operating points using the proposed approach. During the first iteration, θ_1^* is identified (solution of flexibility index problem) which can be projected using direction A to identify critical point A, $\theta_{c,A}$. A second iteration is necessary to first identify θ_2^*, which can be projected using direction B to identify critical point B, $\theta_{c,B}$.

The necessary adjustment of the critical parameter values of some uncertain parameters can be visualized using the theoretical example shown in Figure 2.3. When assuming a dependency (with a negative correlation coefficient) between the uncertain parameters θ_1 and θ_2, Figure 3.6a illustrates how the critical operating point for the box representation and the critical operating point for boundary functions relate to each other. More precisely, Figure 3.6a shows that the critical parameter value of θ_1 needs to be adjusted in order to identify the critical operating point for the uncertainty space defined by boundary functions. On the other hand, Figure 3.6b shows a slight modification of the same theoretical example where the slope of the design constraint is different. Due to this modification, the critical parameter value of θ_2 (and not θ_1) needs to be adjusted. To identify which critical parameter values need to be adjusted, the proposed updating algorithm starts with the assumption that the critical parameter values of the dependent uncertain parameters, θ_{dep}, need to be adjusted (scenario shown in Figure 3.6a). This assumption is tested using the (aforementioned) simplified design under uncertainty problem in combination with the flexibility index problem[4]. The solution of the flexibility index problem indicates if the made assumption was correct, or provides insights regarding which corrective measures (adjustment of critical values of some independent parameters) should be taken to eventually identify the updated critical operating points in an iterative procedure.

[4]Note that the flexibility index problem must be defined for the expected uncertainty space modelled using boundary functions, i.e., the MI(N)LP mentioned in Section 3.1.1.



(a) Example 1: Critical parameter value of dependent parameter θ_1 needs to be adjusted.

(b) Example 2: Critical parameter value of independent parameter θ_2 needs to be adjusted.

Figure 3.6: Examples to illustrate how the critical parameter values for the box representation can be updated to achieve the critical parameter values for the boundary functions representation.

To summarize this section, a strategy comprising three steps has been outlined to identify critical operating points when independent operating periods and/or parameter dependencies are expected. Additionally, a new approach was presented to identify the critical operating points of a hyperbox uncertainty space. The proposed strategy including the new approach were successfully applied to identify the critical operating points of the industrial case study presented in Section 1.1.1 and the results are presented in **Paper VI**. Note that a rigorous comparison of the proposed approach to identify the critical operating points of a hyperbox uncertainty space with the available approaches in the literature remains for future work.

3.2 Flexibility analysis considering uncertain long-term development

In Section 1.1.2, a situation is described where the nominal operating conditions of an industrial heat recovery system are expected to change due to a planned switch in feedstocks. Further examples for such planned long-term development are a change of operational parameters required to comply with new emission legislation and/or a change in the production rate. All these events may have a lasting effect on the operation of the process in question and a possible consequence is that the nominal operating conditions change temporarily or even permanently. Section 1.1.2 further illustrates that operational disturbances around the nominal point (current and future) may also be expected. Such operational disturbances are comparable to the traditional interpretation of uncertainty which, e.g., Swaney and Grossmann [2] aimed to analyse by means of the flexibility index.

The situation resulting from planned long-term development is comparable to the pres-

ence of independent operating periods (compare Section 3.1.2), meaning that such a future event divides the time-horizon of the flexibility analysis into two independent operating periods. If the expected uncertainty space before and after the change can be quantified with high certainty, the flexibility of the respective process can be evaluated as described in Section 3.1.2, i.e., by calculating the flexibility index for each individual uncertainty space to eventually determine the overall flexibility index following Eq. (3.2). In **Paper III** such an approach was utilized to analyse if currently installed process equipment of the industrial case study presented in Section 1.1.2 is able to operate at the expected conditions after the implementation of biomass co-processing.

On the other hand, such (planned) long-term development usually comes with high levels of uncertainty, and the impact on both the nominal operating conditions as well as the expected uncertainty space before and after the change must be estimated. Consequently, (planned) long-term development introduces a new dimension of uncertainty which interferes with the (traditional) operational uncertainty in the short-term[5]. To illustrate this, Figure 3.7 shows a situation in which the nominal value of one of two uncertain parameters of a theoretical example is expected to change. The new nominal value after the implementation of the long-term development is uncertain but it is expected that the change has an upper bound - indicated in Figure 3.7 as the expected maximum change. Furthermore, Figure 3.7 shows that the expected uncertainty space for the short-term operational disturbances around the current nominal point is within the feasible region, i.e., steady-state flexible operation is possible at the current nominal operating point. On the other hand, the change of the nominal value can result in a situation where the process cannot handle the expected short-term operational disturbances if the (absolute) deviations, $\Delta\theta^-$ and $\Delta\theta^+$, remain constant even when the nominal operating point changes. In this context, Figure 3.7 shows the maximum feasible change of the nominal value such that all expected short-term operational disturbances are (exactly) feasible.

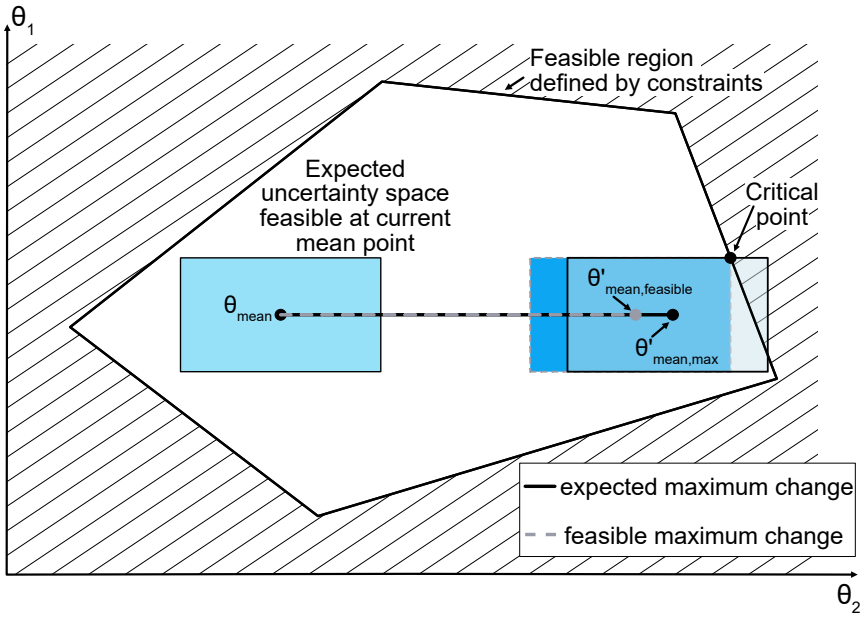


Figure 3.7: A theoretical example illustrating the situation when (planned) long-term development interferes with short-term operational disturbances.

[5]Note that in **Paper IV** this phenomenon is denoted as overlaying uncertainty sources.

As shown in Figure 3.7, when uncertainty in long-term development interferes with operational short-term uncertainty, this can lead to situations where feasible operation is no longer possible. The (traditional) flexibility index as defined by Swaney and Grossmann [2], can be utilized to identify if steady-state flexible operation is possible at the current nominal operating point. However, it is not well-defined how to apply this traditional formulation of the flexibility index if also aiming to consider uncertainty due to long-term development. In Section 3.2.1, strategies to consider uncertain long-term development in the original formulation of the flexibility index are investigated (the original formulation is given in Problem (A.1) in the Appendix). However, these strategies present several shortcomings which include wrong conclusions as well as ineffective iterative procedures. To overcome these shortcomings, this thesis proposes novel reformulations which are presented in Section 3.2.2. For readability, hereafter, the different uncertainty sources (short-term, long-term) are classified as shown in Table 3.1.

Table 3.1: Different classes of uncertainty based on the source or origin of the uncertainty.

Conventional operational disturbances (included in the original flexibility index formulation)	Uncertainty class A
Uncertainty due to (planned) long-term development (i.e., nominal operating point varies or changes)	Uncertainty class B

3.2.1 Original flexibility index formulation and overlaying uncertainty sources

It may be intuitive to include (additional) uncertainty sources (of any kind and nature) in a similar fashion to that proposed for uncertainty *class A* by Swaney and Grossmann [2] (see Problem (A.1)). Thus, the uncertain parameters are grouped with respect to the uncertainty classes given in Table 3.1 (*class A* and *class B*: $\theta_{class\ A}$ and $\theta_{class\ B}$) and the scalable expected uncertainty space can be formulated using the hyperbox approach or the approach based on boundary functions. Note that a physical uncertain parameter (e.g., an uncertain temperature) may be present in both sets, $\theta_{class\ A}$ and $\theta_{class\ B}$, while the expected variation range or change, $\Delta\theta^+$ and $\Delta\theta^-$, differs for each class. In the example shown in Figure 3.7, θ_1 belongs to $\theta_{class\ A}$ while θ_2 is present in $\theta_{class\ A}$ and $\theta_{class\ B}$. Furthermore, the expected variation range for *class A* is given by the light blue rectangle around the current nominal point (for both, θ_1 and θ_2, short-term operational disturbances are expected) while the expected variation or change for *class B* is the maximum expected change of the nominal value of θ_2[6]. To exemplify this strategy, Eq. (3.3) shows the scalable uncertainty space based on the hyperbox approach considering $\theta_{class\ A}$ and $\theta_{class\ B}$.

$$T_{box}(\delta) = \begin{cases} \left\{\theta_i \mid \theta_{i,N} - \delta\Delta\theta^-_{i,A} \leq \theta_i \leq \theta_{i,N} + \delta\Delta\theta^+_{i,A}\right\} \ \forall\ \theta_i \text{ in } \theta_{class\ A} \\ \left\{\theta_k \mid \theta_{k,N} - \delta\Delta\theta^-_{k,B} \leq \theta_k \leq \theta_{k,N} + \delta\Delta\theta^+_{k,B}\right\} \ \forall\ \theta_k \text{ in } \theta_{class\ B} \end{cases} \tag{3.3}$$

This strategy yields the single scalar, δ, which expresses the maximum feasible variation/change for each uncertain parameter in $\theta_{class\ A}$ and $\theta_{class\ B}$. However, with such a strategy it is not possible to gain information about the feasibility of operational short-term disturbances (*class A*) after the nominal operating point has changed from

[6] When assuming dummy values for the example shown in Figure 3.7, the following could be true: $\Delta\theta^+_{1,A} = \Delta\theta^-_{1,A} = 5$, $\Delta\theta^+_{2,A} = \Delta\theta^-_{2,A} = 10$, $\Delta\theta^+_{2,B} = 40$ and $\Delta\theta^-_{2,B} = 0$

its original value (*class B*). For example, in the example shown in Figure 3.7, the value of the flexibility index achieved with this strategy would be > 1 since both the short-term disturbances around the current nominal point as well as the change from the current nominal point to the new nominal point (without considering the short-term variations) are within the feasible region. The reason for this is that this strategy does not include the possibility to consider the (rectangular) uncertainty space (in Figure 3.7) representing the short-term operational disturbances at the new nominal operating point.

The conclusions drawn from the previous paragraph illustrate that the traditional definition of deterministic flexibility analysis to evaluate the ability to cope with (operational short-term) uncertainty is insufficient if uncertainty due to long-term development is expected. More specifically, a strategy is needed to identify the maximum feasible change of the nominal values such that all expected short-term operational disturbances are (exactly) feasible. To address this challenge, the flexibility index problem needs to be formulated for different discrete values within the expected uncertainty space of *class B*, e.g., discrete points on the line illustrating the expected (maximum) change of the nominal value of θ_2 in Figure 3.7, and the resulting formulations can then be solved in an iterative scheme. If sufficiently many points are tested, such a strategy should eventually reveal an acceptable approximation of the maximum feasible change of the uncertain parameters belonging to *class B* while ensuring that the expected variation of the uncertain parameters belonging to *class A* is exactly feasible. However, this iterative scheme can be impractical and time-consuming; thus, in the next section, this thesis proposes reformulations of the original deterministic flexibility index problem.

3.2.2 Proposed reformulations of the deterministic flexibility index problem

In line with Figure 3.7, the proposed reformulations aim to find the maximum feasible variation/change for uncertain parameters of *class B* which guarantees that a pre-defined flexibility target of the uncertain parameters of *class A* (operational short-term disturbances) is feasible, e.g., the expected variations of the uncertain parameters of *class A* are exactly feasible. Thus, an explicit distinction is made between the two uncertainty classes and the proposed reformulation of the flexibility index problem is given in Problem (3.4).

$$
\begin{aligned}
FI = & \max \delta_B \\
& s.t. \max_{\theta \in T(\delta_A, \delta_B)} \min_{z} \max_{j \in J} f_j(d, z, \theta_{class\ A}, \theta_{class\ B}) \leq 0 \\
& T_{box}(\delta_A, \delta_B) = \left\{ \begin{array}{l} \left\{ \theta_i \mid \theta_{i,N} - \delta_A \Delta\theta_i^- \leq \theta_i \leq \theta_{i,N} + \delta_A \Delta\theta_i^+ \right\} \forall\ \theta_i \in \theta_{class\ A} \\ \left\{ \theta_k \mid \zeta_k - \delta_A \Delta\theta_{k,A}^- \leq \theta_k \leq \zeta_k + \delta_A \Delta\theta_{k,A}^+ \right\} \\ \left\{ \theta_k \mid \theta_{k,N} - \delta_B \Delta\theta_{k,B}^- \leq \zeta_k \leq \theta_{k,N} + \delta_B \Delta\theta_{k,B}^+ \right\} \end{array} \right\} \forall\ \theta_k \in \theta_{class\ B} \\
& \delta_A = specific\ target \\
& \delta_B \geq 0
\end{aligned}
\tag{3.4}
$$

Problem (3.4) includes several reformulations compared to the original flexibility index problem (Problem (A.1)). When distinguishing between uncertain parameters of *class A*

and of *class B*, the maximum feasible variation/change for each class is respected individually by defining a scalar, δ, for each class (δ_A, δ_B). As aforementioned, this thesis proposes searching for the maximum feasible variation/change for uncertain parameters of *class B* which allows for a pre-defined flexibility target of the uncertain parameters of *class A*. Consequently, the constraint $\delta_A = specific\ target$ is included in Problem (3.4) while searching for the maximum value of δ_B is formulated as the objective function. Furthermore, the uncertainty of the nominal value(s) for the uncertain parameters of *class B* is respected by defining the variables $\zeta_k \ \forall \ \theta_k \in \theta_{class\ B}$. Consequently, the first and the second line in $T_{box}(\delta_A, \delta_B)$ guarantee that expected short-term disturbances remain feasible when the nominal operating point varies or changes. The maximum feasible variation/change of the nominal operating point is found by including the third line in $T_{box}(\delta_A, \delta_B)$.

Problem (3.4) was formulated using the traditionally established hyperbox approach. If parameter dependencies are expected, these can be modelled using boundary functions or single equation models. Note that parameter dependencies can influence the expected uncertainty space of *class A* as well as *class B*. For example, the expected change of the nominal values of two (or more) uncertain parameters due to uncertain long-term development could be dependent in some way. In this context, **Paper IV** presents a theoretical example where the expected change of the nominal values of two uncertain parameters is correlated, and a single equation model is used to express this dependency.

For solving Problem (3.4), the active constraint strategy can be utilized which was proposed for solving the original flexibility index problem (Problem (A.1)) by Grossmann and Floudas [41]. This requires that an upper bound for the operational flexibility target value ($\delta_A = specific\ target$) is pre-defined. This upper bound can be obtained in a first step by considering only the uncertain parameters of *class A* and thus formulating and solving the original flexibility index problem (Problem (A.1)). In a second step, the uncertain parameters of *class B* are added, and Problem (3.4) can be formulated and solved (for a flexibility target value for δ_A which is lower or equal than the identified upper bound).

If the expected variation range of operational uncertainty (*class A*) is itself expected to change (when the nominal operating point varies/changes), $\Delta\theta^-_{k,A}$ and $\Delta\theta^+_{k,A}$ can be defined as functions depending on the nominal operating point ζ_k. A probable scenario could be that the expected variation range of short-term disturbances is expected to be a percentual share of the nominal operating point, such as $\pm 5\%$. By means of a second theoretical example, the flexibility analysis of such a scenario where the expected variation range of short-term disturbances is expected to be a percentual share of the nominal operating point is presented in **Paper IV**. Finally, the industrial case study presented in Section 1.1.2 is also investigated in **Paper IV** and selected results are presented in Section 5.2.

4

Methodological Development

4.1 Automated heat exchanger network modelling

As illustrated by the motivational examples presented in Section 1.1, industrial HENs can be of very different size and complexity. One issue with (industrial) HENs is that the mathematical modelling can be time-consuming and also error-prone, especially when conducted manually. On the other hand, as outlined in Section 2.3, the variety of physical components commonly present in a HEN is limited while the arrangement (i.e., the existence and the position) of these physical components can differ substantially. To model a HEN, this arrangement of the physical components in a HEN needs to be described[1]. In this thesis, a generic methodology is proposed to describe the topology of any given HEN in a straightforward way. The proposed methodology includes routines to describe stream splits, re-circulation of streams and closed circulation loops.

In a first step, each stream of the HEN is numbered. To handle stream splits, re-circulation of streams, closed circulation loops, etc., streams are numbered with respect to their CP-values and identities. This means that a stream in the modelling sense needs to have a constant CP-value and fixed identity. For example, a stream split where one stream is split into two streams is counted as three individual streams. If the two split streams are mixed back together, a fourth stream is counted. For re-circulation of streams, closed circulation loops, or other components which cause changes in the identities of streams, switches are used which define the locations where the identity of a stream changes (see Section 2.3.3). Consequently, with each switch a new stream is counted. To illustrate the proposed methodology, a theoretical HEN example was developed which requires the application of all developed modelling routines. The example HEN comprises stream splits as well as the re-circulation of a stream. Figure 4.1 shows the classical grid representation of this example network. A switch model is used to describe the stream re-circulation. The streams are numbered according to the proposed numbering procedure. Furthermore, all (unknown) network temperatures with respect to Eq. (2.5) are marked.

Besides stream numbering, a generalized way to refer to the location of a specific process-to-process HEX (P2P) or utility HEX (utility) is needed. Basically, this includes the (process) streams which are exchanging heat by means of the HEX as well as the location of the specific HEX on those streams. For this reason, a system to specify the position of the HEXs was developed which is based on the stream numbering. For each HEX, the hot and cold stream as well as the position of the HEX on these streams must be known. For utility HEXs, a target temperature is specified in order to calculate the duty of each

[1] The arrangement of the physical components in a HEN is also known as the structural layout or the network topology.

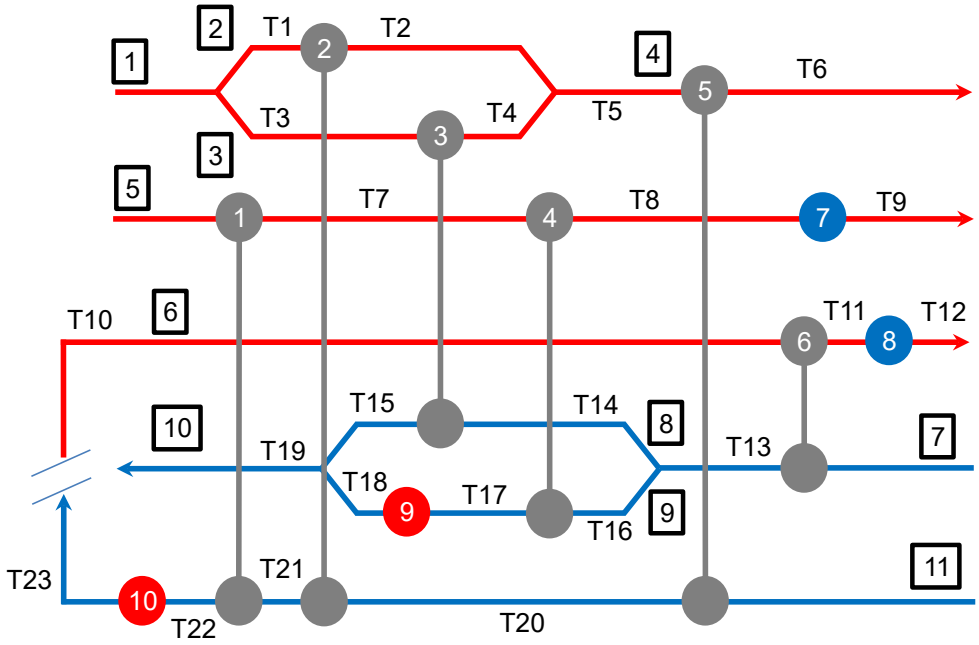


Figure 4.1: An example heat exchanger network illustrating the stream numbering and (unknown) network temperatures.

utility HEX (see Section 2.3.1). For the HEN depicted in Figure 4.1, the location matrix for the HEXs is given in Table 4.1.

Table 4.1: Exchanger location matrix of the example HEN for defining the position of all process-to-process (P2P) and utility exchangers and the target temperatures of streams with utility exchangers.

		Hot streams		Cold streams			Only for type "utility"
Exchanger (HEX)	UA	Stream number	HEX number on stream	Stream number	HEX number on stream	Type (P2P or utility)	Target temperature (after exchanger)
1	UA1	5	1	11	2	P2P	
2	UA2	2	1	11	3	P2P	
3	UA3	3	1	8	1	P2P	
4	UA4	5	2	9	2	P2P	
5	UA5	4	1	11	4	P2P	
6	UA6	6	1	7	1	P2P	
7		5	3			utility	Ttarget,5
8		6	2			utility	Ttarget,6
9				9	1	utility	Ttarget,9
10				11	1	utility	Ttarget,11

In Table 4.1 and Figure 4.1, the position of the HEXs is determined by reading from left to right which in this case is the grid flow direction of the hot streams. In general, different ways are possible and can be implemented in the modelling strategy. A similar positioning

system is used to describe the location of stream splits and mixing of streams. It is also based on the introduced stream numbering system. Table 4.2 presents the location matrix for the stream splits and mixes for the example HEN shown in Figure 4.1.

Table 4.2: Split/mix location matrix describing the position of all stream splits and mixes of the example network.

	Split streams			Mix streams		
Index	In	Out1	Out2	In1	In2	Out
1	1	2	3	2	3	4
2	7	8	9	8	9	10

To describe the location of switches, the ingoing and the outgoing streams must be specified, and a similar location matrix can be derived. Table 4.3 shows the location matrix for switches of the example HEN shown in Figure 4.1.

Table 4.3: Switch location matrix describing the position of all switches of the example network.

	Switches	
Index	In	Out
1	11	6

Based on the exchanger location matrix, the split/mix location matrix and the switch location matrix, the vector of the (unknown) network temperatures, T, can be sorted in a data structure called temperature matrix. In this data structure, the (unknown) network temperatures are allocated to the process streams. The allocation process is automated in the proposed modelling strategy. The basis for this automatization is Eq. (2.5). More specifically, one element of T is allocated to a stream if the stream is connected to a HEX (P2P or utility). Furthermore, one element of T is allocated to each stream resulting from a split, mix or switch. This way, all (unknown) network temperatures can be allocated to the different streams. Therefore, one dimension in the temperature matrix represents the streams present in the HEN.

The temperature matrix for the example HEN is shown in Table 4.4. As mentioned previously, the rows of the temperature matrix represent the different process streams (row number equivalent to stream number in Figure 4.1), while the number of columns represent the number of (unknown) network temperatures of the corresponding process stream. As the number of (unknown) network temperatures may be different for the individual process streams, a data structure which accounts for this must be used (e.g., “Cell Array” in MATLAB or “List of Lists” in PYTHON).

The temperature matrix is the basis for the automated HEN modelling strategy as it enables the allocation of the elements of T (unknown network temperatures, see Figure 4.1) to specific HEXs (P2P and utility), stream splits, mixing points and switches. The allocation of the (unknown) network temperatures is essential to be able to derive the set of energy and mass balances presented in Section 2.3 automatically with the correct elements of T. This is illustrated by means of HEX 1 of the example network shown in Figure 4.1:

Based on the exchanger location matrix (Table 4.1), the two process streams which are connected by HEX 1 are hot stream 5 and cold stream 11. Additionally, the position of HEX 1 on these two streams is specified in the exchanger location matrix. For hot stream 5, HEX 1 is the first exchanger (counting from left to right) on this stream. Thus, the hot stream inlet temperature of HEX 1 is the inlet temperature of stream 5, $T_{in,5}$, while the hot stream outlet temperature of HEX 1 is $T7$ (elements 1 and 2 for stream 5 in the temperature matrix, Table 4.4). For cold stream 11, HEX 1 is the second exchanger on this stream. Thus, the cold stream outlet temperature of HEX 1 is $T22$ while the cold stream inlet temperature of HEX 1 is $T21$ (elements 2 and 3 for stream 11 in the temperature matrix, Table 4.4). Consequently, the unique elements of T which are necessary to calculate the logarithmic mean temperature difference for HEX 1 can be identified.

As the example for HEX 1 shows, by means of the exchanger location matrix and the temperature matrix, specific elements of T are allocated to each HEX. Similar allocations can be achieved by means of the split/mix or the switch location matrix to allocate (unknown) network temperatures to the corresponding split, mix or switch. Consequently, the set of equations to describe the heat and mass balances for an arbitrary HEN (see Section 2.3) can be derived automatically.

Table 4.4: Temperature matrix to allocate (unknown) network temperatures of the example network to specific exchangers (process-to-process and utility), stream splits, mixing points and switches.

Stream	(Unknown) network temperatures				
1	$T_{in,1}$				
2	$T1$	$T2$			
3	$T3$	$T4$			
4	$T5$	$T6$			
5	$T_{in,5}$	$T7$	$T8$	$T9$	
6	$T10$	$T11$	$T12$		
7	$T13$	$T_{in,7}$			
8	$T15$	$T14$			
9	$T18$	$T17$	$T16$		
10	$T19$				
11	$T23$	$T22$	$T21$	$T20$	$T_{in,11}$

4.2 A framework for guiding design under uncertainty problems

As outlined in Section 1.2.2, a major difficulty in design under uncertainty is to guarantee that the final (process) design guarantees steady-state flexible operation within the expected uncertainty space. Therefore, the majority of the available approaches to solve a

design under uncertainty problem guarantee the lowest investment as well as feasible and cost optimal operation for pre-defined discrete operating points while feasible operation at the remaining expected operating points is not addressed. A possible explanation for this shortcoming is that the rigorous assurance of steady-state flexible operation is computationally burdensome which limits the application in large and complex case studies. To avoid such computational burdens when dealing with design under uncertainty, this thesis proposes to divide the design process into two steps which are decoupled from each other.

In the first step, different structural design proposals of the process in question are developed. When following the proposed approach, it is worth mentioning that steady-state flexible operation does not need to be addressed during the definition of structural design proposals. Thus, it is possible to utilize proven design synthesis methodologies to generate the basic structure of the design. Note that methodologies based on graphical-insights or other design synthesis approaches considering only a single or a low number of steady-state operating point(s) are also applicable. Furthermore, this setup allows incorporation of non-quantifiable knowledge such as experience-based heuristics in the design proposals, which can be advantageous for complex industrial cases. The different structural design proposals can be collected in a superstructure which comprises all structural alternatives. For the second step, this thesis proposes a step-wise framework to evaluate the structural design proposals collected in the superstructure to identify:

- if any of the structural design proposals is not structurally feasible,
- the most cost-efficient overdesign of equipment size required to guarantee steady-state flexible operation (for each structurally feasible design proposal), and
- a basis to compare the different structural design proposals with respect to a given objective.

The framework proposed in this work can be seen as a tool to analyse the results of early design stage screening processes for chemical processes subject to uncertain parameters. Since the (structural) design proposals are defined prior to the application of the framework, the framework can be utilized to guide both greenfield and retrofit design projects. The different steps of the framework are presented in Figure 4.2. The framework was initially presented in **Paper II** as a tool to guide HEN retrofit projects. An extended version of the framework which is applicable to both greenfield and retrofit design problems of chemical processes was later presented in **Paper VI**. In the following sections, a comprehensive overview of the individual steps of the framework is provided.

4.2.1 Identification of uncertain parameters

In this work, uncertainty in the input parameters of a given process or system is considered. Since these parameters are determined outside the system's boundary, they cannot be affected by measures within the system. However, recourse actions can be taken to ensure that the operational target of the system of interest is met. These recourse actions are limited by the structure of a design and/or by the size/capacity of the equipment. Note that uncertain parameters which are not measurable cannot be addressed by means of the presented framework, since recourse action is not possible in such cases. Once the uncertain parameters are identified, the expected uncertainty space needs to be modelled,

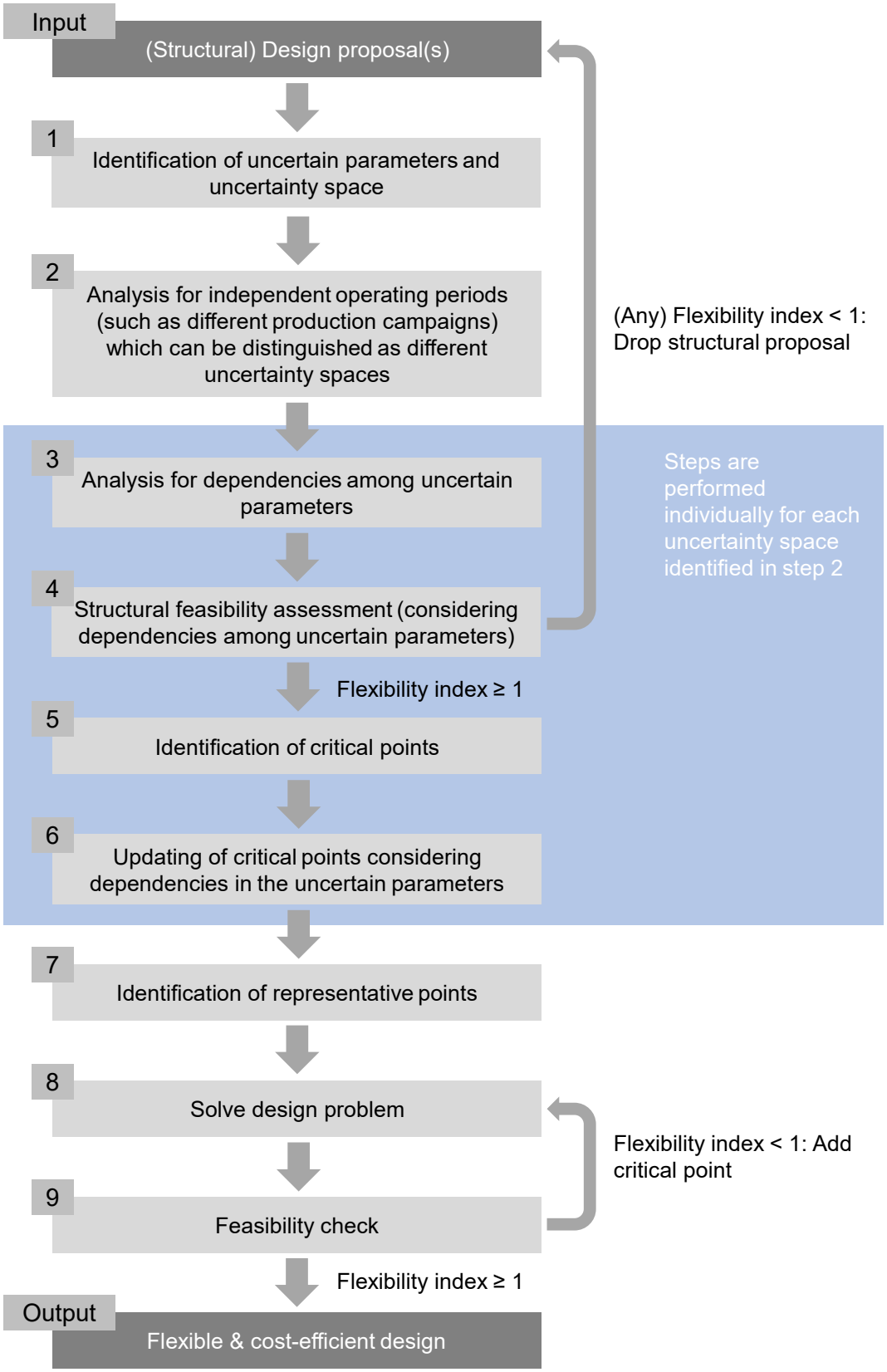


Figure 4.2: Proposed framework for addressing design under uncertainty problems when designing or redesigning chemical processes/plants.

i.e., the space in which all possible combinations of operating values for these parameters are expected. As shown in Section 3.1, dependencies in the uncertain parameters as well as independent operating periods should be considered to ensure a good resemblance between the modelled and the actual uncertainty space. Therefore, additional data analysis is necessary to model the uncertainty space with good accuracy, i.e., checking for such dependencies and independent operating periods.

4.2.2 Analysis for independent operating periods

In this step the existence of potential independent operating periods is identified. These independent operating periods can be identified manually by analysing if (potential) causes for independent operating periods are present such as:

- seasonality in the operating data (e.g., differences between day-time and night-time operation or seasonal variation),
- different production campaigns,
- uncertain long-term development resulting in changes in the nominal operating conditions (overlaying uncertainty sources).

Additionally, automated approaches such as data clustering may be utilized to identify if the operating points can be divided into separate subsets. Note that automated approaches rely to a larger extent on the availability of (good-quality) operating data, i.e., measurements of the uncertain parameters which is often only the case in retrofit projects. On the other hand, anticipating different operating cases such as those described in the list above is commonly done for greenfield design projects where no historic operating data is available. In such cases, the expected disturbance ranges of the different (presumably) uncertain parameters must also be anticipated.

Since the identified operating periods or data clusters are (ideally) independent of each other, a separate uncertainty space should be defined for each period/cluster, e.g., as shown in Figure 3.4b. The following steps of the framework (Steps 3-6) are then performed for each identified uncertainty space.

4.2.3 Analysis for dependencies among uncertain parameters

For each identified uncertainty space (overall, or for each individual independent operating period), the shape of the uncertainty space needs to be approximated. If the uncertain parameters are non-correlated, the shape can be described by the expected lower and upper bound values yielding several hyperbox spaces in the overall space of uncertainty[2]. If dependencies (correlations) are observed or expected, the real uncertainty space would be smaller than the space approximated by the hyperbox representation, since regions will exist within the hyperbox in which no operating points are expected. By considering dependencies when modelling the uncertainty (sub-)spaces, such 'empty' regions may be identified and removed, achieving a more accurate approximation of the actual uncertainty (sub-)spaces. To identify dependencies between different uncertain parameters,

[2]As previously mentioned, this step can be performed relying on historic operating data or by anticipating the uncertainty space.

Paper VI proposes two alternative approaches.

Once there is an indication that a dependency exists between two or more uncertain parameters, it has to be determined if the dependency should be considered when modelling the uncertainty space[3]. Such a decision can be made based on different criteria, e.g., the strength of the dependency. Next, for each identified pair of uncertain parameters (which show a dependency/correlation), there is a need to determine which of the two parameters is to be classified as dependent and independent parameter[4]. This decision can sometimes be enhanced by background knowledge about the process to identify the origin of the dependencies. Finally, the identified dependencies which should be considered when modelling the expected uncertainty space can be formulated in the following general form:

$$\theta_j = f(\theta_{i,1}, \theta_{i,2}, ...),\ \theta_i \in \theta_{ind},\ \theta_j \in \theta_{dep}. \tag{4.1}$$

4.2.4 Structural flexibility analysis

The next step analyses whether the structural layout of each design proposal allows for steady-state flexible operation. As previously outlined, this can be done by calculating the flexibility index discarding all design constraints and considering only structural constraints. If design proposals can be identified that are structurally infeasible for at least some operating points within the expected uncertainty space (i.e., structural flexibility index smaller than 1), these design proposals are discarded.

If independent operating periods and/or dependencies in the uncertain parameters are identified in Steps 2 and 3, it is necessary to adjust the calculation procedure of the flexibility index, compared to the conventional procedure based on a single hyperbox uncertainty space. In order to assess the structural feasibility in the presence of independent operating periods, the identified uncertainty spaces of the different periods need to be analysed individually (as discussed in Section 3.1.2) to eventually determine the overall flexibility index following Eq. (3.2). In the special case that uncertain long-term development is expected to lead to permanent or temporary changes in the nominal operating conditions there is an inevitable need for assumptions regarding the future change of operation (see also Section 3.2). If the change of the nominal operating point and short-term disturbances around the new nominal operating point can be predicted with high certainty, the situation is similar to the other cases of independent operating periods, and the overall flexibility index can be calculated according to Eq. (3.2). However, if there is a high level of uncertainty with respect to how the nominal operating point and/or the short-term disturbances are affected by the uncertain long-term development, it may be desirable to identify the maximum change of the nominal operating point which is structurally feasible (compare Section 3.2 and Figure 3.7).

If parameter dependencies (correlations between uncertain parameters) have been identified in the previous step of the framework (see Section 4.2.3), they should be considered

[3]Note that considering a parameter dependency requires (more) advanced modelling which means that (simply) considering all identified parameter dependencies may increase the problem complexity of the subsequent flexibility index problem and the identification of critical operating points.

[4]Note that the dependent parameter is modelled in terms of the other (independent) parameter, e.g., by using boundary functions which are formulated based on the independent parameters.

when calculating the structural flexibility index. Considering parameter dependencies by means of boundary functions requires more advanced modelling of the uncertainty space compared to the hyperbox approach (see Section 3.1.1), which can lead to more complex problems to solve. Therefore, it can be advantageous to know beforehand whether considering a given dependency when computing the (structural) flexibility index is likely to have an influence on the results. Such knowledge is usually very much dependent on the process, but it is possible to establish guidelines for specific types of processes. Such guidelines are presented for heat exchanger networks in **Paper VI**.

4.2.5 Identification of critical operating points

For each of the remaining structural design proposals, the framework proposes to identify the respective critical operating points. Considering the critical operating points in the discretized design under uncertainty problem (i.e., Problem (2.4), see also Section 2.2) ensures that each piece of equipment is of sufficient size to allow for steady-state flexible operation within the entire expected uncertainty space. As outlined in Section 3.1.3, the combinations of uncertain parameter values which are identified to be the critical operating points depend on how the expected uncertainty space is modelled. If independent operating periods and/or parameter dependencies are expected, Section 3.1.3 presents an approach which comprises three steps, as shown in Figure 4.2:

Step 1: Select one uncertainty space representing an independent operating period for which the critical operating points should be identified (if no independent operating periods have been identified, this step can be omitted).
Step 2: Identify the critical operating points for the respective uncertainty space when modelled as a hyperbox.
Step 3: Update or adjust the critical parameter values of those uncertain parameters which show correlating trends[5].

If different sets of critical operating points were identified due to the presence of independent operating periods, these sets must be considered simultaneously in the discretized design under uncertainty problem (see Section 4.2.7).

4.2.6 Identification of representative operating points

To decrease the number of scenarios or operating periods considered in the discretized design under uncertainty problem, (i.e., Problem (2.4)), the framework proposes to identify representative operating points. As mentioned in Section 2.2, the motivation for defining representative operating points is to enable a fair approximation of the objective function value (compared to considering all operating conditions) while avoiding issues with problem complexity.

Note that the identification of representative operating points is not equivalent to the identification of independent operating periods, even though the methodological approaches show similarities (see next paragraph). There is a fundamental difference concerning the

[5]Note that the critical parameter values of independent and/or dependent parameter may need to be adjusted while the critical parameter values of uncertain parameters which are not involved in a boundary function model remain unchanged (see Section 3.1.3 for further details).

origin and the objective of independent operating periods on the one hand and representative operating points on the other hand. More specifically, when identifying representative operating points, the probability that certain operating conditions will occur is the essential criteria - for example the mean operating point is the most trivial interpretation of a representative operating point since it represents the mean value of all operating conditions. This means that operating conditions with high probability are to be represented while operating conditions with low probability are omitted. On the other hand, when defining independent operating periods, the aim is to approximate the actual (i.e., expected or observed) uncertainty space as accurately as possible (see Section 4.2.2). Commonly, this uncertainty space is influenced or even defined by extreme operating conditions with (usually) low occurrence over the entire operating period and thereby low probability. Therefore, for defining independent operating periods, the essential criterion is that all expected/observed combinations of uncertain parameter values are captured (independent of probability) while all other combinations are excluded. Further explanation can be found in the Appendix of **Paper VI**.

To identify representative operating points, clustering algorithms such as Lloyd's clustering algorithm [85] commonly known as *K-means* can be utilized to identify centroids among all the operating points (for further information see, e.g., Jin and Han [86]). Alternatively, mean values for certain time-periods (daily, monthly, etc.) may be considered. If uncertain long-term development is expected to affect (future) operation, assumptions regarding how the operating conditions change are necessary. In this context, note that also the exact time-point of the implementation of the long-term development may be uncertain. Consequently, if cost for the entire life-time is to be considered in the final design problem, e.g., Total annualized cost, TAC, further assumptions are necessary. In this context, **Paper III** presents a strategy based on different extreme scenarios which can be utilized to obtain a lower and upper bound for the expected objective function value.

4.2.7 Design problem

After the critical operating points, CP, as well as the representative operating points are identified, OP, the discretized design under uncertainty problem (i.e., Problem (2.4)) can be solved for the remaining structural design proposals[6]. The solution yields the optimal design parameter values as well as an approximation of the expected objective function value, usually TAC, by optimizing the settings for the control variables. In this context, data for the investment and installation cost of the equipment as well as data for the operating costs for the process and the maintenance costs of the equipment are necessary. Additionally, possible cash flows generated by selling products or by avoiding expenses need to be considered.

4.2.8 Feasibility check

The final step of the framework is a feasibility assessment to check that the obtained values for the design parameters allow for steady-state flexible operation within the expected uncertainty space, i.e., a final validation of the previously identified critical operating points.

[6]Note that the critical operating points which were identified for different independent operating periods must be combined to one set to be included in Problem (2.4).

In this context, the flexibility index including design constraints may be utilized considering the identified independent operating periods and possible parameter dependencies. Alternatively, the optimal operation problem (see Problem (4.2)) may be solved for discrete operating points (OP_{all}). These operating points may be generated by means of sampling methods. Also, historical measurement data may be utilized, if available. Problem (4.2) is a single-period optimization problem to identify the optimal control variable settings for fixed design parameter values.

$$\left.\begin{aligned} &\min_{z_s} C_s(d, z_s, \theta_s) \\ &s.t.\ h_i(x_s, z_s, d, \theta_s) = 0;\ i \in I \\ &\quad g_j(x_s, z_s, d, \theta_s) \leq 0;\ j \in J \\ &\quad x_s, z_s, d, \theta_s \in \mathbb{R} \end{aligned}\right\} \ s \in OP_{all}. \tag{4.2}$$

In general, the (N)LP defined in Problem (4.2) is significantly easier to solve compared to the discretized design under uncertainty problem (i.e., Problem (2.4)) since the design specifications are fixed, i.e., the number of optimization variables is lower. If a feasible solution, $C_s(d, z_s, \theta_s) \in \mathbb{R}$, can be identified for each operating point, $s \in OP_{all}$, the design parameters derived in step 8 allow for operation at all expected/observed operating points. If the final feasibility check reveals that operation is not possible at certain operating points, i.e., that parts of the uncertainty space(s) are outside of the feasible region, the previously identified set of critical operating points is incomplete. In such a case, the infeasible operating points need to be added to the previously identified set of critical operating points, and steps 8 and 9 need to be repeated until a final indication of feasibility of the design is achieved.

Note, that a feasibility check based on solving Problem (4.2) for a discrete number of operating points cannot guarantee that the derived values for the design parameters are feasible for all possible operating points within the expected uncertainty space. Such a proof can only be obtained by solving the flexibility index problem, including design constraints. However, for the case study presented in **Paper VI**, it was observed that the flexibility index problem for the final feasibility check can be significantly more computationally burdensome compared to the flexibility index problem in the two-stage algorithm for identifying critical operating points. In this context, other work in the field which is based on the identification of critical operating points, e.g., Pintarič and Kravanja [34], suggests to conduct the final feasibility check based on Monte-Carlo process simulation for discrete operating points. Note, that the validity of the suggested feasibility check based on the evaluation of discrete operating points increases with the number of discrete operating points considered.

5

Selected results from the motivational examples

5.1 Motivational example 1

The structural layout of a retrofit proposal for increased heat recovery in a Swedish pulp mill was presented in Section 1.1.1. The retrofit proposal involves several new HEX units and Figure 5.1 shows a selected part of the process flow diagram presented in Section 1.1.1 which shows the proposed placement of the new HEX units. The identification of the optimal heat transfer areas of the different new HEX units in this heat recovery system is a design under uncertainty problem. The framework presented in Section 4.2 was utilized to identify the optimal solution to this problem, i.e., the solution that guarantees that the operational targets of the respective heat recovery system are met at all expected operating points while minimizing both investment cost as well as operational cost. A detailed documentation of the solution process and the results can be found in **Paper VI**. An overview of the most important results is provided hereafter.

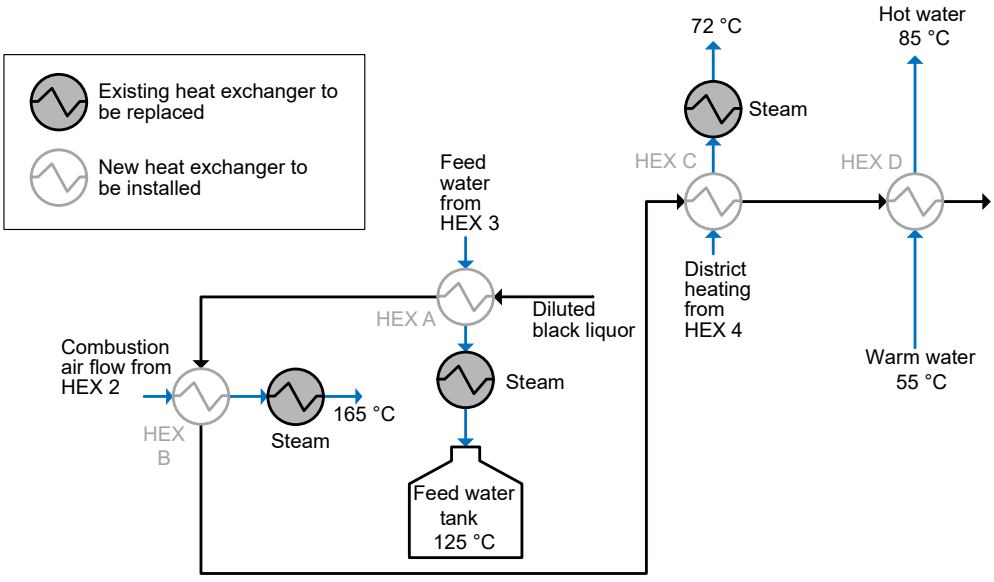


Figure 5.1: Selected part of the process flow diagram of the case study presented in Section 1.1.1 which shows the proposed placement of the new heat exchanger units according to the suggested retrofit for increased heat recovery.

In total, 15 uncertain parameters were identified, and their respective (measurement) locations are indicated in Figure 1.3. After filtering, 3415 historic operating points representing measurement values with hourly resolution were retained for the analysis. The

mill in question produces softwood and hardwood pulp in campaigns according to a regular schedule. Consequently, each operating point was allocated to one of two independent operating periods, corresponding to one of the production campaigns. Furthermore, an analysis of the available measurement data revealed that two parameter dependencies can be expected during softwood campaigns, while during hardwood campaigns an additional third parameter dependency can be expected. A structural flexibility analysis revealed that the design proposal of the retrofit is structurally feasible. The critical operating points for the two independent operating periods were identified using the approach presented in Section 3.1.3. Note that the critical operating points were identified when modelling the expected uncertainty spaces as hyperboxes but also when modelling the expected uncertainty spaces using boundary functions.

The discretized design under uncertainty problem (i.e., Problem (2.4)) was solved considering different numbers of representative operating points, namely 1 (mean value), 5, 25, 50 and 75 which were identified using the *K-means* clustering algorithm. For each set of representative operating points, three different cases which differ with respect to the critical operating points included in the problem formulation were considered:

Case 1: No critical operating points and only representative operating points.

Case 2: Critical operating points of uncertainty space modelled as hyperbox as well as representative operating points.

Case 3: Critical operating points of uncertainty space modelled as hyperpolygon using boundary functions as well as representative operating points.

The numerical results indicate that including critical operating points in the design problem significantly reduces the number of representative operating points that need to be considered to obtain a fair approximation of the expected objective function value. More precisely, for cases 2 and 3, a set of 25 representative operating points is sufficient to achieve a good approximation of the expected total annual cost (TAC). For case 1, the results indicate that the expected TAC obtained for the largest set of representative operating points considered in this work (75 points) may not be representative, i.e., the expected TAC may change significantly when considering sets with more than 75 representative operating points. Since higher numbers of representative operating points lead to increasing problem complexity, the additional effort for identifying critical operating points can be justified.

Furthermore, the results confirm that considering (only) representative operating points in Problem (2.4) (i.e., case 1) is problematic when aiming for steady-state flexible operation. Table 5.1 shows the results of the final feasibility check (see Section 4.2.8). More specifically, Table 5.1 shows the share of the 3415 expected operating points at which feasible operation was possible when solving Problem (2.4) for the investigated case study considering the different cases. The results indicate that steady-state flexible operation can only be guaranteed when considering the identified critical operating points in the constraints of the design problem (cases 2 and 3). When considering only representative operating points (i.e., case 1), feasible steady-state operation for all expected operating points was not possible for the investigated sets of representative operating points. Note that the absolute number of operating points at which feasible steady-state operation is not possible is bigger than the number of representative operating points - also for the

largest set investigated (75 points)[1].

Table 5.1: Percentage share of the 3415 expected operating points at which feasible operation was possible when solving Problem (2.4) for the investigated case study considering the different cases (case 1: only representative operating points; case 2: representative and critical operating points for hyperbox approach; case 3: representative and critical operating points for hyperpolygon approach).

Feasibility share [%]	Number of representative operating points considered in Problem (2.4)				
	1	5	25	50	75
Case 1	20	45	92	95	96
Case 2	100	100	100	100	100
Case 3	100	100	100	100	100

Moreover, the expected TAC as well as the optimized design specifications which were obtained for cases 2 and 3 were analysed and compared. Differences between cases 2 and 3 are of special interest since such differences originate from the different approaches to model the uncertainty space. Note that the final feasibility check revealed that steady-state flexible operation is possible for both cases independently of the chosen set of representative operating points (see Table 5.1). Table 5.2 presents the results obtained for cases 2 and 3 for the set of 25 representative operating points. The numbers show that the total optimized heat transfer area is lower when the uncertainty space is modelled as hyperpolygon (case 3) compared to the hyperbox approach (case 2). Additionally, the total heat transfer area is distributed differently among the different HEX units for cases 2 and 3.

Table 5.2: Comparison of optimal design specifications and expected total annualized cost, TAC, when modelling the expected uncertainty space as a hyperpolygon (case 3) or as hyperbox (case 2). Results obtained with 25 representative operating points.

Heat transfer area [m^2]	Case 2	Case 3	abs(Difference)
HEX A	1099	1195	96
HEX B	603	0	603
HEX C	0	0	0
HEX D	361	319	42
Feed water heater	1108	1035	73
District heating water heater	136	136	0
Combustion air heater	11 974	11 847	127
Total heat transfer area	15 281	14 532	749
Expected TAC [€]	6 484 349	6 313 190	171 159

Table 5.2 shows a 5% lower requirement of total heat transfer area as well as a 3% decrease in the expected TAC when boundary functions are used to model the uncertainty space(s), compared to results obtained for the hyperbox approach. These findings support the assumption that a more accurate model of the uncertainty space guarantees steady-state flexible operation with less overdesign and reduced cost. Note that the difference in the expected TAC is relatively small in this specific case. However, the difference may be

[1]Therefore, the strategy outlined in Section 4.2.8 may be considered ineffective since it proposes to include all operating points at which feasible operation was not possible as critical operating points in Problem (2.4).

significantly more pronounced for other cases studies.

Finally, as mentioned in Section 1.2, the traditional approach for dealing with uncertainty in (chemical) process design is to approximate the design specifications for a single operating point, such as the mean operating point, and apply empirical overdesign factors to these values. For illustrative purposes, this procedure was investigated. The design obtained by solving Problem (2.4) considering (only) the mean operating point was used as a basis. The obtained design specifications suggest the installation of HEX A and HEX D as well as steam heaters for the feed water and the combustion air (compare the process flow diagram of the suggested retrofit in Figure 5.1). Based on these design specifications, different overdesign factors were investigated. Note that only the initially suggested heat exchanger units were considered. The resulting share of operating points at which feasible operation can be obtained are presented in Table 5.3. The results show clearly that even when the heat transfer areas of the suggested HEX units are increased by additional 50%, steady-state operation at all expected operating points is not feasible.

Table 5.3: Impact of different overdesign factors on the feasibility share (share of the total operating points at which feasible operation is possible) as well as the TAC. The design obtained by solving Problem (2.4) considering (only) the mean operating point was used as a basis. Note that only the initially suggested heat exchanger units were considered.

Overdesign factor	20%	30%	40%	50%
Feasibility share [%]	59	67	76	83
Δ TAC (to case 1) [%]	15	23	30	38
Δ TAC (to case 3) [%]	-4	3	9	15

Table 5.3 also reports the difference in the expected TAC for a design based on overdesign factors compared to the expected TAC for a design with no overdesign (and no critical operating points, i.e., case 1, for one representative operating point) as well as compared to a design obtained by considering critical operating points (case 3, for one representative operating point). The results show that considering overdesign factors leads to a significant increase in the expected TAC which can be higher than the expected TAC for case 3 (overdesign factor $\geq$ 30%) while steady-state flexible operation still cannot be achieved. This illustrates the potential ineffectiveness of overdesign factors for design under uncertainty problems.

5.2 Motivational example 2

The case study presented in Section 1.1.2 illustrates the challenges that arise when planned but uncertain long-term development interferes with short-term operational disturbances. More specifically, with the currently installed process equipment, it is uncertain if steady-state operation will be feasible at all expected operating points after the nominal values change due to long-term development. For the analysis conducted in **Paper IV**, it was assumed that the UA-value[2] of HEX units HX $1-1$ and $1-2$ is 850 kW/K and 110

[2]The UA-value is the product of the heat transfer area, A, and the overall heat transfer coefficient, U.

kW/K for HEX unit HX 2. With the assumed design data and the variation data given in Table 1 in **Paper III**, the flexibility index for the operation at the current nominal operating point (values given in Figure 1.4) was calculated to be $\delta_{A,max} = 1.38$.

As mentioned in Section 1.1.2, a future implementation of biomass feedstock co-processing is considered by the refinery operating company, which represents an example for planned long-term development with uncertain consequences. More specifically, introducing biomass feedstock in the oil refinery is expected to cause a substantial increase of the nominal flow rates of streams 1 and 2 (see Figure 1.4). While it was assumed in **Paper III** that the exact values of the nominal flow rates after the feedstock switch can be predicted with good accuracy, in **Paper IV** it was assumed that the exact increase is uncertain. More specifically, in **Paper IV**, it was assumed that the nominal value of the flow rates of streams 1 and 2 may increase by up to 100% compared to current operation. When applying the reformulations of the flexibility index problem proposed in Section 3.2, a solution value of 0.5 is obtained for δ_B. This result was obtained when requiring that the expected operational short-term disturbances are (exactly) feasible (assuming that these expected short-term disturbances are not affected by the implementation of the biomass co-processing). Consequently, it can be concluded that the process configuration would be able to handle an increase of 50% in the flow rates while for any larger increase in flow rates the expected operational short-term disturbances may result in infeasible operating conditions.

6
Discussion

In this chapter, some of the assumptions, terminology and methodological choices underlying the contributions presented in this thesis are discussed.

Availability and quality of data for modelling the expected uncertainty space
In Section 3.1, it is proposed to analyse, prior to modelling the uncertainty space, if there are indications that certain combinations of the (individually expected) uncertain parameter values are not expected to occur. Based on this strategy, two cases were identified, namely the presence of dependencies in the uncertain parameters as well as the presence of independent operating periods. In both cases, combinations of the individually expected uncertain parameter values are not expected to occur, meaning that the hypervolume of the actual uncertainty space can be significantly smaller than the hypervolume of the hyperbox defined by the expected lower and upper bound values. For both cases, an approach was presented to model the expected uncertainty space aiming to enhance the resemblance with the actual uncertainty space compared to the hyperbox model.

The presented modelling approaches rely upon availability of good-quality data, especially data related to parameter dependencies. The availability of such data is commonly limited, especially for the greenfield design of chemical plants. If high-quality operating data is available, it should be noted that such data usually represents historic measurements. Thus, there is uncertainty regarding the extent to which such historic data also represents future operating conditions. However, even if no historic operating data is available, it is usually possible to assume that certain independent operating periods occur, such as differing operating conditions during summer and winter time (typical example of seasonality). Nevertheless, in certain projects, a rough estimation of the uncertainty space such as upper and lower bound values for the uncertain parameters for the entire expected operating period may be the best approximation possible at the early design planning stage.

Definition of uncertainty
This thesis aims to propose theoretical development in the field of deterministic flexibility analysis and methodological development in the field of design under uncertainty. For example, the framework proposed in Section 4.2 allows uncertainties to be considered when identifying the optimal design specifications for structural design proposals. These uncertainties may result from uncertain parameters which are defined outside the system boundary and can thereby not be affected by the respective system, while recourse action is possible to counteract changes in the uncertain parameters. The proposed development does not cover uncertain parameters which are not measurable, since recourse action is not possible in such cases. Future work may address such uncertainties based on the work of, e.g., Ostrovsky et al. [46] as well as Ochoa and Grossmann [12].

Usage of the terminology *critical (operating) points*
In each of the Figures, 2.1, 2.2, 2.3, 3.3, 3.4a, 3.4b and 3.7 the points at which the feasible uncertainty space coincides with the boundary of the feasible region are marked as critical points. Each of these points represents the solution point of a flexibility index problem, and it is common practise to denote such a point as critical point (compare, e.g., Ochoa and Grossmann [12] or Grossmann et al. [11]). This practise goes back to the work by Halemane and Grossmann [29] who defined critical points as those combinations of the uncertain parameter values for which the feasible region of operation is the smallest.

As outlined in Section 2.1.3, the terminology critical (operating) points was also used by Pintarič and Kravanja [32] to denote those combinations of the uncertain parameter values within the expected uncertainty space which demand the largest overdesign. However, the solution point(s) of the flexibility index problem must not necessarily be the point(s) demanding the largest overdesign. In fact, if the structural constraints of a process do not allow for feasible steady-state operation at all operating points within the expected uncertainty space, a critical point with respect to the definition by Pintarič and Kravanja [32] does not exist[1]. When steady-state operation is feasible with respect to the structural constraints, equipment (size) may limit the operation. Such a situation is visualized in Figure 2.3. Note that the point marked as critical point in Figure 2.3 is the solution point for the flexibility index considering a specific design parameter value (e.g., equipment size). However, this design parameter value does not allow for feasible steady-state operation at all operating points within the expected uncertainty space which is why the feasible uncertainty space is smaller than the expected uncertainty space. Therefore, the marked critical point does not match the definition by Pintarič and Kravanja. However, as outlined in Section 3.1.3, the solution point of the flexibility index (marked point in Figure 2.3), can be utilized to identify the critical point with respect to the definition by Pintarič and Kravanja. More specifically, the direction from the mean operating point to the marked critical point in Figure 2.3 also points to the point Pintarič and Kravanja defined as critical point. In fact, this linkage between the two points is utilized in the proposed approach to identify the critical operating points of a structural design proposal which demand the largest overdesign (see Section 3.1.3). In summary, the usage of the terminology critical operating points is not rigorously consistent (also in this work) while the context usually makes the intended meaning clear.

Sensitivity analysis vs. flexibility analysis
As mentioned in Section 1.4, in **Paper I** a systematic methodology is presented for applying automated sensitivity analyses in HEN retrofitting processes to evaluate the operability and energy efficiency of different retrofit proposals. The methodology is based on the idea that sensitivity analysis can be utilized to identify the influence of changes in the control settings on the operational target(s). The methodology aims to identify control settings which optimize an objective function such as energy efficiency or cost while maintaining feasible operation when operating conditions vary. It was further proposed in **Paper I** that the influence of (minor) design changes can be analysed by means of sensitivity analyses. The automated sensitivity analysis is based on (Monte Carlo) simulation for discrete operating conditions and/or (minor) design changes. If sufficiently many discrete simulation runs are conducted, an indication can be obtained about whether a

[1]Note that for this reason, structural feasibility is a premise for identifying critical points as defined by Pintarič and Kravanja [32].

process/plant can achieve steady-state flexible operation, i.e., if feasible steady-state operation can be maintained/achieved for all expected operating points by adjusting the control settings and/or investing in design changes. However, it is not possible to rigorously confirm that a process/plant can achieve steady-state flexible operation by means of (automated) sensitivity analyses which is why such strategies were not further pursued in this thesis. As an alternative to (automated) sensitivity analyses, flexibility analysis can be exploited to rigorously confirm that a process/plant can achieve steady-state flexible operation. In comparison to (automated) sensitivity analyses based on (Monte Carlo) simulation, flexibility analysis can lead to problem formulations which are computational burdensome to solve. For this reason, flexibility analysis may be of limited use, particularly for large-scale industrial applications. However, it was demonstrated in this thesis that flexibility analysis can also be applied for industrial case studies (see, e.g., the industrial case studies in **Papers III - VI**).

Outsourcing of design synthesis - smart move or ineffective procedure?
In Section 4.2, it is proposed to divide the design process of a design under uncertainty problem into two steps which are decoupled from each other. In the first step, different structural design proposals are developed for the process in question. Since steady-state flexible operation does not need to be rigorously addressed in this step, it is possible to utilize methodologies based on graphical-insights or other design synthesis approaches considering only a single or a low number of steady-state operating point(s). Thereafter, the different structural design proposals can be collected in a superstructure. For the second step, this thesis proposes a step-wise framework which helps identifying the best design proposal from the superstructure ensuring steady-state flexible operation.

Since the design synthesis step is not included in the framework proposed in this thesis, it can be used for both retrofit (re-design) as well as greenfield design problems. However, the proposed strategy also relies on the application of existing methodologies for generating good-quality structural design proposals. It should be noted that when following the different steps of the framework, the optimal solution with respect to a defined objective can (only) be found among the design proposals considered in the superstructure. Additionally, since the design synthesis and the framework are decoupled, there exist no feedback loops to modify a structural design based on results obtained during the analysis steps of the framework. It is the author's assumption that such feedback loops cause the computational burdens which limit the application of design synthesis approaches which rigorously address steady-state flexible operation during the design process.

Parameter dependencies - higher or lower deterministic flexibility index?
As illustrated by the numerical examples and the industrial case study presented in **Paper V**, dependencies between (some) of the uncertain parameters can have a significant influence on the deterministic flexibility index if the respective dependency is satisfactorily captured by means of a mathematical model. Grossmann and Floudas [41] also assumed that dependencies between (some of) the uncertain parameters should have a significant influence on the flexibility index calculated for a process. More precisely, the authors assumed that considering parameter dependencies in the model of the expected uncertainty space results in a higher flexibility index, compared to the flexibility index calculated for the hyperbox approach.

As mentioned in Section 2.1.2, Grossmann and Floudas [41] suggested to utilize single equation models for modelling the expected uncertainty space in the presence of parameter dependencies. To illustrate this, the authors presented a HEN example where one uncertain temperature (*here:* for simplicity T_B) was expressed as a linear function of another uncertain temperature (*here:* for simplicity T_A) (compare [41]). When including this linear equation, the authors reported that compared to the hyperbox approach (where T_A and T_B are assumed to vary independently by ± 10 K), the flexibility index increased by 0.08824. However, as outlined in **Paper V**, it can easily be identified that due to the assumed linear equation, the (absolute) expected disturbance range of the dependent parameter, T_B, decreased from ± 10 K to ± 8 K, i.e., for the case that T_A is 10 K higher than the nominal value, the linear equation returns a value for T_B which is not 10 K but only 8 K higher than the nominal value. Consequently, to allow a fair comparison, the flexibility index obtained for the linear equation needs to be compared to the flexibility index obtained for a modified hyperbox model of the expected uncertainty space. More specifically, for this modified hyperbox, T_A and T_B are assumed to vary independently but variations of ± 10 K are assumed for T_A whereas variations of ± 8 K are assumed for T_B. **Paper V** reports that for the modified hyperbox model, the flexibility index is also 0.08824 higher compared to the flexibility index calculated for the initial hyperbox model, i.e., the case where both parameters vary independently by ± 10 K. Therefore, it can be concluded that the example presented in Grossmann and Floudas [41] illustrates how changes in the maximum and minimum extreme values influence the result of the flexibility analysis. On the other hand, the same value for the flexibility index was obtained when:

1. including the linear equation to model a dependency between the uncertain parameters T_A and T_B,
2. assuming independent variation of T_A and T_B while adjusting the expected maximum and minimum extreme values of T_B (modified hyperbox uncertainty space).

Consequently, the linear equation did not help to exclude subsets of the hyperbox uncertainty space which limit the flexibility metric but within which no operating points are expected/observed. In fact, the consequence of including the linear equation is a reduction of the expected extreme values of the uncertain parameters. However, the absolute extreme values of the uncertain parameters are independent of possible parameter dependencies since a parameter dependency (only) describes the relation between the extreme values, e.g., if they occur at the same time point(s). Based on these results, it was concluded in **Paper V** that the example of Grossmann and Floudas [41] does not explicitly show the impact on the results of the flexibility analysis when subsets of the hyperbox model of the uncertainty pace which limit the flexibility metric (e.g. flexibility index) are excluded since no operating points are expected/observed in these subsets.

As mentioned in the previous paragraph, Grossmann and Floudas [41] assumed that considering dependencies in the uncertain parameters when modelling the uncertainty space results in a higher flexibility index, compared to the flexibility index calculated for a hyperbox model. This assumption is intuitive and was shared by other authors in the field, such as Pulsipher and Zavala [51] as well as Rooney and Biegler [75]. However, in **Paper V**, and in an additional conference contribution [74], it was demonstrated that when considering dependencies (by means of boundary functions) the flexibility index can be lower compared to the flexibility index calculated using a hyperbox uncertainty space:

1. Langner et al. [74]: a case was identified for which $\delta(T_{boundary}) = 0.27$ is smaller than $\delta(T_{box}) = 0.36$,
2. **Paper V** presents a case for the industrial case study where the $\delta(T_{boundary}) = 0.55$ is smaller than $\delta(T_{box}) = 0.64$

These observations contradict the intuitive assumption that considering parameter dependencies when modelling the expected uncertainty space results in a higher flexibility index, compared to the flexibility index based on the hyperbox model. In this context, **Paper V** concluded that in certain cases, the proposed usage of boundary functions to model the expected uncertainty space in the presence of parameter dependencies can bias the result of the flexibility analysis to put more emphasis on the (previously) selected dependent parameters. Note, that in certain situations this bias may be desirable, e.g., when the feasibility towards the uncertainty of selected parameters is prioritized. In case the bias should be avoided, the identified dependencies can be excluded (as, e.g., done in the industrial case study presented in **Paper V**) or a different approach needs to be utilized to model the expected uncertainty space. In this context, single equation models should, however, be avoided due to the risk of overestimating the flexibility (see Section 2.1.2).

7
Conclusion and Outlook

The purpose of this thesis was to present theoretical development in the field of deterministic flexibility analysis as well as methodological development in the field of design under uncertainty. Four research objectives were formulated, and hereafter concluding remarks are provided with respect to each of the objectives.

Modelling of the uncertainty space for deterministic flexibility analysis (Research Objective 1):
In this thesis, a strategy was defined to identify situations where a single hyperbox model of the uncertainty space can be inaccurate. The main idea is to investigate, prior to modelling the uncertainty space, if there are indications that certain combinations of the individual expected uncertain parameter values (i.e., values within the respective lower and upper bounds) are not expected to occur. Two situations were identified which potentially lead to such an occurrence:

- (some of) the uncertain parameters show correlating trends,
- (some) parameter values are only expected during certain operating periods (independent operating periods).

Several approaches were outlined to identify these situations and to build upon these insights when modelling the uncertainty space.

It was shown that existing approaches to model the expected uncertainty space when parameter dependencies are expected have significant shortcomings. More precisely, the value of the deterministic flexibility index can over- or underestimate the feasible disturbance range in which uncertain parameters may vary while feasible operation is achieved when using aforementioned approaches. It was concluded that overestimation of the feasible disturbance range occurs when the expected uncertainty space is underestimated using the selected model. Conversely, underestimation of the feasible disturbance range can occur when the expected uncertainty space is overestimated. An important insight of this work is that overestimating the flexibility of a process can have severe consequences since the infeasibility of certain operating conditions may not be identified before actual operation, and (very) costly retrofits may be required.

A new strategy was outlined to model the expected uncertainty space in the presence of parameter dependencies based on the definition of boundary functions utilizing the polygon convex hull of the expected uncertainty space. The proposed strategy allows irrelevant sub-spaces of the hyperbox model to be excluded, i.e., sub-spaces in which no operation is expected. It was concluded that the resulting shape of the modelled uncertainty space can be represented as a hyperpolygon (or irregular square in two dimensions).

A generic MI(N)LP formulation for the flexibility index problem was derived adopting the suggested strategy. It was concluded that overestimation of the feasible variation range can be avoided since all expected operating conditions are included in the analysis. At the same time, it should be noted that underestimation is inherent in any approach where geometrical shapes are used to approximate the actual uncertainty space. However, by adopting the proposed strategy, the level of underestimation is likely to be lower compared to existing approaches as long as a tight convex hull representation can be found, since in such cases the resemblance of the modelled uncertainty space with the expected uncertainty space should be satisfactory.

In addition to parameter dependencies, the occurrence of independent operating periods was investigated when modelling the expected uncertainty space. Independent operating periods were defined based on the observation that certain realizations of the uncertain parameters only occur during certain periods of the time-horizon considered for the flexibility analysis. It was shown that the overall uncertainty space which is defined by all expected operating conditions can be divided into several sub-spaces representing the different independent operating periods. It was further shown that this structuring of operating conditions into individual uncertainty spaces can exclude sub-spaces of the overall uncertainty space. These excluded sub-spaces represent combinations of the uncertain parameter values that are not expected, leading to a more accurate modelling of the expected uncertainty space. Possible causes for independent operating periods were presented, such as:

- seasonality in the operating data (e.g., differences between day-time and night-time operation or seasonal variation)
- different production campaigns

It was concluded that by allocating realizations of the uncertain parameters to different independent operating periods and modelling those periods as individual uncertainty spaces, some time-dependency can be enabled in the flexibility analysis.

If parameter dependencies and/or independent operating periods are present, modelling the expected uncertainty space as a single hyperbox can lead to situations in which identified critical operating points correspond to combinations of uncertain parameter values which are not expected to occur. To avoid such problems, a new strategy for identifying critical operating points was presented. The proposed strategy was used to identify the critical operating points for the industrial case study presented in Section 1.1.1. The results show that different combinations of uncertain parameter values constitute critical operating points when modelling the expected uncertainty space as a hyperbox or using boundary functions. While steady-state flexible operation of the respective process design was achieved for both cases, the critical operating points identified for the hyperbox model demanded more heat transfer area, resulting in higher projected total annualized cost. Consequently, it can be concluded that an overly simplified model of the uncertainty space may lead to overdesign and unnecessary costs.

Deterministic flexibility analysis considering uncertain long-term development (Research Objective 2):
In this thesis, situations were outlined in which planned long-term development of pro-

cesses interferes with short-term operational disturbances. It was shown that such instances can be interpreted as special cases of independent operating periods. In such cases, it can be concluded that the traditional definition of flexible processes, i.e., allowing for steady-state flexible operation around a single nominal operating point, can be insufficient. Planned long-term development can result in temporary or permanent changes of the nominal operating conditions. Consequently, it may be relevant to identify which changes in the nominal operating conditions connected to long-term development are feasible, i.e., allowing for feasible operation considering short-term operational disturbances. In line with this conclusion, a new approach was outlined to perform flexibility analysis when long-term development interferes with operational short-term disturbances. The presented approach was utilized to analyse the industrial case study presented in Section 1.1.2.

Automated HEN modelling (Research Objective 3):
In this thesis, an automated HEN modelling strategy was proposed which can be implemented in any high-level programming language. The modelling strategy builds upon the observation that HENs are usually based on similar sets of equations which implies that large parts of the modelling process can be automated. The proposed strategy can handle complexities commonly present in industrial HENs such as stream splits, closed loops or re-circulation. A table-based representation of the HEN was proposed, which can be applied directly to a process flowsheet and a transformation to the commonly used but limited grid-diagram representation is not necessary. As a result, error-prone manual definition of mathematical constraints can be avoided easing the mathematical modelling, especially of large-scale HENs.

A framework combining valuable designer input with the efficiency of mathematical programming for design under uncertainty problems (Research Objective 4):
In this thesis, a framework was presented which aims to support designers in (early) design stage screening processes when dealing with chemical process design under uncertainty. The framework is a step-wise approach which identifies:

- whether structural design proposals are structurally feasible,
- the most cost-efficient overdesign of equipment size required to guarantee steady-state flexible operation,
- a basis to compare different structural design proposals with respect to a given objective.

The proposed framework can be used to evaluate and improve one or several structural design proposals which are provided by the designer. In this way, it is possible to utilize proven design methodologies available for single steady-state operating conditions to generate the basic structure of the design. This setup also enables incorporating non-quantifiable knowledge such as experience-based heuristics in the design proposals, which can be advantageous for complex industrial case studies. Since the (structural) design proposals are defined prior to the application of the framework, the framework can be utilized to guide both greenfield and retrofit design projects. As steady-state flexible operation is difficult to address using the aforementioned design synthesis methodologies, the framework comprises several steps based on deterministic flexibility analysis, i.e., cal-

culation of the flexibility index and the identification of critical operating points. In this context, the proposed framework evaluates if dependencies in the uncertain parameters as well as independent operating periods are present to adapt the modelling of the expected uncertainty space accordingly. The proposed framework was utilized to identify the cost-optimal solution for the retrofit proposal presented in Section 1.1.1. More specifically, the proposed framework was used to identify the optimal heat transfer areas of the different new HEX units that always meet the operational target at the lowest cost.

7.1 Considerations for future work

In Section 3.1, two situations were identified for which there are indications that certain combinations of the individually expected uncertain parameter values (i.e., values within the respective lower and upper bounds) are not expected to occur, namely parameter dependencies and independent operating periods. In future work, it would be valuable to identify additional situations, which may also be case-specific.

Furthermore, in Section 3.1, two approaches were presented to enhance the modelling of the expected uncertainty space for deterministic flexibility analysis when parameter dependencies and/or independent operating periods are expected. As these approaches allow for a better resemblance with the actual expected uncertainty space, the approaches may be considered for other tasks which demand a model of the expected uncertainty space, such as the volumetric flexibility index proposed by Lai and Hui [52].

Section 3.2 outlined a new approach to perform deterministic flexibility analysis when planned but uncertain long-term development interferes with operational short-term disturbances. To the author's knowledge such overlaying uncertainty sources have not been considered in any work connected to flexibility analysis, and future work is needed to explore if such considerations are of relevance even for stochastic [48, 49] or volumetric flexibility analysis [52].

This thesis proposed theoretical development in the field of deterministic flexibility analysis applicable to cases with measurable uncertain parameters. However, uncertain parameters which are not measurable were not addressed since in such cases recourse action is not possible. Future work may address such uncertainties based on the work of, e.g., Ostrovsky et al. [46] as well as Ochoa and Grossmann [12].

Finally, the design strategy presented in Section 4.2 proposed to divide a design under uncertainty problem into a design synthesis step which enables direct input from the designer, followed by a number of subsequent steps which are summarized in a framework presented in this thesis. The proposed framework and the design synthesis step are decoupled from each other considering no feedback loops since the author assumes that such feedback loops may limit the application to small scale examples. Future work may identify straight-forward procedures to allow feedback from the analysis steps to be translated into design modifications while maintaining the computational effort within acceptable limits.

Bibliography

[1] IPCC, In *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Shukla, P., Skea, J., Slade, R., Khourdajie, A. A., van Diemen, R., McCollum, D., Pathak, M., Some, S., Vyas, P., Fradera, R., Belkacemi, M., Hasija, A., Lisboa, G., Luz, S., Malley, J., Eds.; Cambridge University Press: Cambridge, UK and New York, NY, USA, 2022.

[2] Swaney, R. E.; Grossmann, I. E. An index for operational flexibility in chemical process design. Part I: Formulation and Theory. *AIChE Journal* **1985**, *31*, 621–630.

[3] UNFCCC, The Paris Agreement. 2015; `https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement`.

[4] IEA, *World Energy Outlook 2022*; 2022; p 524.

[5] Friedlingstein, P. et al. Global Carbon Budget 2020. *Earth System Science Data* **2020**, *12*, 3269–3340.

[6] Minx, J. C. et al. A comprehensive and synthetic dataset for global, regional, and national greenhouse gas emissions by sector 1970–2018 with an extension to 2019. *Earth System Science Data* **2021**, *13*, 5213–5252.

[7] IEA, *Energy Efficiency 2019*; 2019; p 110.

[8] Papapetrou, M.; Kosmadakis, G.; Cipollina, A.; La Commare, U.; Micale, G. Industrial waste heat: Estimation of the technically available resource in the EU per industrial sector, temperature level and country. *Applied Thermal Engineering* **2018**, *138*, 207–216.

[9] van Dyk, S.; McMillan, J. D.; Saddler, J. Potential synergies of drop-in biofuel production with further co-processing at oil refineries. **2019**,

[10] Grossmann, I.; Halemane, K.; Swaney, R. Optimization strategies for flexible chemical processes. *Computers & Chemical Engineering* **1983**, *7*, 439–462.

[11] Grossmann, I. E.; Calfa, B. A.; Garcia-Herreros, P. Evolution of concepts and models for quantifying resiliency and flexibility of chemical processes. *Computers and Chemical Engineering* **2014**, *70*, 22–34.

[12] Ochoa, M. P.; Grossmann, I. E. Novel MINLP formulations for flexibility analysis for measured and unmeasured uncertain parameters. *Computers and Chemical Engineering* **2020**, *135*.

[13] Grossmann, I. E.; Sargent, R. W. Optimum design of chemical plants with uncertain parameters. *AIChE Journal* **1978**, *24*, 1021–1028.

[14] Steimel, J.; Engell, S. Conceptual design and optimization of chemical processes under uncertainty by two-stage programming. *Computers & Chemical Engineering* **2015**, *81*, 200–217.

[15] Pistikopoulos, E. N.; Ierapetritou, M. G. Novel approach for optimal process design under uncertainty. *Computers & Chemical Engineering* **1995**, *19*, 1089–1110.

[16] Acevedo, J.; N. Pistikopoulos, E. A Parametric MINLP Algorithm for Process Synthesis Problems under Uncertainty. *Industrial & Engineering Chemistry Research* **1996**, *35*, 147–158.

[17] Rooney, W. C.; Biegler, L. T. Optimal process design with model parameter uncertainty and process variability. *AIChE Journal* **2003**, *49*, 438–449.

[18] Sahinidis, N. V. Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering* **2004**, *28*, 971–983.

[19] Birge, J. R.; Louveaux, F. *Springer Series in Operations Research and Financial Engineering*; 2011.

[20] Wei, J.; Realff, M. J. Sample average approximation methods for stochastic MINLPs. *Computers & Chemical Engineering* **2004**, *28*, 333–346.

[21] Karuppiah, R.; Martín, M.; Grossmann, I. E. A simple heuristic for reducing the number of scenarios in two-stage stochastic programming. *Computers and Chemical Engineering* **2010**,

[22] Pintarič, Z. N.; Kasaš, M.; Kravanja, Z. Sensitivity Analyses for Scenario Reduction in Flexible Flow Sheet Design with a Large Number of Uncertain Parameters. *AIChE Journal* **2013**, *59*, 2862–2871.

[23] Alvarez, R.; Moser, A.; Rahmann, C. A. Novel methodology for selecting representative operating points for the TNEP. *IEEE Transactions on Power Systems* **2017**, *32*, 2234–2242.

[24] Salameh, M.; Brown, I. P.; Krishnamurthy, M. Fundamental Evaluation of Data Clustering Approaches for Driving Cycle-Based Machine Design Optimization. *IEEE Transactions on Transportation Electrification* **2019**, *5*, 1395–1405.

[25] Reichenberg, L.; Siddiqui, A. S.; Wogrin, S. Policy implications of downscaling the time dimension in power system planning models to represent variability in renewable output. *Energy* **2018**, *159*, 870–877.

[26] Teichgraeber, H.; Brandt, A. R. Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. *Applied Energy* **2019**, *239*, 1283–1293.

[27] Chen, H.; Liu, X.; Demerdash, N. A.; El-Refaie, A. M.; Chen, Z.; He, J. Computationally efficient optimization of a five-phase flux-switching PM machine under different operating conditions. *IEEE Transactions on Vehicular Technology* **2019**, *68*, 6495–6508.

[28] Mozafari, Y.; Rosehart, W. Representative operating scenario selection with algebraic multi-grid clustering for integrated energy systems planning. *International Journal of Electrical Power & Energy Systems* **2023**, *146*, 108767.

[29] Halemane, K. P.; Grossmann, I. E. Optimal process design under uncertainty. *AIChE Journal* **1983**, *29*, 425–433.

[30] Pintarič, Z. N.; Kravanja, Z. A strategy for MINLP synthesis of flexible and operable processes. *Computers and Chemical Engineering* **2004**,

[31] Pintarič, Z. N.; Kravanja, Z. A-priori identification of critical points for the design and synthesis of flexible process schemes. *Computer Aided Chemical Engineering* **2006**, *21*, 503–508.

[32] Pintarič, Z. N.; Kravanja, Z. Identification of critical points for the design and synthesis of flexible processes. *Computers & Chemical Engineering* **2008**, *32*, 1603–1624.

[33] Pintarič, Z. N.; Kravanja, Z. A multilevel MINLP methodology for the approximate stochastic synthesis of flexible chemical processes. *Revista de Chimie* **2007**, *58*, 387–391.

[34] Pintarič, Z. N.; Kravanja, Z. A methodology for the synthesis of heat exchanger networks having large numbers of uncertain parameters. *Energy* **2015**, *92*, 373–382.

[35] Zirngast, K.; Kravanja, Z.; Novak Pintarič, Z. An improved algorithm for synthesis of heat exchanger network with a large number of uncertain parameters. *Energy* **2021**, *233*, 121199.

[36] Saboo, A. K.; Morari, M.; Woodcock, D. C. Design of resilient processing plants-VIII. A resilience index for heat exchanger networks. *Chemical Engineering Science* **1985**, *40*, 1553–1565.

[37] Di Pretoro, A.; Negny, S.; Montastruc, L. The traveling deliveryman problem under uncertainty: Fundamentals for flexible supply chains. *Computers & Chemical Engineering* **2022**, *160*, 107730.

[38] da Silva, P. R.; Aragão, M. E.; Trierweiler, J. O.; Trierweiler, L. F. Integration of hydrogen network design to the production planning in refineries based on multi-scenarios optimization and flexibility analysis. *Chemical Engineering Research and Design* **2022**, *187*, 434–450.

[39] Luo, J.; Moncada, J.; Ramirez, A. Development of a Conceptual Framework for Evaluating the Flexibility of Future Chemical Processes. *Industrial & Engineering Chemistry Research* **2022**, *61*, 3219–3232.

[40] Zhang, Q.; Grossmann, I. E.; Lima, R. M. On the relation between flexibility analysis and robust optimization for linear systems. *AIChE Journal* **2016**, *62*, 3109–3123.

[41] Grossmann, I. E.; Floudas, C. A. Active constraint strategy for flexibility analysis in chemical processes. *Computers & Chemical Engineering* **1987**, *11*, 675–693.

[42] Raspanti, C.; Bandoni, J.; Biegler, L. New strategies for flexibility analysis and design under uncertainty. *Computers & Chemical Engineering* **2000**, *24*, 2193–2209.

[43] Li, J.; Du, J.; Zhao, Z.; Yao, P. Efficient Method for Flexibility Analysis of Large-Scale Nonconvex Heat Exchanger Networks. *Industrial and Engineering Chemistry Research* **2015**, *54*, 10757–10767.

[44] Floudas, C. A.; Grossmann, I. E. Synthesis of flexible heat exchanger networks with

uncertain flowrates and temperatures. *Computers and Chemical Engineering* **1987**, *11*, 319–336.

[45] Papalexandri, K. P.; Pistikopoulos, E. N. Synthesis and Retrofit Design of Operable Heat Exchanger Networks. 1. Flexibility and Structural Controllability Aspects. *Industrial & Engineering Chemistry Research* **1994**, *33*, 1718–1737.

[46] Ostrovsky, G. M.; Datskov, I. V.; Achenie, L. E.; Volin, Y. M. Process uncertainty: Case of insufficient process data at the operation stage. *AIChE Journal* **2003**, *49*, 1216–1232.

[47] Ochoa, M. P.; García-Muñoz, S.; Stamatis, S.; Grossmann, I. E. Novel flexibility index formulations for the selection of the operating range within a design space. *Computers & Chemical Engineering* **2021**, *149*, 107284.

[48] Pistikopoulos, E. N.; Mazzuchi, T. A. A novel flexibility analysis approach for processes with stochastic parameters. *Computers & Chemical Engineering* **1990**, *14*, 991–1000.

[49] Straub, D. A.; Grossmann, I. E. Integrated stochastic metric of flexibility for systems with discrete state and continuous parameter uncertainties. *Computers & Chemical Engineering* **1990**, *14*, 967–985.

[50] Straub, D. A.; Grossmann, I. E. Design optimization of stochastic flexibility. *Computers & Chemical Engineering* **1993**, *17*, 339–354.

[51] Pulsipher, J. L.; Zavala, V. M. A mixed-integer conic programming formulation for computing the flexibility index under multivariate gaussian uncertainty. *Computers and Chemical Engineering* **2018**, *119*, 302–308.

[52] Lai, S. M.; Hui, C. W. Process flexibility for multivariable systems. *Industrial and Engineering Chemistry Research* **2008**, *47*, 4170–4183.

[53] Zheng, C.; Zhao, F.; Zhu, L.; Chen, X. Analytical solution of volumetric flexibility through symbolic computation. *Chemical Engineering Science* **2021**, *239*, 116643.

[54] Sreepathi, B. K.; Rangaiah, G. P. Review of Heat Exchanger Network Retrofitting Methodologies and Their Applications. *Industrial & Engineering Chemistry Research* **2014**, *53*, 11205–11220.

[55] Kang, L.; Liu, Y. Synthesis of flexible heat exchanger networks: A review. *Chinese Journal of Chemical Engineering* **2019**, *27*, 1485–1497.

[56] Kotjabasakis, E.; Linnhoff, B. Sensitivity tables for the design of flexible processes (1)–How much contingency in heat exchanger networks is cost-effective? *Chemical engineering research & design* **1986**, *64*, 197–211.

[57] Hafizan, A. M.; Klemeš, J. J.; Alwi, S. R. W.; Manan, Z. A.; Hamid, M. K. A. Temperature disturbance management in a heat exchanger network for maximum energy recovery considering economic analysis. *Energies* **2019**, *12*.

[58] Linnhoff, B.; Vredeveld, D. R. Pinch Technology Has Come of Age. *Chemical Engineering Progress* **1984**, *80*, 33–40.

[59] Floudas, C. A.; Grossmann, I. E. Automatic generation of multiperiod heat exchanger network configurations. *Computers and Chemical Engineering* **1987**, *11*, 123–142.

[60] Aaltola, J. Simultaneous synthesis of flexible heat exchanger network. *Applied Thermal Engineering* **2002**, *22*, 907–918.

[61] Yee, T. F.; Grossmann, I. E. Simultaneous optimization models for heat integration. II. Heat exchanger network synthesis. *Computers & Chemical Engineering* **1990**, *14*, 1165–1184.

[62] Verheyen, W.; Zhang, N. Design of flexible heat exchanger network for multi-period operation. *Chemical Engineering Science* **2006**, *61*, 7730–7753.

[63] Short, M.; Isafiade, A. J.; Fraser, D. M.; Kravanja, Z. Two-step hybrid approach for the synthesis of multi-period heat exchanger networks with detailed exchanger design. *Applied Thermal Engineering* **2016**, *105*, 807–821.

[64] Tveit, T. M.; Aaltola, J.; Laukkanen, T.; Laihanen, M.; Fogelholm, C. J. A framework for local and regional energy system integration between industry and municipalities-Case study UPM-Kymmene Kaukas. *Energy* **2006**, *31*, 2162–2175.

[65] Čuček, L.; Boldyryev, S.; Klemeš, J. J.; Kravanja, Z.; Krajačić, G.; Varbanov, P. S.; Duić, N. Approaches for retrofitting heat exchanger networks within processes and Total Sites. *Journal of Cleaner Production* **2019**, *211*, 884–894.

[66] Persson, J.; Berntsson, T. Influence of seasonal variations on energy-saving opportunities in a pulp mill. *Energy* **2009**, *34*, 1705–1714.

[67] Kamel, D. A.; Gadalla, M. A.; Abdelaziz, O. Y.; Labib, M. A.; Ashour, F. H. Temperature driving force (TDF) curves for heat exchanger network retrofit – A case study and implications. *Energy* **2017**, *123*, 283–295.

[68] Bonhivers, J. C.; Srinivasan, B.; Stuart, P. R. New analysis method to reduce the industrial energy requirements by heat-exchanger network retrofit: Part 1 – Concepts. *Applied Thermal Engineering* **2017**, *119*, 659–669.

[69] Lal, N. S.; Atkins, M. J.; Walmsley, T. G.; Walmsley, M. R.; Neale, J. R. Insightful heat exchanger network retrofit design using Monte Carlo simulation. *Energy* **2019**, *181*, 1129–1141.

[70] Kang, L.; Liu, Y. Retrofit of heat exchanger networks for multiperiod operations by matching heat transfer areas in reverse order. *Industrial & Engineering Chemistry Research* **2014**, *53*, 4792–4804.

[71] Kang, L.; Liu, Y. Minimizing investment cost for multi-period heat exchanger network retrofit by matching heat transfer areas with different strategies. *Chinese Journal of Chemical Engineering* **2015**, *23*, 1153–1160.

[72] Kang, L.; Liu, Y. A systematic strategy for multi-period heat exchanger network retrofit under multiple practical restrictions. *Chinese Journal of Chemical Engineering* **2017**, *25*, 1043–1051.

[73] Stampfli, J. A.; Ong, B. H.; Olsen, D. G.; Wellig, B.; Hofmann, R. Applied heat exchanger network retrofit for multi-period processes in industry: A hybrid evolutionary algorithm. *Computers & Chemical Engineering* **2022**, *161*, 107771.

[74] Langner, C.; Svensson, E.; Harvey, S. Flexibility analysis of chemical processes con-

sidering dependencies between uncertain parameters. *Computer Aided Chemical Engineering* **2021**, *50*, 1105–1110.

[75] Rooney, W. C.; Biegler, L. T. Incorporating joint confidence regions into design under uncertainty. *Computers & Chemical Engineering* **1999**, *23*, 1563–1575.

[76] Marselle, D. F.; Morari, M.; Rudd, D. F. Design of resilient processing plants—II Design and control of energy management systems. *Chemical Engineering Science* **1982**, *37*, 259–270.

[77] Li, J.; Du, J.; Zhao, Z.; Yao, P. Structure and area optimization of flexible heat exchanger networks. *Industrial and Engineering Chemistry Research* **2014**, *53*, 11779–11793.

[78] Smith, R. *Chemical process design and integration*, 2nd ed.; John Wiley & Sons, Ltd, 2016; p 896.

[79] Shah, R. K.; Sekulic, D. P. *Fundamentals of Heat Exchanger Design*; John Wiley & Sons, Ltd, 2012; pp 127–211.

[80] VDI-Gesellschaft Verfahrenstechnik und Chemieingenieurwesen, *VDI Heat Atlas*, 2nd ed.; Springer-Verlag, 2010; p 1585.

[81] Escobar, M.; Trierweiler, J. O.; Grossmann, I. E. Simultaneous synthesis of heat exchanger networks with operability considerations: Flexibility and controllability. *Computers and Chemical Engineering* **2013**,

[82] Boslaugh, S.; Watters, P. A. *Statistics in a nutshell - a desktop quick reference.*; O'Reilly, 2008; pp I–XXI, 1–452.

[83] Hyndman, R. J.; Athanasopoulos, G. Forecasting: principles and practice. 2nd edition, OTexts: Melbourne, Australia, 2018; `https://otexts.com/fpp2/`, Accessed on 2022-11-11.

[84] Dimitriadis, V. D.; Pistikopoulos, E. N. Flexibility Analysis of Dynamic Systems. *Industrial & Engineering Chemistry Research* **1995**, *34*, 4451–4462.

[85] Lloyd, S. P. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory* **1982**, *28*, 129–137.

[86] Jin, X.; Han, J. In *Encyclopedia of Machine Learning*; Sammut, C., Webb, G. I., Eds.; Springer US: Boston, MA, 2010; pp 563–564.

[87] Biegler, L.; Grossmann, I. E.; Westerberg, A. W. *Systematic Methods of Chemical Process Design*, 1st ed.; Pearson Education, 1997.

[88] Floudas, C. A.; Gümüş, Z. H.; Ierapetritou, M. G. Global Optimization in Design under Uncertainty: Feasibility Test and Flexibility Index Problems. *Industrial and Engineering Chemistry Research* **2001**, *40*, 4267–4282.

A

Theory on the deterministic flexibility index

The problem formulation of the flexibility index for a hyperbox uncertainty space is given in Problem (A.1) as derived by Swaney and Grossmann [2]. Note that in Problem (A.1), the uncertainty space is scaled with the scalar δ, $T_{box}(\delta)$. In Problem (A.1), Swaney and Grossmann [2] included a *max min max* constraint which is known as the feasibility constraint to find the largest value for the scaling parameter, δ so that all constraint functions are feasible, $f_j(d, z, \theta) \leq 0$, while accounting for recursive actions, i.e., adjustments of the control variables z. Since the uncertainty space is scaled, $T_{box}(\delta)$ can be interpreted as the largest scaled hyperbox which can be inscribed in the feasible region.

$$\begin{aligned} FI = & \max \delta \\ & s.t. \ \max_{\theta \in T} \min_{z} \max_{j \in J} f_j(d, z, \theta) \leq 0 \\ & \quad T_{box}(\delta) = \left\{ \theta_i \mid \theta_{i,N} - \delta \Delta \theta_i^- \leq \theta_i \leq \theta_{i,N} + \delta \Delta \theta_i^+ \right\} \forall \ \theta_i \in \theta \\ & \quad \delta \geq 0 \end{aligned} \tag{A.1}$$

To determine the flexibility index, Swaney and Grossmann [2] proposed search procedures for the special case of exclusively convex constraint functions, $f_j(d, z, \theta)$, in which the solution of (A.1) corresponds to vertices of $T_{box}(\delta)$. To overcome the limitations of vertex exploration (exponential number of vertices, 2^θ, and convexity of constraint functions $f_j(d, z, \theta)$), Grossmann and Floudas [41] proposed a reformulation of Problem (A.1) to solve the feasibility constraint explicitly without relying on the assumption of critical points corresponding to vertices of the hyperbox uncertainty space. They reformulated the feasibility constraint in Problem (A.1) to explicitly search for the solution on the boundary of the feasible region, $\psi(d, \theta) = 0$, yielding the bi-level optimization problem given in Problem (A.2).

$$\begin{aligned} FI = & \min \delta \\ & s.t. \ \psi(d, \theta) = 0 \\ & \quad \psi(d, \theta) = \min_{z,u} u \\ & \qquad s.t. \ f_j(d, z, \theta) \leq u, \ j \in J \\ & \quad T_{box}(\delta) = \left\{ \theta_i \mid \theta_{i,N} - \delta \Delta \theta_i^- \leq \theta_i \leq \theta_{i,N} + \delta \Delta \theta_i^+ \right\} \forall \ \theta_i \in \theta \\ & \quad \delta \geq 0 \end{aligned} \tag{A.2}$$

The advantage of Problem (A.2) is that the feasibility constraint can be solved explicitly by replacing the lower-level optimization problem, $\psi(d, \theta) = \min u$, by its Karush-Kuhn-

Tucker (KKT) optimality conditions. Consequently, a single-level non-linear program (NLP) for the deterministic flexibility index is obtained. Grossmann and Floudas [41] further identified that if the gradients, $\partial f_j/\partial z$, are linearly independent, there will be $n_z + 1$ active inequality constraints at the solution. Based on this observation, the authors formulated a search strategy, the active constraint strategy, which sequentially calculates the flexibility index for each active set, and eventually returns the smallest solution value found during the search[1]. To implement the active constraint strategy, the authors simplified the complementarity conditions with mixed-integer constraints. Later, Biegler et al. [87] found that if linear independence of the gradients, $\partial f_j/\partial z$, cannot be guaranteed, the number of active inequality constraints can be relaxed. The mixed-integer representation of the KKT-conditions and the relaxed constraint on the number of active inequality constraints is given by Eq. (A.3a) to Eq. (A.3g). Note that the $Big - M$ parameter in Eq. (A.3d) represents an upper bound to the slack variables of the inequality constraints. Furthermore, it can be noted that the resulting mixed-integer linear/non-linear program (MI(N)LP) does not require strict convexity of the constraint functions, $f_j(x, z, \theta) \leq 0,\ j \in J$, and thereby does not rely on the assumption of critical points corresponding to vertices. To also be able to solve non-convex system formulations in general, the active constraint strategy was later extended by Floudas et al. [88] to a global solution algorithm.

$$\sum_{j \in J} \lambda_j = 1 \tag{A.3a}$$

$$f_j(d, z, \theta) + s_j = 0,\ j \in J \tag{A.3b}$$

$$\sum_{j \in J} \lambda_j \frac{\partial f_j}{\partial z} = 0 \tag{A.3c}$$

$$s_j - M(1 - y_j) \leq 0,\ j \in J \tag{A.3d}$$

$$\lambda_j - y_j \leq 0,\ j \in J \tag{A.3e}$$

$$\sum_{j \in J} y_j \leq n_z + 1 \tag{A.3f}$$

$$s_j \geq 0, \lambda_j \geq 0, y_j \in 0, 1\ j \in J \tag{A.3g}$$

[1]Note that the number of potential active sets is limited by the constraint that exactly n_z+1 inequality constraints must be active at the solution.