

Nodal Control and Probabilistic Constrained Optimization using the Example of Gas Networks

Nodale Steuerung und Optimierung unter
Wahrscheinlichkeitsrestriktionen am
Beispiel von Gasnetzwerken

Der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg

zur
Erlangung des Doktorgrades

Dr. rer. nat.

vorgelegt von
Michael Schuster
aus Nürnberg

Als Dissertation genehmigt
von der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung: 23. Juli 2021

Vorsitzender des Promotionsorgans: Prof. Dr. Wolfgang Achtziger

Gutachter: Prof. Dr. Martin Gugat

Prof. Dr. Rüdiger Schultz

Zusammenfassung

In dieser Arbeit analysieren wir den stationären und dynamischen Gasfluss mit zufallsbehafteten Randdaten in Pipelinenetzwerken, der mit Hilfe der isothermalen Eulergleichungen modelliert wird. Die zufälligen Randdaten werden mittels Wahrscheinlichkeitsverteilungen abgebildet. Sie repräsentieren dabei den a priori unbekanntem Gasverbrauch der Kunden. Das Ziel der Arbeit ist die Analyse von ausgewählten Optimierungsproblemen mit probabilistischen Nebenbedingungen im Kontext des Gastransports.

Um die Wahrscheinlichkeit dafür zu berechnen, dass ein zufälliger Gasverbrauch zulässig ist, nutzen wir sowohl die Idee des Kerndichteschätzers, als auch die der sphärisch radialen Zerlegung. Zulässig heißt in diesem Kontext, dass das geforderte Gas so durch das Netzwerk transportiert werden kann, dass gegebene Druckschranken an den Knoten eingehalten werden. Außerdem diskutieren wir die Vor- und Nachteile beider Verfahren. Im stationären Fall erweitern wir unser Model für den Gastransport in Pipelinenetzwerken um Kompressorsteuerungen und Druckschranken an den Eingangsknoten und berechnen ebenfalls die Wahrscheinlichkeit dafür, dass ein zufälliger Gasverbrauch zulässig ist. Im dynamischen Fall ist der zufällige Gasverbrauch zeitabhängig. Dieser wird mit zufälligen Fourierreihen modelliert.

Des Weiteren analysieren wir ausgewählte Optimierungsprobleme mit Wahrscheinlichkeitsrestriktionen, bei denen wir die Wahrscheinlichkeitsrestriktionen mittels Kerndichteschätzer approximieren. Wir zeigen einerseits die Existenz von Lösungen der exakten und ihrer approximierten Optimierungsprobleme und andererseits zeigen wir, dass bei hinreichend genauer Approximation der Wahrscheinlichkeitsrestriktionen auch die Lösung des approximierten Optimierungsproblems nah an der des exakten Problems liegt. Der Kerndichteschätzer ermöglicht die Berechnung der Ableitung der approximierten Wahrscheinlichkeitsrestriktion. Dies ermöglicht es uns, notwendige Optimalitätsbedingungen für die approximierten Optimierungsprobleme mit Wahrscheinlichkeitsrestriktionen herzuleiten.

Abstract

In this thesis we analyze stationary and dynamic gas flow with uncertain boundary data in networks of pipelines. The gas flow in pipeline networks is modeled by the isothermal Euler equations. The uncertain boundary data is modeled by probability distributions, they represent the a priori unknown gas demand of the consumers. The aim of this work is the analysis of optimization problems with probabilistic constraints in the context of gas transport.

For computing the probability that an uncertain gas demand is feasible we use both, a kernel density estimator approach and the spheric radial decomposition. Feasible in this context means, that the demanded gas can be transported through the network, s.t. bounds for the pressure at the nodes are satisfied. Moreover we discuss advantages and disadvantages of both methods. In the stationary case we extend our model by compressor control and bounds for the pressure at the entry nodes, and we also compute the probability for an uncertain gas demand to be feasible. In the dynamic setting the uncertain gas demand is time dependent, which is modeled by randomized Fourier series.

Further we analyze certain optimization problems with probabilistic constraints, in which the probabilistic constraints are approximated by the kernel density estimator approach. On the one hand we show the existence of optimal solutions for both, the exact and the approximated problems, and on the other hand we show that if the approximation is sufficiently accurate, then the optimal solutions of the approximated problems are close to the solutions of the exact problems. With the approximation of the probabilistic constraints via the kernel density estimator we are able to compute derivatives of the approximated optimization constraints, which allows us to derive necessary optimality conditions for the approximated optimization problems with probabilistic constraints.

Contents

Zusammenfassung	iv
Abstract	v
Danksagung	ix
Preface	x
Motivation and Introduction	xii
1. Mathematical Background	1
1.1. Kernel Density Estimation (KDE)	2
1.1.1. Univariate Kernel Density Estimation	2
1.1.2. Multivariate Kernel Density Estimator	4
1.1.3. Convergence of the Kernel Density Estimator	9
1.2. The Spheric Radial Decomposition (SRD)	15
2. Gas Flow in Pipes	19
2.1. The isothermal Euler equations	19
2.2. The isothermal Euler equations in terms of Riemann invariants	22
2.3. Literature survey	26
3. Probabilistic Constrained Optimization on Stationary Gas Net-	
works	30
3.1. Mathematical Modelling and Problem Description	31
3.2. Application of the SRD	36
3.3. Application of the KDE	41
3.4. Numerical Examples	46
3.4.1. A Single Edge	46
3.4.2. A Network with two parallel edges	48
3.4.3. A Network with two linear edges	51
3.5. A Model Extension: Unknown Inlet Pressure and Compressor	
Control	53
3.6. Optimization Problems with Probabilistic Constraints	72

3.7. Summary, Advantages and Disadvantages of the SRD and the KDE	89
4. Probabilistic Constrained Optimization on Dynamic Flow Networks	92
4.1. Time dependent Probabilistic Constraints	92
4.2. Uncertain Time dependent Functions	94
4.3. Mathematical Modelling and Problem Description	101
4.3.1. Model Derivation	101
4.3.2. The Model in Terms of Riemann Invariants	103
4.3.3. Problem Description	108
4.4. Application of the KDE	109
4.5. Optimization Problems with Probabilistic Constraints	115
4.6. A Numerical Example	119
Summary, Conclusion and Perspectives	126
A. Appendix	128
A.1. Comparison of radial symmetric and product kernel functions .	128
A.2. Solution of ISO1 on one edge	130
Bibliography	132

Danksagung

Zu Beginn der Dissertation möchte ich mich bei einigen Freunden*innen und Kollegen*innen bedanken, die in den letzten Jahren eine zentrale Rolle in meinem Leben gespielt haben.

Ein besonderer Dank geht an mein Betreuer, Martin Gugat, ohne den diese Arbeit nicht existieren würde. Er hat mich seit meiner Masterarbeit stets unterstützt, sowohl mit detaillierten Hinweisen und hilfreichen Diskussionen, als auch mit seiner Zeit und Geduld. Außerdem hätte ich viele meiner Kollegen*innen und Freunde*innen ohne ihn wahrscheinlich gar nicht kennen gelernt. Vielen Dank dafür!

Ich möchte mich auch bei meinem Mentor Rüdiger Schultz bedanken. Er hatte immer ein offenes Ohr für mich, sowohl für Mathematisches als auch für Privates. Er hat mich die letzten Jahre nicht nur immer unterstützt, sondern allem auch etwas mehr Humor gegeben.

Darüber hinaus möchte ich noch Günther Leugering, Michael Stingl, Enrique Zuazua und ihren Arbeitsgruppen für die vielen hilfreichen Diskussionen und das angenehme Arbeitsumfeld danken. Insbesondere bedanke ich mich bei Yan und Elisa für die angenehme Zusammenarbeit, sie waren maßgeblich am guten Arbeitsklima beteiligt.

Danke auch an meine Freunde, meine Band und meine Leichtathleten*innen, die sich notgedrungen meine mathematischen Probleme anhören mussten (wahrscheinlich öfter als sie es wollten). Mein „betreutes Lernen“ (Andi, Domi) hat mich bis hierher gebracht. Danke auch an Bernd und das Flatted für mein zweites „Zuhause“. Mit euch ist es nie langweilig!

Zum Schluss möchte ich natürlich noch meiner Familie und meiner Freundin danken. Mama, Papa, Chrisi und Yvi, ihr habt mich immer unterstützt in dem, was ich tue und tun wollte. Ihr wart und seid immer für mich da! Ich bin glücklich darüber, dass ihr ein Teil meines Lebens seid. **Danke!**

Preface

Before we dive in the depth of mathematics I want to thank some friends and colleagues who played a outstanding role in my life for the last few years.

Special thanks go to my adviser, Martin Gugat, without whom this work would not exist. Since I wrote my master thesis he supported me with detailed hints, helpful discussions but also with his time and patience. Furthermore I probably never would have met some of my colleagues and friends. Thank you for everything!

I also want to thank my mentor Rüdiger Schultz. He always payed attention to both, mathematical problems and private issues. Not only that he always supported me, he also was never too good for a joke.

Next I want to thank Günther Leugering, Michael Stingl, Enrique Zuazua and their research groups for the helpful discussions and the comfortable atmosphere in the last years. Especially I want to thank Yan and Elisa for the pleasant collaboration, they basically were part of the comfortable atmosphere.

Thanks to all of my friends, my band and my athletes, who carefully listened to my mathematical problems (probably more often than they wanted). The „supervised studying“ (Andi, Domi) brought me where I am now. Thanks also to Bernd and the Flatted Fifth, you gave me a „second home“. Without you a lot more would be humdrum everyday.

Finally, of course, I want to thank my family and my girlfriend. Mom, Dad, Chrisi and Yvi, you always support me in whatever I am doing and whatever I was doing. You've always been there for me when I needed you! I'm beyond grateful for having you in my life! **Thank you!**

This thesis arose during my time in the CRC¹ Transregio 154: *Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks*, subproject C03: *Nodal Control and the Turnpike Phenomenon*. I want to thank the German Research Foundation DFG² for funding my PhD position.

This thesis is mainly based on the following articles, which form the core for the content of this thesis:



M. Schuster, E. Strauch, M. Gugat, J. Lang:
Probabilistic Constrained Optimization on Flow Networks.
Optim. Eng. 2021, <https://doi.org/10.1007/s11081-021-09619-x>.



M. Gugat, M. Schuster:
Stationary Gas Networks with Compressor Control and Random Loads: Optimization with Probabilistic Constraints.
Math. Prob. Eng., vol 2018, Article-ID 9784079, 17 pages, 2018.

The following articles that have a minor contribution to this thesis empowered a lot of new ideas, broadened my horizon and introduced me to different fields of mathematics:



M. Gugat, M. Schuster, R. Schultz:
Convexity and Starshapedness of Feasible Sets in Stationary Flow Networks.
Netw. Heterog. Media 15(2), 2020, pp. 171-195.



M. Gugat, M. Schuster, E. Zuazua:
The Finite-Time Turnpike Phenomenon for Optimal Control Problems: Stabilization by Non-Smooth Tracking Terms.
Stabilization of Distributed Parameter Systems: Design Methods and Applications 2021, pp. 17-41. Springer.

¹Collaborative Research Center

²Deutsche Forschungsgemeinschaft

Motivation and Introduction

At least since the disaster of Fukushima in 2011 the discussion about the nuclear power phase-out is more current than ever. Also due to the aim to reduce the number of coal-fired power plants, natural gas as energy source becomes more important. Of course the long term goal is a complete change to renewable sources, but until this is possible, energy based on natural gas is a wise interim solution. Also pipelines for natural gas, because of its history, are available in many German resp. European countries and cities. As well gas is used for heating and hot water in many German houses, e.g. in 2014 almost half of the German households used natural gas for heating¹. A scheme of the German pipeline network is shown in *Figure 0.1*. In the future, when the pipelines are no longer needed for gas transport, they can be used for the transport of other gases like hydrogen, e.g. with respect to cars with fuel cells.

Inter-city resp. inter-country pipelines are important because Germany and most of the European countries produce less gas than they consume. Inner-city pipelines are the main method to deliver gas to the households. In 2005, the *Federal Network Agency*² was founded with the aim of liberalization of the German gas market. Since 2006 German households can choose the gas provider of their choice³. The European Union supported and still supports the liberalization of gas and energy market and essentially contributed to the separation of energy production, energy supply and energy transportation⁴. Further the Federal Network Agency demands, that every declared nomination⁵ must be transportable in the network at any time.

One of the most up-to-date gas projects in Germany is the *North Stream 2* pipeline, which is currently built (see *Figure 0.2*). The North Stream 2

¹Federal Ministry of Economic Affairs and Energy: <https://www.bmwi-energiewende.de/EWD/Redaktion/EN/Newsletter/2015/09/Meldung/infografik-heizsysteme.html>, last checked: 20.04.2021

²Bundesnetzagentur, BNetzA

³https://www.gesetze-im-internet.de/enwg_2005/EnWG.pdf, last checked: 20.04.2021

⁴<http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009L0073&from=EN>, last checked, 20.04.2021

⁵A nomination is the demanded (in common with the agreement) gas at a moment

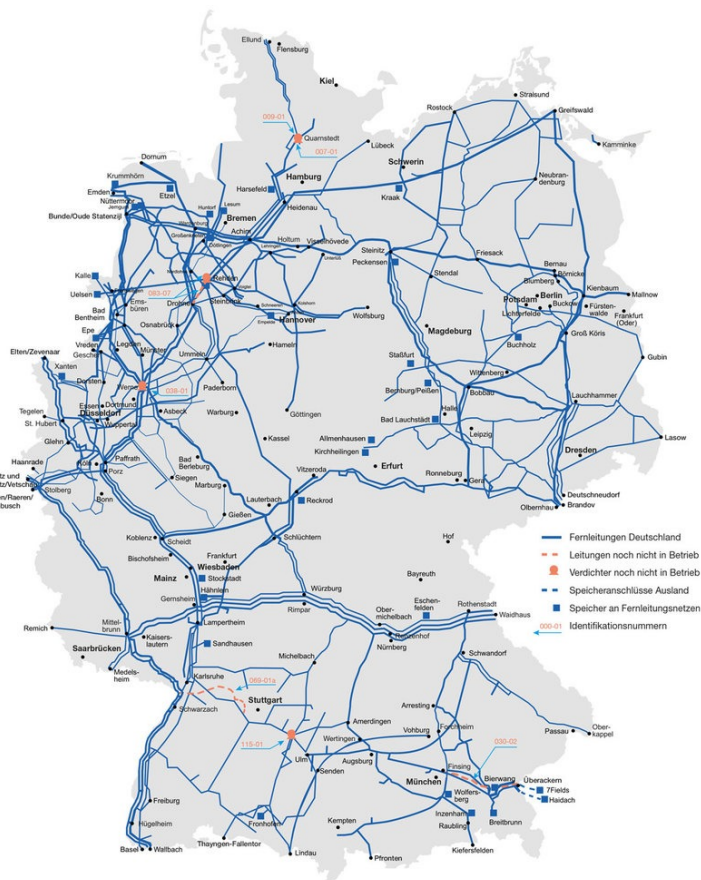


Figure 0.1.: German pipeline network⁶
 ©BMWi, Fernleitungsnetzbetreiber

pipeline should transport natural gas through the Baltic see for stabilizing the German gas market and guaranteeing enough gas in German pipes and tanks. This is discussed frequently in politics. So it gets a fortiori important to understand the process of gas transport in pipeline networks in all fields.

Since gas transport is an up-to-date topic, many scientists (engineers and physicists as well as mathematicians) study the modelling, simulation and optimization of gas networks and develop different models to deal with the problem of gas transport. From a mathematical point of view, [Koch et al., 2015, Chapter 2] gives a great overview about the fundamentals of gas transport in pipeline networks and [Domschke et al., 2017] gives a great overview

⁶<https://www.bmwi.de/Redaktion/DE/Artikel/Energie/gas-erdgasversorgung-in-deutschland.html>, last checked: 20.04.2021



Figure 0.2.: Process of building the Nord Stream 2 pipeline⁷
© Nord Stream 2, Axel Schmidt

about existing models with their areas of application. For mathematicians gas transport in pipeline networks is quite interesting because it combines many fields of mathematics, e.g. the theory of partial differential equations including existence and uniqueness of solutions, optimization with partial differential equations, control theory, mixed integer programming, uncertainty, robustness and much more. In this thesis we focus on the isothermal Euler equations on networks, which model the gas transport through the pipelines, combined with uncertain boundary data and controllable elements like compressor stations. This thesis is structured as follows:

In *Chapter 1* we give a basic introduction into *kernel density estimation* and the *spheric radial decomposition*. Both tools are used in this thesis to compute probabilities and to analyze probabilistic constraints.

In *Chapter 2* we introduce the isothermal Euler equations, which model the physics of gas flow in pipelines with constant temperature. We write the isothermal Euler equations in terms of Riemann invariants and give an overview about the existing literature on gas flow in pipeline networks and related topics.

In *Chapter 3* we introduce a stationary mathematical model including the

⁷<https://www.nord-stream2.com/media-info/images/connecting-pipe-sections-above-water-800/?query=&category=&language=&country=Germany&year=>, last checked: 20.04.2021

gas network, the stationary isothermal Euler equations and coupling conditions at the network junctions. We assume that the gas outflows of the network, i.e. the loads, are uncertain with a given probability distribution. We define the set of feasible loads as the outflows that can be transported through the network and satisfy pressure bounds at the network junctions. Further we develop two different methods (based on the kernel density estimation and the spheric radial decomposition which are introduced in *Chapter 1*) to compute the probability whether a random load vector is feasible. We extend the model by compressor stations and an unknown inlet pressure and again answer the question whether a random load vector is feasible. Then we consider certain optimization problems with probabilistic constraints in the context of gas transport and we derive necessary optimality conditions for these problems using kernel density estimation. We prove the existence of optimal solutions under weak assumptions and state the key theorem of this chapter, whose gist consists in the accuracy of the solutions of the approximated problems with respect to the solutions of the exact problems. This is completely new in optimization with probabilistic constraints in the context of gas transport and one of the key results of this thesis. In the end of this chapter we discuss the advantages and disadvantages of both methods for computing the probability whether a random load vector is feasible.

Chapter 4 deals with the dynamics of gas transport and uncertain boundary data. We first discuss time dependent probabilistic constraints and introduce the recently developed **probust** (**probabilistic** and **robust**) constraints. Further we discuss time dependent uncertainty and introduce random Fourier series. Then (similar to *Chapter 3*) we introduce a dynamic mathematical model including the gas network, the isothermal Euler equations (written in terms of Riemann invariants) and coupling conditions at the network junctions. We assume that the time dependent gas outflow is uncertain. We define the set of feasible loads as all outflows that satisfy the dynamic model with additional box constraints for the pressure. We apply the kernel density estimation approach (that is introduced in *Chapter 1*) to compute the probability whether a random outflow is feasible for all times. We consider an optimization problem with a probust constraint and use the kernel density estimator approach to approximate the probability in the constraint. This leads to the fact that we can use the theory of the stationary case (developed in *Chapter 3*) in the dynamic context as well. The application of a kernel density estimator approach in the context of dynamic gas transport has not been used yet and is a novelty in the mathematics of gas transport resp. optimization with probabilistic constraints in the context of gas transport.

*"Mathematics is the most beautiful and most powerful
creation of the human spirit."*

- Stefan Banach -

1. Mathematical Background

In this chapter we introduce two mathematical techniques that are highly used in this thesis. The first is the *kernel density estimator*, which provides a method for estimating unknown probability density functions. The kernel density estimator is introduced in *Section 1.1* for both, one dimension and higher dimensions. Further we state some convergence results for the kernel density estimator in *Section 1.1.3*, adapted to our setting. In *Section 1.2* we introduce the *spheric radial decomposition*, which is an integral transformation that we will use to avoid the computation of high dimensional integrals.

Let an objective function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, an inequality constraint $g : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^m$ and a probability level $\alpha \in (0, 1)$ be given. Further let a Gaussian distributed random variable $\xi \sim \mathcal{N}(\mu, \Sigma)$ with mean value $\mu \in \mathbb{R}^p$ and positive definite covariance matrix $\Sigma \in \mathbb{R}^{(p \times p)}$ be given on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. In general, we consider optimization problems of the form

$$\begin{cases} \min_{x \in \mathbb{R}^n} & f(x), \\ \text{s.t.} & \mathbb{P}(\omega \in \Omega \mid g_i(x, \xi(\omega)) \leq 0 \quad \forall i = 1, \dots, m) \geq \alpha. \end{cases} \quad (1.1)$$

This optimization problem means that we minimize the objective function with respect to the decision variable $x \in \mathbb{R}^n$, s.t. the probability, that the constraints are satisfied, is at least α . The problem (1.1) is a probabilistic constrained optimization problem. For a general introduction to probabilistic constraints and optimization problems with probabilistic constraints we refer to [Prékopa, 1995], which gives a great introduction and overview about this topic. For $\omega \in \Omega$, we define the set

$$M(x) := \{ \xi(\omega) \in \mathbb{R}^p \mid g_i(x, \xi(\omega)) \leq 0 \quad \forall i = 1, \dots, m \}. \quad (1.2)$$

Thus the probabilistic constraint of (1.1) can be equivalently written as

$$\mathbb{P}(\omega \in \Omega \mid \xi(\omega) \in M(x)).$$

We identify $\zeta = \xi(\omega)$ with a realization of the random variable ξ . For simplicity we write

$$\mathbb{P}(\zeta \in M(x)).$$

The main task in this chapter is the computation of this probability. In general, the probability can be computed by integrating the exact probability density ρ_ξ of the random variable ξ over the set $M(x)$, i.e. for $x \in \mathbb{R}^n$ we have

$$\mathbb{P}(\zeta \in M(x)) = \int_{M(x)} \rho_\xi(z) dz. \quad (1.3)$$

This can be very difficult since e.g. the set $M(x)$ can be high dimensional and have an arbitrary non convex and non smooth shape or the exact probability density function ρ_ξ can be unknown. In the next subsections, we introduce a *kernel density estimator* approach to estimate unknown probability density functions and we introduce the *spheric radial decomposition* which we use to reduce the dimension of integration.

1.1. Kernel Density Estimation (KDE)

Kernel density estimation is a non-parametric way to estimate the (unknown) probability density function of a random variable using given statistical data. Kernel density estimation as we use it in this thesis was mainly developed in the early 60s (see e.g. [Parzen, 1962]). In this section, we introduce a univariate and a multivariate kernel density estimator that we will apply to our setting in the next chapters. From now on we write KDE instead of kernel density estimator. In general [Gramacki, 2018] gives a very good overview and detailed introduction on this topic.

Definition 1.1. A kernel function is a integrable, borel measurable function $K : \mathbb{R}^n \rightarrow [0, \infty)$, s.t.

$$\int_{\mathbb{R}^n} K d\mu = 1,$$

with respect to the Lebesgue measure μ .

These kernel functions are used to smooth the empirical measure. The idea is closely related to the idea of Dirac sequences in the analysis. In this section we introduce both, a one-dimensional (univariate) KDE and a n -dimensional (multivariate) KDE.

1.1.1. Univariate Kernel Density Estimation

In this subsection, we introduce the univariate KDE to approximate one-dimensional probability density functions. Let Y be a one-dimensional random variable on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We assume that

Y has an absolutely continuous distribution function with probability density function ϱ . This assumption is necessary since a probability density function exists if and only if the distribution function is absolutely continuous with respect to the Lebesgue measure μ . We introduce the univariate KDE as follows:

Definition 1.2. Let $\mathcal{Y} = \{y_1, \dots, y_N\} \subseteq \mathbb{R}$ be independent and identically distributed samples of the random variable Y , which has an absolutely continuous distribution function with probability density function ϱ . Let K be a one-dimensional kernel function. Then, the univariate kernel density estimator ϱ_N corresponding to the bandwidth $h \in (0, \infty)$ is defined as

$$\varrho_N(z) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{z-y_i}{h}\right). \quad (1.4)$$

The KDE is basically a weighted sum of single densities. Since K is a kernel function, we have $\varrho_N \geq 0$. We use the transformation

$$\xi_i := \left(\frac{z-y_i}{h}\right),$$

which leads to

$$\begin{aligned} \int_{-\infty}^{\infty} \varrho_N(z) dz &= \int_{-\infty}^{\infty} \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{z-y_i}{h}\right) dz \\ &= \frac{1}{Nh} \sum_{i=1}^N h \underbrace{\int_{-\infty}^{\infty} K(\xi_i) d\xi_i}_{=1} \\ &= 1. \end{aligned}$$

Thus ϱ_N is a probability density function. A choice of common kernel functions is given in *Table 1.1* and the corresponding graphs are shown in *Figure 1.1*. For more kernels, see e.g. [Langrené and Warin, 2019] or [Gramacki, 2018].

As one can see in *Table 1.1*, some of the kernel functions have a compact support, others do not. Since our sample is Gaussian distributed later, we will choose the Gaussian kernel. Even though the Gaussian kernel has no compact support, it has good smoothing properties. The optimal choice of the bandwidth $h \in (0, \infty)$ is a separate topic. In *Figure 1.2*, an estimated density of 25 (standard) Gaussian distributed points is shown for different bandwidths. One

Uniform:	$K(x) = \begin{cases} \frac{1}{2} & x \leq 1 \\ 0 & \text{else} \end{cases}$
Epanechnikov:	$K(x) = \begin{cases} \frac{3}{4}(1-x^2) & x \leq 1 \\ 0 & \text{else} \end{cases}$
Gaussian:	$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$
Cauchy:	$K(x) = \frac{1}{\pi(1+x^2)}$

Table 1.1.: Example of univariate kernel Functions

can see, that if h gets larger, the estimated density gets smoother. Roughly speaking for small h one can almost see the single terms of the sum (1.4), while for large h the full curve is close to a classical Gaussian distribution. In general it is difficult to find the optimal bandwidth $h \in (0, \infty)$, for an analysis of the optimal bandwidth we refer to [Duller, 2018, Chapter 8] and [Härdle et al., 2004, Chapter 3]. Here we use a heuristic formula depending on the standard deviation σ_N of a sampling of length N given by

$$h = 1.06 \frac{\sigma_N}{\sqrt[5]{N}}. \quad (1.5)$$

This choice of the bandwidth is often proposed in the literature (see e.g. [Gramacki, 2018, Chapter 4] and [Silverman, 1986, Chapter 3]). The term coincides with the choice of bandwidth matrix in the multivariate case as we will see in the next section. The equation (1.5) guarantees \mathbb{P} -almost surely convergence for the KDE as we also see later. In *Figure 1.2* the value $h = 0.48512$ was computed with (1.5).

1.1.2. Multivariate Kernel Density Estimator

We introduce the multivariate KDE in this section. Our aim is to approximate a n -dimensional probability density function. Let Y be a n -dimensional random variable on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Equally to the last subsection, we assume that Y has an absolutely continuous probability distribution function with probability density function ϱ . A probability density function of Y exists if and only if the distribution function of Y is absolutely continuous.

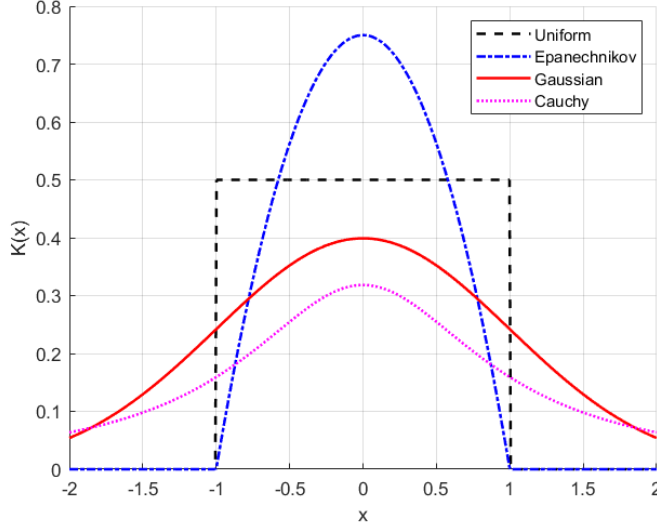


Figure 1.1.: Graphs of the kernel functions given in *Table 1.1*

Definition 1.3. Let $\mathcal{Y} = \{y_1, \dots, y_N\} \subseteq \mathbb{R}^n$ be independent and identically distributed samples of the random variable Y , which has an absolutely continuous distribution function with probability density ϱ . Let K be a n -dimensional kernel function. Then, the multivariate kernel density estimator ϱ_N corresponding to the symmetric positive definite bandwidth matrix $H \in \mathbb{R}^{n \times n}$ is defined as

$$\varrho_N(z) = \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N K\left(H^{-\frac{1}{2}}(z - y_i)\right). \quad (1.6)$$

In general, it is not clear that $H^{\frac{1}{2}}$ exists, so H has to be chosen suitably, e.g. symmetric positive definite.

We have $\varrho_N(z) \geq 0$ and with

$$\xi_i := \left(H^{-\frac{1}{2}}(z - y_i)\right),$$

and it follows

$$\begin{aligned} \int_{\mathbb{R}^n} \varrho_N(z) dz &= \int_{\mathbb{R}^n} \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N K\left(H^{-\frac{1}{2}}(z - y_i)\right) dz \\ &= \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N \det\left(H^{-\frac{1}{2}}\right) \underbrace{\int_{\mathbb{R}^n} K(\xi_i) d\xi_i}_{=1} \\ &= 1. \end{aligned}$$

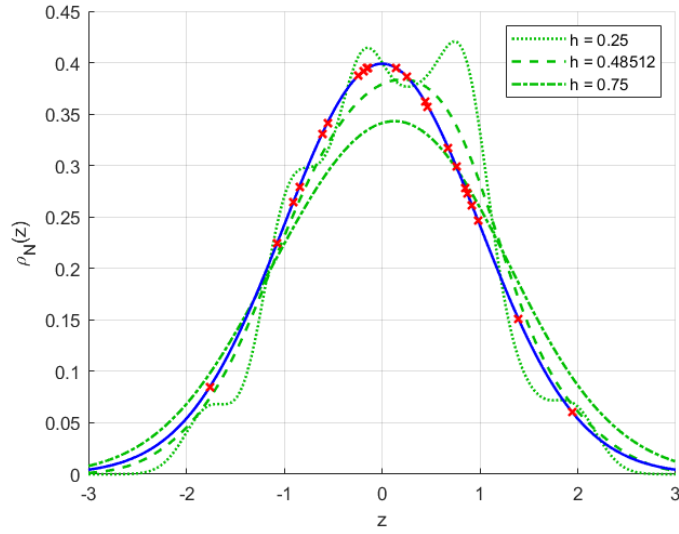


Figure 1.2.: Estimated standard Gaussian distribution for 25 sample points

Thus, $\varrho_N(z)$ is a probability density. The multivariate KDE is quite similar to the univariate KDE. One can extend the one-dimensional kernels to n -dimensional kernels in different ways (see e.g. [Langrené and Warin, 2019], [Härdle and Müller, 1997] or [Gramacki, 2018]). A selection of n -dimensional radial symmetric kernel functions are shown in *Table 1.2*. The corresponding graphs in two dimensions are shown in *Figure 1.3*.

For the multivariate KDE (1.6) we have

$$\varrho_N(z) = \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N K \left(\begin{bmatrix} \sum_{k=1}^n (H^{-\frac{1}{2}})_{1,k} (z_k - y_{i,k}) \\ \vdots \\ \sum_{k=1}^n (H^{-\frac{1}{2}})_{n,k} (z_k - y_{i,k}) \end{bmatrix} \right),$$

where $y_{i,k}$ is the k -th component of y_i . Another typical choice of a multivariate kernel functions K is a product kernel

$$K(z) = \prod_{j=1}^n \mathcal{K}(z_j),$$

with univariate kernel functions $\mathcal{K} : \mathbb{R} \rightarrow \mathbb{R}$ (see e.g. [Langrené and Warin, 2019] or [Gramacki, 2018]). To be more general, one could use different univariate kernel functions for every dimension, but this would blow up the difficulty to analytically work with the KDE, so we use the same univariate

Uniform:	$K(x) = \begin{cases} \frac{1}{\pi} & \ x\ \leq 1 \\ 0 & \text{else} \end{cases}$
Epanechnikov:	$K(x) = \begin{cases} \frac{\Gamma(2+\frac{n}{2})}{\pi^{\frac{n}{2}}}(1 - \ x\ ^2) & \ x\ \leq 1 \\ 0 & \text{else} \end{cases}$
Gaussian:	$K(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp(-\frac{1}{2}\ x\ ^2)$
Cauchy:	$K(x) = \frac{\Gamma(\frac{1+n}{2})}{\Gamma(\frac{1}{2})\pi^{\frac{n}{2}}(1+\ x\ ^2)}$

Table 1.2.: Example of radial symmetric kernel functions

kernel function in every dimension. Also for this case the convergence theory works as we see in the next subsection. For the multivariate KDE with product kernel we have

$$\varrho_N(z) = \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N \prod_{j=1}^n \mathcal{K} \left(\sum_{k=1}^n (H^{-\frac{1}{2}})_{j,k} (z_k - y_{i,k}) \right).$$

A selection of n -dimensional product kernels is shown in *Table 1.3* and the corresponding graphs are shown in *Figure 1.4*. All radial symmetric kernels and the corresponding product kernels listed in this section are juxtaposed in *Figure A.1* and *Figure A.2* in the *Appendix A.1*.

The next challenge is to choose a suitable bandwidth matrix $H \in \mathbb{R}^{n \times n}$. As it is mentioned in the last subsection, one can use a whole thesis discussing how to choose a good or even the best bandwidth matrix, but this topic is not of main interest here. We refer to [Gramacki, 2018, Chapter 4] and [Silverman, 1986, Chapter 4] for a detailed analysis of this topic. For our analysis in the next sections, it is essential that the bandwidth matrix is diagonal. Since the convergence theory also holds for suitably chosen diagonal bandwidth matrices, this is not a strong restriction. In [Silverman, 1986] and in [Gramacki, 2018], the authors suggest choosing either a dense bandwidth matrix $H = h^2 \Sigma$ or a diagonal bandwidth matrix $H_{i,i} = h^2 \Sigma_{i,i}$ ($i = 1, \dots, n$) with

$$h = \left(\frac{4}{(n+2)N} \right)^{\frac{1}{n+4}}. \quad (1.7)$$

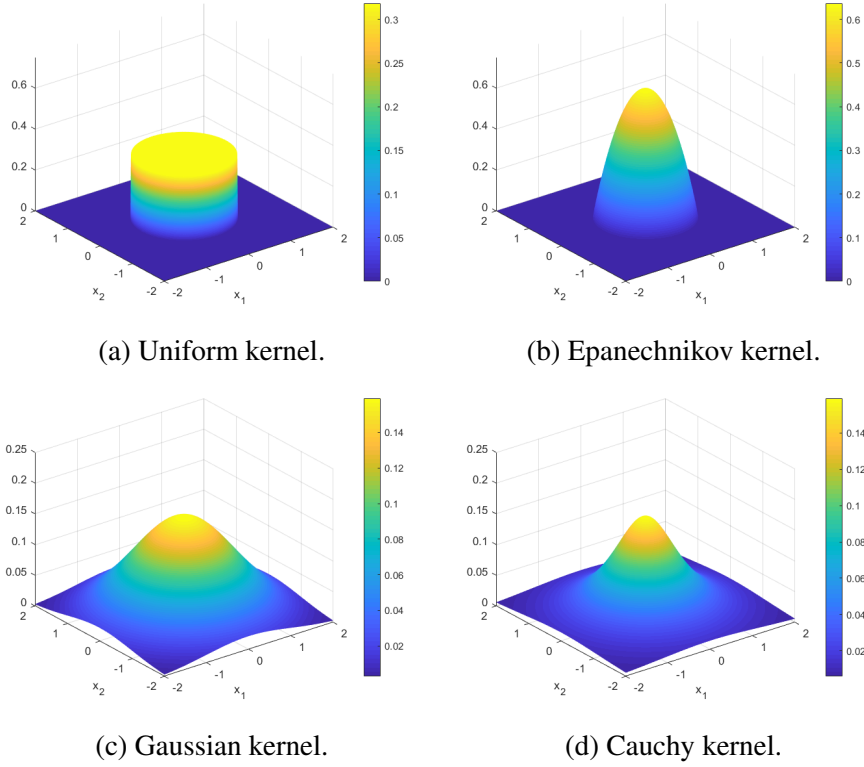


Figure 1.3.: Graphs of the kernel functions given in *Table 1.2* in 2D

Here Σ is the covariance matrix of the sampling. As mentioned before it is essential for our later analysis that the bandwidth matrix is diagonal, hence we use the diagonal form $H_{i,i} = h^2 \Sigma_{i,i}$ ($i = 1, \dots, n$) with h as in (1.7) in the next chapters. Another advantage of the diagonal form is, that $H^{\frac{1}{2}}$ always exists since $\Sigma_{i,i} > 0$ for all $i = 1, \dots, n$. With a diagonal bandwidth matrix, for the KDE with a product kernel we have

$$\varrho_N(z) = \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N \prod_{j=1}^n \mathcal{K} \left(\frac{z_j - y_{i,j}}{\sqrt{H_{j,j}}} \right).$$

This is basically the multivariate KDE that we use for our analysis later.

Remark 1.4. For a diagonal bandwidth matrix $H_{i,i} = h^2 \Sigma_{i,i}$ with h as in (1.7), the radial symmetric Gaussian kernel stated in *Table 1.2* and the Gaussian product kernel stated in *Table 1.3* are equal.

Uniform:	$K(x) = \begin{cases} \frac{1}{4} & x_i \leq 1 \quad \forall i = 1, \dots, n \\ 0 & \text{else} \end{cases}$
Epanechnikov:	$K(x) = \begin{cases} \left(\frac{3}{4}\right)^n \prod_{i=1}^n (1 - x_i^2) & x \leq 1 \quad \forall i = 1, \dots, n \\ 0 & \text{else} \end{cases}$
Gaussian:	$K(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \prod_{i=1}^n \exp\left(-\frac{1}{2}x_i^2\right)$
Cauchy:	$K(x) = \frac{1}{\pi^n} \frac{1}{\prod_{i=1}^n (1+x_i^2)}$

Table 1.3.: Example of product kernel functions

Remark 1.5. For one dimension (i.e. $n = 1$), it holds

$$h = \left(\frac{4}{3N}\right)^{\frac{1}{5}} \approx 1.06 \frac{1}{\sqrt[5]{N}}.$$

Then $H^{\frac{1}{2}} = h\Sigma^{\frac{1}{2}}$ is equal to (1.5).

In the next section we show, that both, the univariate KDE and the multivariate KDE with certain bandwidths \mathbb{P} -almost surely converge to the exact probability density of the data.

1.1.3. Convergence of the Kernel Density Estimator

In this subsection, we state some convergence results on the kernel density estimator which can be applied to our setting. Let $\mathcal{Y} = \{y_1, \dots, y_N\} \subseteq \mathbb{R}^n$ be an independent and identically distributed sample of the random variable Y , which has an absolutely continuous distribution function with probability density ϱ . Further let K be a kernel function and $h = h(N, y_1, \dots, y_N)$ be a borel measurable function of N and the samples.

An important convergence result is stated in [Devroye and Gyorf, 1985, Chapter 6, Theorem 1]. The authors give a general L^1 convergence result for this setting. We adapt the theorem to our notation (note that they use n for the number of samples and d for the dimension, we use N for the number of samples and n for the dimension).

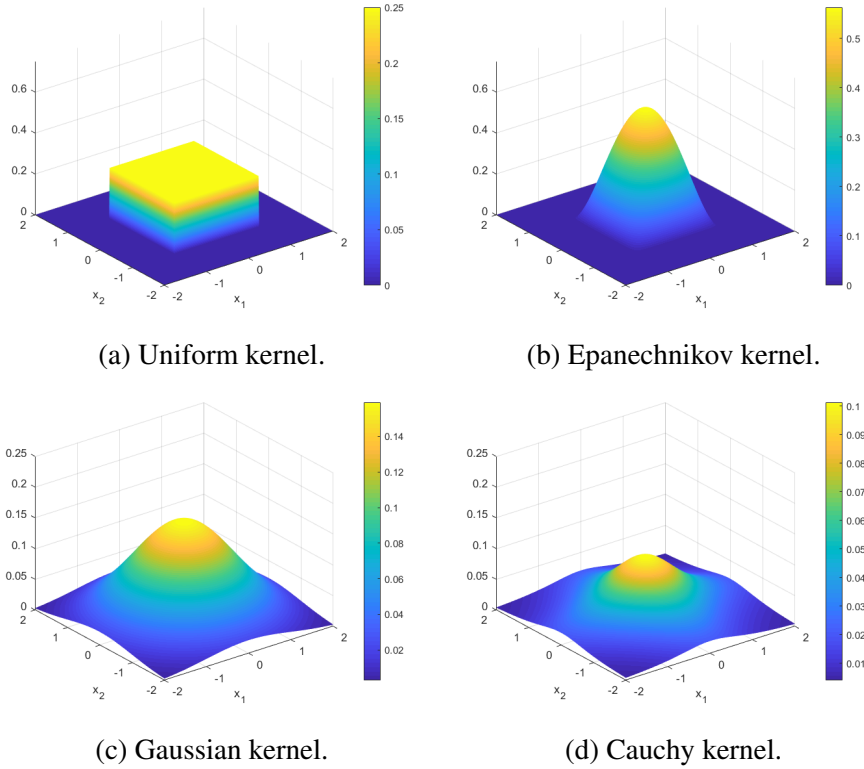


Figure 1.4.: Graphs of the kernel functions given in *Table 1.3* in 2D

Theorem 1.6. Let ϱ_N be an automatic kernel estimate, i.e. ϱ_N is given by

$$\varrho_N(z) = (N h^n)^{-1} \sum_{i=1}^N K\left(\frac{z - y_i}{h}\right),$$

with arbitrary kernel function K . If $h + (N h^n)^{-1} \rightarrow 0$ for $N \rightarrow \infty$ \mathbb{P} -almost surely, then

$$\int_{\mathbb{R}^n} |\varrho_N(z) - \varrho(z)| \xrightarrow{N \rightarrow \infty} 0,$$

\mathbb{P} -almost surely.

The proof of this theorem can be found in [Devroye and Györfi, 1985, Chapter 6]. Since the data and the random variable are n -dimensional, the theorem holds for both, the univariate KDE and the multivariate KDE, which is an enormous benefit. To apply this to our setting we need to guarantee that the assumption in *Theorem 1.6* holds.

Lemma 1.7. Consider the univariate kernel density estimator (1.4). For h defined in (1.5) we have

$$h + (N h)^{-1} \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P} - \text{almost surely.}$$

For the convenience of the reader we present the proof of *Lemma 1.7* even though it is a special case of *Lemma 1.9*.

Proof of Lemma 1.7. For h as in (1.5), we have

$$\begin{aligned} \lim_{N \rightarrow \infty} h + (N h)^{-1} &= \lim_{N \rightarrow \infty} 1,06 \frac{\sigma_N}{N^{\frac{1}{5}}} + \lim_{N \rightarrow \infty} \left(N 1,06 \frac{\sigma_N}{N^{\frac{1}{5}}} \right)^{-1} \\ &= \lim_{N \rightarrow \infty} 1,06 \sigma_N N^{-\frac{1}{5}} + \lim_{N \rightarrow \infty} \left(1,06 \sigma_N N^{\frac{4}{5}} \right)^{-1} \\ &= \lim_{N \rightarrow \infty} 1,06 \sigma_N N^{-\frac{1}{5}} + \lim_{N \rightarrow \infty} (1,06 \sigma_N)^{-1} N^{-\frac{4}{5}} \end{aligned}$$

Since σ_N is the standard deviation of the sampling, it is

$$\lim_{N \rightarrow \infty} \sigma_N = \sigma \quad \mathbb{P} - \text{almost surely,}$$

where $\sigma > 0$ is the standard deviation of the random variable Y . Thus, we have

$$\lim_{N \rightarrow \infty} h + (N h)^{-1} = 0 \quad \mathbb{P} - \text{almost surely.}$$

□

So we can apply *Theorem 1.6* to our setting and the univariate kernel density estimator (1.4) with bandwidth (1.5) L^1 -converges \mathbb{P} -almost surely to the probability density function ϱ of Y , i.e.

$$\| \varrho_N - \varrho \|_{L^1} \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P} - \text{almost surely.}$$

The convergence of the multivariate KDE does not follow directly. *Theorem 1.6* only holds for constant bandwidth h in every dimension, e.g. a bandwidth matrix $H = h^2 \mathbb{E}_{n \times n}$ where $\mathbb{E}_{n \times n}$ is the unit matrix. To be able to use *Theorem 1.6* for general dense bandwidth matrices and diagonal matrices like (1.7), [Gramacki, 2018] and [Silverman, 1986] provide a transformation algorithm which consists in the following three steps:

Algorithm 1.8.

1. Apply the Whitening transformation: $X = \Sigma^{-\frac{1}{2}} Y$
2. Smooth the pre-scaled data: $\varrho_N^W(\xi) = \frac{1}{Nh^n} \sum_{i=1}^N K\left(\frac{\xi - x_i}{h}\right)$
3. Transform back: $\varrho_N(\xi) = \varrho_N^W(z) \det(\Sigma)^{-\frac{1}{2}}$

Step 1 in Algorithm 1.8 consists in applying the *Whitening* transformation to the random variable resp. to the sampled data. The *Whitening* transformation is a linear transformation which transforms a n -dimensional random variable with arbitrary covariance matrix into another n -dimensional random variable whose covariance matrix is the unit matrix (for more details see [Kessy et al., 2018]). Let Σ be the covariance matrix of the sampling $\mathcal{Y} = \{y_1, \dots, y_N\}$. With the Whitening transformation we get a sampling $\mathcal{X} = \{x_1, \dots, x_N\}$ with $x_i = \Sigma^{-\frac{1}{2}} y_i$ and covariance matrix $\mathbb{E}_{n \times n}$.

Then in Algorithm 1.8, step 2, we estimate the probability density of the sampling \mathcal{X} using the multivariate KDE (1.6) with the bandwidth matrix $H = h^2 \mathbb{E}_{n \times n}$ with h defined in (1.7), that is

$$\varrho_N^W(\xi) = \frac{1}{Nh^n} \sum_{i=1}^N K\left(\frac{\xi - x_i}{h}\right). \quad (1.8)$$

Last, step 3 of Algorithm 1.8 states the inverse whitening transformation. Hence we have

$$\begin{aligned} \varrho_N(z) &= \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N K\left(H^{-\frac{1}{2}}(z - y_i)\right) \\ &= \frac{1}{Nh^n \det(\Sigma)^{\frac{1}{2}}} \sum_{i=1}^N K\left(\Sigma^{-\frac{1}{2}} \left(\frac{z - y_i}{h}\right)\right) && (H=h^2\Sigma) \\ &= \frac{1}{Nh^n \det(\Sigma)^{\frac{1}{2}}} \sum_{i=1}^N K\left(\frac{\xi - x_i}{h}\right) && (\text{Whitening}) \\ &= \frac{1}{\det(\Sigma)^{\frac{1}{2}}} \varrho_N^W(\xi). \end{aligned}$$

We need to guarantee, that $\Sigma^{-\frac{1}{2}}$ exists. We will later always assume, that the random variable has a symmetric positive definite covariance matrix. Thus, the covariance matrix is invertible, the whitening transformation is bijective and we can apply Algorithm 1.8 to the multivariate KDE stated in (1.6) with dense bandwidth matrix $H = h^2 \Sigma$.

Lemma 1.9. Consider the transformed multivariate KDE (1.8) with bandwidth matrix $H = h^2 \mathbb{E}_{n \times n}$. For h as in (1.7) we have

$$h + (N h)^{-1} \xrightarrow{N \rightarrow \infty} 0.$$

Proof. For h as in (1.7) we have

$$\begin{aligned} \lim_{N \rightarrow \infty} h + (N h)^{-1} &= \lim_{N \rightarrow \infty} \left(\frac{4}{(n+2)N} \right)^{\frac{1}{n+4}} + \lim_{N \rightarrow \infty} N^{-1} \left(\frac{4}{(n+2)N} \right)^{-\frac{1}{n+4}} \\ &= \left(\frac{4}{n+2} \right)^{\frac{1}{n+4}} \lim_{N \rightarrow \infty} N^{-\frac{1}{n+4}} + \left(\frac{4}{n+2} \right)^{-\frac{1}{n+4}} \lim_{N \rightarrow \infty} N^{-1+\frac{1}{n+4}} \\ &= \left(\frac{4}{n+2} \right)^{\frac{1}{n+4}} \lim_{N \rightarrow \infty} N^{-\frac{1}{n+4}} + \left(\frac{4}{n+2} \right)^{-\frac{1}{n+4}} \lim_{N \rightarrow \infty} N^{-\frac{n+3}{n+4}}. \end{aligned}$$

Since $n \geq 1$ it follows

$$\lim_{N \rightarrow \infty} h + (N h)^{-1} = 0.$$

□

So we can apply *Theorem 1.6* to the KDE (1.8) with bandwidth matrix $H = h^2 \mathbb{E}_{n \times n}$, hence we have

$$\| \varrho_N^W - \varrho^W \|_{L^1} \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P} - \text{almost surely,}$$

where ϱ^W is the probability density function of the transformed random variable $X = \Sigma^{-\frac{1}{2}} Y$. So if $\Sigma^{-\frac{1}{2}}$ exists, the Whitening transformation is bijective and the convergence of the multivariate KDE stated in (1.6) directly follows. A schematic representation of *Algorithm 1.8* is shown in *Figure 1.5*.

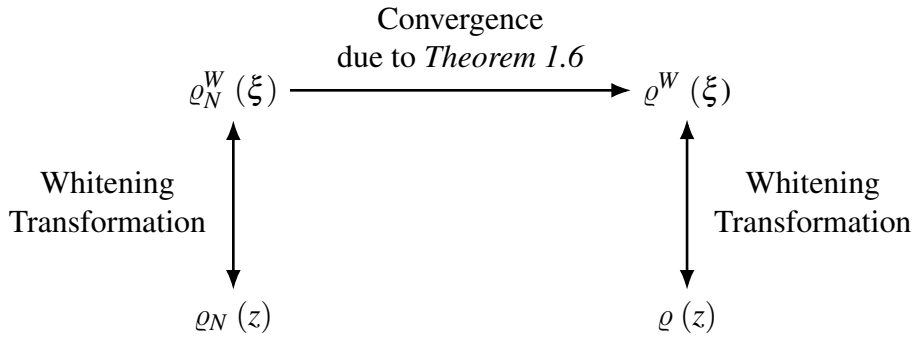


Figure 1.5.: Convergence scheme for the multivariate KDE

This convergence theory also holds for the multivariate KDE with product kernel and diagonal bandwidth matrix H with $H_{i,i} = h^2 \Sigma_{i,i}$. Let σ_i^2 be the variance of the i -th variable of Y . Define the diagonal matrix $V := \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$. Then the bandwidth matrix for the multivariate KDE with product kernel is $H = h^2 V$ with h defined in (1.7). Then *Algorithm 1.8* can be applied analogously using the Whitening transformation

$$X = V^{-\frac{1}{2}} Y.$$

In step 3 of *Algorithm 1.8*, this leads to

$$\varrho_N(\xi) = \varrho_N^W(z) \det(V)^{-\frac{1}{2}},$$

and consequently the multivariate KDE with product kernel and diagonal bandwidth matrix $H = h^2 V$ (with h defined in (1.7)) converges \mathbb{P} -almost surely to the exact probability density function.

Remark 1.10. Let $\sigma_{N,i}^2$ be the variance of the i -th variable of the sampling \mathcal{Y} . When we choose

$$X = V_N^{-\frac{1}{2}} Y,$$

with $V_N = \text{diag}(\sigma_{N,1}^2, \dots, \sigma_{N,n}^2)$, the convergence scheme in *Figure 1.5* also holds.

We now know that both, the univariate KDE and the multivariate KDE converge to the exact probability density function in our setting. Last in this section, we use *Scheffé's lemma* to show that the probabilities also converge. Let \mathcal{B} be the class of all Borel sets in \mathbb{R}^n . Then *Scheffé's lemma* (see [Devroye and Györfi, 1985, Chapter 1, Theorem 1]) guarantees the \mathbb{P} -almost surely convergence for the probabilities.

Theorem 1.11 (Scheffé's lemma). For all densities f and g on \mathbb{R}^n , it holds

$$\int_{\mathbb{R}^n} |f - g| = 2 \sup_{B \in \mathcal{B}} \left| \int_B f - \int_B g \right|.$$

A proof of *Scheffé's lemma* can be found in [Devroye and Györfi, 1985, Chapter 1].

For a set $M(x)$ as it is defined in (1.2), let

$$\mathbb{P}(\zeta \in M(x)) = \int_{M(x)} \varrho(z) dz,$$

denote the exact probability for the random variable Y to be in $M(x)$ (ζ is a realization of Y) and let

$$\mathbb{P}_{\text{KDE}}(\zeta \in M(x)) = \int_{M(x)} \varrho_N(z) dz,$$

denote the probability for the random variable to be in $M(x)$ computed by the KDE. Then we have

$$\begin{aligned} |\mathbb{P}(\zeta \in M(x)) - \mathbb{P}_{\text{KDE}}(\zeta \in M(x))| &= \left| \int_{M(x)} \varrho(z) dz - \int_{M(x)} \varrho_N(z) dz \right| \\ &\leq \frac{1}{2} \int_{\mathbb{R}^n} |\varrho(z) - \varrho_N(z)| dz, \end{aligned}$$

and this directly implies

$$|\mathbb{P}(\zeta \in M(x)) - \mathbb{P}_{\text{KDE}}(\zeta \in M(x))| \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P} - \text{almost surely.}$$

Thus we have shown in this subsection that both, the univariate KDE and the multivariate KDE converge \mathbb{P} -almost surely to the exact probability density functions in our setting resp. that the approximated probabilities converge to the exact probabilities.

1.2. The Spheric Radial Decomposition (SRD)

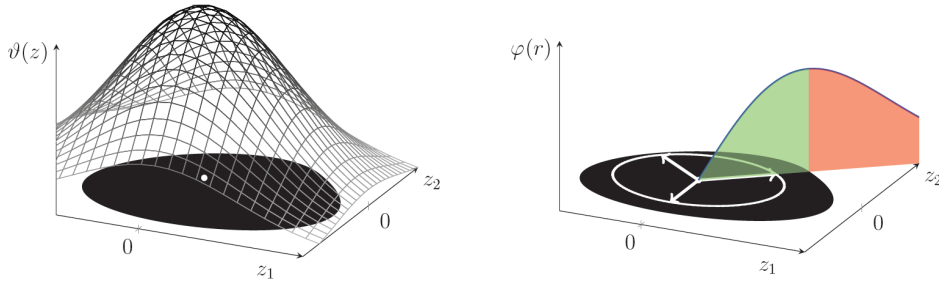
The evaluation of the probabilistic constraint (1.3) can be very difficult due to a very high dimensional integral resp. a high dimensional random variable. In this section, we introduce the *spheric radial decomposition* (see e.g. [Farshbaf-Shaker et al., 2018], [Gotzes et al., 2016], [González-Gradón et al., 2017], [van Ackooij et al., 2018] and [van Ackooij and Henrion, 2014a]), which is basically an integral transformation for Gaussian distributed random variables. As mentioned before, the random variables that are considered later are always Gaussian distributed, so we focus on Gaussian distributed random variables in this section. If the set $M(x)$ in (1.3) is polyhedral, the computation of probability can be done easier (see [Genz and Bretz, 2009]) but in our applications the set $M(x)$ is not even close to a polyhedral set.

Theorem 1.12. (*spheric radial decomposition, see [Gotzes et al., 2016], Theorem 2*) Let $\xi \sim \mathcal{N}(0, R)$ be the n -dimensional standard Gaussian distribution with zero mean and positive definite correlation matrix R . Then, for any Borel measurable subset $M \subseteq \mathbb{R}^n$ it holds that

$$\mathbb{P}(\xi \in M) = \int_{\mathbb{S}^{n-1}} \mu_\chi \{r \geq 0 \mid rLv \in M\} d\mu_\eta(v), \quad (1.9)$$

where \mathbb{S}^{n-1} is the $(n-1)$ -dimensional sphere in \mathbb{R}^n , μ_η is the uniform distribution on \mathbb{S}^{n-1} , μ_χ denotes the χ -distribution with n degrees of freedom and L is such that $R = LL^\top$ (e.g., Cholesky decomposition).

Our aim is to use the SRD for computing the probabilistic constraint (1.3). Here the dependence of the set M (defined in (1.2)) with respect to the decision variable $x \in \mathbb{R}^n$ is not essential for the analysis, so we neglect this dependence here. The idea of the SRD is graphically shown in *Figure 1.6*.



(a) Direct approach: Integrate the probability density function over a certain set

(b) SRD: Integrate rays evaluated at the χ -distribution over the unit sphere

Figure 1.6.: Computing the probability for a random vector to be in a certain set in 2D: Direct approach vs. SRD¹

This result of *Theorem 1.12* can be applied to general Gaussian distributions as follows: For $\xi \sim \mathcal{N}(\mu, \Sigma)$, set $\xi^* = D^{-1}(\xi - \mu) \sim \mathcal{N}(0, R)$ with $D = \text{diag}(\sqrt{\Sigma_{ii}})$ and $R = D^{-1}\Sigma D^{-1}$. Then it follows

$$\mathbb{P}(\xi \in M) = \mathbb{P}(\xi^* \in D^{-1}(M - \mu)).$$

For applying the SRD to specific problems, plenty of knowledge about the set M (defined in (1.2)) is necessary. If this is not the case, one may get

¹Figure already published in [Schuster et al., 2021, page 4]

in trouble trying to apply the SRD. A significant part of the analysis in this thesis consists in characterizing the set M to apply the SRD to the problems which are considered in this thesis. The following algorithm can be used to apply the SRD and it can be found in [Gotzes et al., 2016, Gugat and Schuster, 2018] with a slightly different notation.

Algorithm 1.13 (see [Gotzes et al., 2016, Algorithm 3]). *Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ and L , s.t. $LL^\top = \Sigma$.*

1. *Sample N uniformly distributed points $\mathcal{S} := \{v_1, \dots, v_N\}$ on the unit sphere \mathbb{S}^{n-1} .*
2. *Compute the one dimensional sets $M_v := \{r \geq 0 \mid rLv + \mu \in M\}$ for all $v \in \mathcal{S}$.*
3. *Set $\mathbb{P}(\xi \in M) \approx \frac{1}{N} \sum_{i=1}^N \mu_{\mathcal{X}}(M_{v_i})$.*

The first step of *Algorithm 1.13* can be done easily by creating a Monte Carlo sampling of the standard Gaussian distribution $\mathcal{N}(0_n, \mathbb{E}_{n \times n})$. The sampling \mathcal{S} is chosen as the normalized points of the obtained sampling. In two dimensions, another possibility is to equidistantly choose N points in $[0, 2\pi]$ and use the circle function (i.e. function of the angle of a point on the unit sphere in polar coordinates)

$$f(\varphi) = \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix}, \quad \varphi \in [0, 2\pi],$$

to obtain the sampling \mathcal{S} . In three dimensions this works for the ball function

$$f(\varphi, \vartheta) = \begin{bmatrix} \sin(\varphi) \cos(\vartheta) \\ \sin(\varphi) \sin(\vartheta) \\ \cos(\varphi) \end{bmatrix}, \quad \varphi, \vartheta \in [0, 2\pi],$$

equivalently.

The second step of *Algorithm 1.13* is the most difficult one. It requires a lot of knowledge about the structure of the set M . Since this is available in our setting, the one-dimensional sets M_v ($v \in \mathcal{S}$) can be represented as union of disjoint intervals, i.e.

$$M_v = \bigcup_{j=1}^{\ell_v} I_{v,j}$$

with

$$I_{v,j} = [\underline{a}_{v,j}, \bar{a}_{v,j}] \quad v \in \mathcal{S}, j = 1, \dots, \ell_v.$$

The third step in *Algorithm 1.13* is an approximation of the integral (1.9). It consists in computing the probabilities of the one-dimensional sets M_v ($v \in \mathcal{S}$) with respect to the χ -distribution. These probabilities can be computed easily by using the cumulative distribution function F_χ (with n degrees of freedom) of the χ -distribution (see e.g. [Shea, 1988] and [Abell et al., 1998]), which is given by

$$F_\chi(x) = \Gamma\left(\frac{k}{2}, \frac{x^2}{2}\right),$$

where the incomplete gamma function Γ is given by

$$\Gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt.$$

Approximations with high precision of these functions are available in almost every programming language. So for $v \in \mathcal{S}$ it is

$$\begin{aligned} \mu_\chi(M_v) &= \mu_\chi\left(\bigcup_{j=1}^{\ell_v} I_{v,j}\right) \\ &= \sum_{j=1}^{\ell_v} \mu_\chi(I_{v,j}) \\ &= \sum_{j=1}^{\ell_v} F_\chi(\bar{a}_{v,j}) - F_\chi(\underline{a}_{v,j}). \end{aligned}$$

The probability $\mathbb{P}(\xi \in M)$ as the average over all computed probabilities can be approximately computed by

$$\mathbb{P}(\xi \in M) \approx \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{\ell_{v_i}} F_\chi(\bar{a}_{v_i,j}) - F_\chi(\underline{a}_{v_i,j}).$$

The convergence here is easy to see. We have

$$\mathcal{S} \xrightarrow{N \rightarrow \infty} \mathbb{S}^{n-1},$$

in the sense that for all $x \in \mathbb{S}^{n-1}$ and for all $\varepsilon > 0$, there exists $y \in \mathcal{S}$ with $\|x - y\| \leq \varepsilon$ and vice versa. Thus it follows

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \mu_\chi(M_{v_i}) = \mathbb{P}(\xi \in M).$$

Differently to the KDE approach (introduced in the previous section) we have a classical convergence result here, the convergence is not \mathbb{P} -almost surely.

2. Gas Flow in Pipes

In this chapter we introduce the *isothermal Euler equations* in one dimension, for both, stationary and dynamic states. The isothermal Euler equations model the gas flow in a pipe and they have been the topic of many research articles and books in the last decades. Further we write the isothermal Euler equations in terms of Riemann invariants and in the end of this chapter we give a detailed survey about existing literature on the isothermal Euler equations, on gas flow in pipeline networks and optimization in the context of gas transport.

The isothermal Euler equations are quite similar to the *Saint-Venant equations*, which model the water flow in an open channel. They can analogously be written in terms of Riemann invariants and also the mentioned literature is closely related to the Saint-Venant equations.

From now on we write f_t for the time derivative of some time and space dependent function $f(t, x)$ and f_x for the spatial derivative of f .

2.1. The isothermal Euler equations

For $(t, x) \in [0, T] \times [0, L]$ and $p = p(t, x)$, $q = q(t, x)$ consider the one-dimensional isothermal Euler equations (see e.g. [Osiaacz, 1987], [Banda et al., 2006a], [Gugat et al., 2018], [Gugat and Ulbrich, 2018]), for horizontal pipes

$$\begin{cases} \rho_t + q_x = 0, \\ q_t + \left(p + \frac{q^2}{\rho} \right)_x = -\frac{\lambda^F}{2D} \frac{q|q|}{\rho}. \end{cases} \quad (\text{qISOreal})$$

These equations model the gas flow in a pipe with length L in the time period $[0, T]$ for constant temperature. The notation and the variables are explained in *Table 2.1*.

Variable	Letter	Unit	Range
density	ρ	kg m^{-3}	$\mathbb{R}_{\geq 0}$
flow	q	$\text{kg s}^{-1} \text{ m}^{-2}$	\mathbb{R}
pressure	p	Pa	$\mathbb{R}_{\geq 0}$
velocity	v	m s^{-1}	\mathbb{R}
speed of in the gas	c	m s^{-1}	$\mathbb{R}_{\geq 0}$
pipe friction coefficient	λ^F		$\mathbb{R}_{\geq 0}$
pipe diameter	D	m	\mathbb{R}_+
specific gas constant	R_S	$\text{J kg}^{-1} \text{ K}^{-1}$	\mathbb{R}
temperature	T	K	$\mathbb{R}_{\geq 0}$

Table 2.1.: Notation and variables in this Chapter

For ideal gases the gas pressure and the gas density are directly proportional by the ideal gas equation

$$p = R_S T \rho.$$

For real gases, the proportionality depends on a compressibility factor $z(p)$. It is given by the state equation for real gas (see e.g. [Gugat et al., 2018], [Gugat and Ulbrich, 2018])

$$p = R_S T \rho z(p).$$

A suggestion for the compressibility factor $z(p)$ is stated e.g. in [Gugat et al., 2018], [Gugat and Ulbrich, 2018] and [Schmidt et al., 2014]. For $\alpha \in (-0.5, 0)$ we set

$$z(p) = 1 + \alpha p. \quad (2.1)$$

This compressibility factor is sufficiently accurate for the models in this thesis (see e.g. [Almeida et al., 2014]) and also the American Gas Association (AGA) uses this compressibility factor (see [Schmidt et al., 2014]). Note that most of the theory for real gases can also be applied to ideal gases, i.e. real gases with $\alpha = 0$. Although we focus on ideal gases in this thesis we state results on the isothermal Euler equations for both, ideal and real gases in the literature survey at end of this chapter.

The velocity of the gas v is given by the quotient of flow and density q/ρ . The speed of sound in the gas is given by (see e.g. [Gugat et al., 2018] and [Gugat and Ulbrich, 2018])

$$\left(\frac{1}{c}\right)^2 = \frac{\partial \rho}{\partial p},$$

which is equivalent to

$$c^2 = R_S T,$$

for ideal gases. Another variable that often occurs in this context is the Mach number

$$M := \frac{v}{c},$$

which gives the relation between the velocity of the gas and the speed of sound in the gas. Using the speed of sound in the gas we can write (qISO) for ideal gases as (see e.g. [Gugat and Ulbrich, 2017])

$$\left\{ \begin{array}{l} \rho_t + q_x = 0, \\ q_t + \left(c^2 \rho + \frac{q^2}{\rho} \right)_x = -\frac{\lambda^F}{2D} \frac{q|q|}{\rho}. \end{array} \right. \quad (\text{qISO})$$

Since the speed of sound c does not depend on x we swap speed of sound and the spatial derivative in (qISO). The resulting quotient $(q^2 \rho)/(c^2 \rho^2)$ is equal to $M^2 \rho$. When the gas flow is quite slow, i.e. when the Mach number is close to zero, this term is negligible. So we get a semi-linear form of the isothermal Euler equations for real gases (see e.g. [Gugat and Ulbrich, 2017])

$$\left\{ \begin{array}{l} \rho_t + q_x = 0, \\ q_t + c^2 \rho_x = -\frac{\lambda^F}{2D} \frac{q|q|}{\rho}. \end{array} \right. \quad (\text{sISO})$$

Since we later consider both, a stationary and a dynamic model, we also state the stationary equations here. The stationary form of (qISO) (see e.g. [Gugat et al., 2015], [Gugat et al., 2018]) is given by

$$\left\{ \begin{array}{l} q_x = 0, \\ \left(c^2 \rho + \frac{q^2}{\rho} \right)_x = -\frac{\lambda^F}{2D} \frac{q|q|}{\rho}, \end{array} \right. \quad (\text{qISOstat})$$

and the stationary form of (sISO) (see e.g. [Gugat and Schuster, 2018] and [Schuster et al., 2021]) is given by

$$\left\{ \begin{array}{l} q_x = 0, \\ c^2 \rho_x = -\frac{\lambda^F}{2D} \frac{q|q|}{\rho}. \end{array} \right. \quad (\text{sISOstat})$$

In the literature survey we also state some results on the stationary equations.

2.2. The isothermal Euler equations in terms of Riemann invariants

In this subsection we write the systems (qISO) and (sISO) in terms of Riemann invariants. The advantage is, that the equations in the resulting systems are decoupled in the sense that every equation contains only derivatives of one variable. Every 2×2 system of hyperbolic balance laws can be written in terms of Riemann invariants without restrictions, for larger systems one need additional assumptions. A detailed introduction to Riemann invariants and hyperbolicity can be found in [Meleshko, 2015, Chapter 2]. Many articles deal with hyperbolic systems of this type. The Riemann invariants for the isothermal Euler equations have already been computed in [Gugat and Ulbrich, 2018], but since the authors use the state variables pressure and flow (while we use density and flow here) our Riemann invariants slightly differ from the Riemann invariants stated in [Gugat and Ulbrich, 2018].

We define

$$F(\rho, q) := -\frac{\lambda^F q|q|}{2D \rho},$$

and we write (qISO) in matrix form:

$$\begin{bmatrix} \rho \\ q \end{bmatrix}_t + \underbrace{\begin{bmatrix} 0 & 1 \\ c^2 - \frac{q^2}{\rho^2} & 2\frac{q}{\rho} \end{bmatrix}}_{=: A} \begin{bmatrix} \rho \\ q \end{bmatrix}_x = \begin{bmatrix} 0 \\ F(\rho, q) \end{bmatrix}. \quad (2.2)$$

Since the matrix A is not diagonal, the derivatives in both equations depend on both variables. Using the velocity of the gas, we can write the matrix A as

$$A = \begin{bmatrix} 0 & 1 \\ c^2 - v^2 & 2v \end{bmatrix}.$$

The eigenvalues of A are given by

$$\lambda_1 = v + c \quad \text{and} \quad \lambda_2 = v - c.$$

Surely the eigenvalues depend on ρ and q but for better readability we neglect this dependence. We compute the corresponding left eigenvectors $\ell^{(1)}$ and $\ell^{(2)}$ by solving

$$\begin{bmatrix} \ell_1^{(i)} & \ell_2^{(i)} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ c^2 - v^2 & 2v \end{bmatrix} = \begin{bmatrix} \ell_1^{(i)} & \ell_2^{(i)} \end{bmatrix} \lambda_i, \quad i = 1, 2.$$

This leads to the eigenvector

$$\ell^{(1)} = \frac{1}{c\rho} [c - v \quad 1] = \frac{1}{c\rho} [-\lambda_2 \quad 1],$$

corresponding to the eigenvalue $\lambda_1 = v + c$ and to the eigenvector

$$\ell^{(2)} = \frac{1}{c\rho} [c + v \quad -1] = \frac{1}{c\rho} [\lambda_1 \quad -1],$$

corresponding to the eigenvalue $\lambda_2 = v - c$. The fact that $\ell^{(1)}$ depends on λ_2 and $\ell^{(2)}$ depends on λ_1 should not confuse the reader. We multiply (2.2) with these left eigenvectors. Hence we have

$$\frac{1}{c\rho} [-\lambda_2 \rho_t + q_t] + \lambda_1 \frac{1}{c\rho} [-\lambda_2 \rho_x + q_x] = \frac{1}{c\rho} F(\rho, q),$$

and

$$\frac{1}{c\rho} [\lambda_1 \rho_t - q_t] + \lambda_2 \frac{1}{c\rho} [\lambda_1 \rho_x - q_x] = \frac{1}{c\rho} F(\rho, q).$$

Thus the Riemann invariants are defined by

$$\frac{\partial}{\partial t} R_1(\rho, q) = \frac{1}{c\rho} [-\lambda_2 \rho_t + q_t] \quad \text{resp.} \quad \frac{\partial}{\partial x} R_1(\rho, q) = \frac{1}{c\rho} [-\lambda_2 \rho_x + q_x], \quad (2.3)$$

and

$$\frac{\partial}{\partial t} R_2(\rho, q) = \frac{1}{c\rho} [\lambda_1 \rho_t - q_t] \quad \text{resp.} \quad \frac{\partial}{\partial x} R_2(\rho, q) = \frac{1}{c\rho} [\lambda_1 \rho_x - q_x]. \quad (2.4)$$

Lemma 2.1. *The Riemann invariants defined by (2.3) and (2.4) are given by*

$$R_1(\rho, q) = \ln(\rho) + \frac{q}{c\rho} = \ln(\rho) + M, \quad (2.5)$$

and

$$R_2(\rho, q) = \ln(\rho) - \frac{q}{c\rho} = \ln(\rho) - M. \quad (2.6)$$

Proof. We compute the time derivative of R_1 :

$$\begin{aligned} \frac{\partial}{\partial t} R_1(\rho, q) &= \frac{1}{\rho} \rho_t + \frac{1}{c} \frac{\rho q_t - q \rho_t}{\rho^2} \\ &\stackrel{v = \frac{q}{\rho}}{=} \frac{1}{c\rho} [c\rho_t + q_t - v\rho_t] \\ &\stackrel{\lambda_2 = v - c}{=} \frac{1}{c\rho} [-\lambda_2 \rho_t + q_t]. \end{aligned}$$

The spatial derivative of $R_1(\rho, q)$ can be computed analogously. For $R_2(\rho, q)$ we have

$$\begin{aligned}\frac{\partial}{\partial t}R_2(\rho, q) &= \frac{1}{c\rho}\rho_t - \frac{1}{c}\frac{\rho q_t - q\rho_t}{\rho^2} \\ &\stackrel{v=\frac{q}{\rho}}{=} \frac{1}{c\rho}[c\rho_t - q_t + v\rho_t] \\ \lambda_1 &\stackrel{v+c}{=} \frac{1}{c\rho}[\lambda_1\rho_t - q_t].\end{aligned}$$

The spatial derivative of $R_2(\rho, q)$ can be computed analogously. \square

Thus, for $R_1(\rho, q)$ and $R_2(\rho, q)$ as in *Lemma 2.1*, the system (qISO) can be equivalently written as

$$\begin{bmatrix} R_1(\rho, q) \\ R_2(\rho, q) \end{bmatrix}_t + \begin{bmatrix} v+c & 0 \\ 0 & v-c \end{bmatrix} \begin{bmatrix} R_1(\rho, q) \\ R_2(\rho, q) \end{bmatrix}_x = \frac{1}{c\rho}F(\rho, q) \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

The system is now decoupled, which means that in the first equation, only the derivatives of $R_1(\rho, q)$ occur and in the second equation, only the derivatives of $R_2(\rho, q)$ occur. Henceforth, for better readability, we neglect the dependencies of the Riemann invariants, so we write R_1 and R_2 instead of $R_1(\rho, q)$ and $R_2(\rho, q)$. Now we write the remaining variables in terms of Riemann invariants. For $v = q/\rho$ we have

$$v = \frac{c}{2}(R_1 - R_2),$$

and for $M = v/c$ we have

$$M = \frac{1}{2}(R_1 - R_2).$$

Then for the eigenvalues $\lambda_{1/2} = v \pm c$ it follows

$$\lambda_{1/2} = \frac{c}{2}(R_1 - R_2) \pm c = c \left[\frac{1}{2}(R_1 - R_2) \pm 1 \right]. \quad (2.7)$$

For the right hand side it follows

$$\begin{aligned}\frac{1}{c\rho}F(\rho, q) &= -\frac{1}{c\rho}\frac{\lambda^F q|q|}{2D\rho} \\ &= -\frac{\lambda^F q|q|}{2cD\rho^2} \\ &\stackrel{(*)}{=} -\frac{\lambda^F}{2cD}v|v| \\ &= -\frac{c\lambda^F}{8D}(R_1 - R_2)|R_1 - R_2|.\end{aligned} \quad (2.8)$$

The step (*) in (2.8) only holds if ρ is positive, which is always true as we will see later. So the whole system (qISO) can be equivalently written as

$$\begin{bmatrix} R_1 \\ R_2 \end{bmatrix}_t + \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}_x = -\frac{c\lambda^F}{8D}(R_1 - R_2)|R_1 - R_2| \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad (\text{qISOR})$$

with R_1, R_2 as in *Lemma 2.1* and $\lambda_{1/2}$ as in (2.7). The last thing missing is the expression of ρ and q in terms of Riemann invariants. It follows

$$\rho = \exp\left(\frac{R_1 + R_2}{2}\right),$$

which is always positive due to the exponential function. So step (*) in (2.8) holds. Further we have

$$q = c \frac{R_1 - R_2}{2} \exp\left(\frac{R_1 + R_2}{2}\right),$$

where the sign on q depends on the velocity term $(R_1 - R_2)/2$.

Next we also write the system (sISO) in terms of Riemann invariants. We have

$$\begin{bmatrix} \rho \\ q \end{bmatrix}_t + \underbrace{\begin{bmatrix} 0 & 1 \\ c^2 & 0 \end{bmatrix}}_{=: A} \begin{bmatrix} \rho \\ q \end{bmatrix}_x = \begin{bmatrix} 0 \\ -\frac{\lambda^F q|q|}{2D\rho} \end{bmatrix}. \quad (2.9)$$

The eigenvalues of A are given by $\lambda_{1/2} = \pm c$ and the corresponding left eigenvectors are given by

$$\ell^{(1)} = [c \quad 1] \quad \text{and} \quad \ell^{(2)} = [-c \quad 1].$$

We multiply (2.9) with these left eigenvectors. Hence we have

$$(q_t + c\rho_t) + c(q_x + c\rho_x) = -\frac{\lambda^F q|q|}{2D\rho}, \quad (2.10)$$

and

$$(q_t - c\rho_t) - c(q_x - c\rho_x) = -\frac{\lambda^F q|q|}{2D\rho}, \quad (2.11)$$

Corollary 2.2. *The Riemann invariants defined in (2.10) and (2.11) are given by*

$$R_1(\rho, q) = q + c\rho,$$

and

$$R_2(\rho, q) = q - c\rho.$$

Thus for $R_1(\rho, q)$ and $R_2(\rho, q)$ as in *Corollary 2.2* the system (sISO) can be equivalently written as

$$\begin{bmatrix} R_1(\rho, q) \\ R_2(\rho, q) \end{bmatrix}_t + \begin{bmatrix} c & 0 \\ 0 & -c \end{bmatrix} \begin{bmatrix} R_1(\rho, q) \\ R_2(\rho, q) \end{bmatrix}_x = -\frac{\lambda^F q|q|}{2D \rho} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

This system is also decoupled now. The eigenvalues are constant and do not depend in the Riemann invariants. Next we need an expression for ρ and q in terms of Riemann invariants. Again we neglect the arguments of R_1 and R_2 . It follows

$$q(t, x) = \frac{1}{2} (R_1 + R_2),$$

and

$$\rho(t, x) = \frac{1}{2c} (R_1 - R_2).$$

Here ρ may not always be positive (contrary to the quasilinear case) but one can find boundary conditions, s.t. the solution of the corresponding boundary value problem has a solution with positive density. For the right hand side it follows

$$\begin{aligned} -\frac{\lambda^F q|q|}{2D \rho} &= -\frac{\lambda^F \frac{1}{4} (R_1 + R_2) |R_1 + R_2|}{2D \frac{1}{2c} R_1 - R_2} \\ &= -\frac{c\lambda^F (R_1 + R_2) |R_1 + R_2|}{4D R_1 - R_2}. \end{aligned}$$

Thus system (sISO) can be equivalently written as

$$\begin{bmatrix} R_1 \\ R_2 \end{bmatrix}_t + \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}_x = -\frac{c\lambda^F (R_1 + R_2) |R_1 + R_2|}{4D R_1 - R_2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad (\text{sISOR})$$

with R_1, R_2 stated in *Corollary 2.2* and $\lambda_{1/2} = \pm c$.

In the next section we give an overview about the existing literature on these systems.

2.3. Literature survey

In this section we give an overview about selected literature related to gas transport and general hyperbolic flow problems. Further some references particularly deal with the existence of solutions and optimal control problems in this context.

A general overview about the physics of gas transportation, about gas networks, different stationary and dynamic models and the modelling of several network elements (compressor stations, valves and more) can be found in [Domschke et al., 2017], in [Koch et al., 2015] and [Míos-Mercado and Borraz-Sánchez, 2015].

For the system (sISOstat) a solution can easily be found:

Corollary 2.3. For $\lambda^F \in \mathbb{R}_{\geq 0}$ and $D \in \mathbb{R}_+$ and $c \in \mathbb{R}$,

$$q(x) = k_1 \quad \text{and} \quad \rho^2(x) = -\frac{\lambda^F}{2Dc}q|q|x + k_2 \quad \text{with} \quad k_1, k_2 \in \mathbb{R}, \quad (2.12)$$

solve (sISOstat).

If we assume positive density for the gas, which of course makes sense from a real world point of view, the solution (2.12) can be uniquely determined by suitable boundary conditions. In [Gotzes et al., 2016], in [Gugat and Schuster, 2018], in [Gugat et al., 2020], in [Schuster et al., 2021] and in the following chapter one can find analytical solutions resp. characterizations of analytical solutions for (sISOstat) on network graphs with appropriate coupling conditions.

In [Gugat et al., 2015] the authors prove the existence of a solution on both, a single interval and a network graph, for the model (qISOstat) with non-horizontal pipes. The proof for networks graphs consists in opening up the cycles inductively until the network does not contain cycles anymore. Further the authors state an explicit expression for the solution of (qISOstat) on a single pipe by using the *Lambert-W* function. In [Schmidt et al., 2014] the authors discuss stationary states on a single pipe and in [Gugat et al., 2018] the existence of a solution for (qISOstat) for real gases on an arbitrary network graph is stated. The authors also analyze the difference between the real gas model and the ideal gas model.

The transient model (qISO) has been analyzed in [Banda et al., 2006b]. In [Gugat and Ulbrich, 2017] the authors show that for both transient models, (qISO) and (sISO), one can construct product solutions for which initial states with arbitrary large C^1 -norm exist, s.t. the solution does not blow up in finite time. Further, the authors present a classical solution of (qISO) where the direction of the gas flow changes. One year later the authors showed the existence of semi-global Lipschitz continuous solutions of (qISOR) on a network with compatible coupling conditions, based upon a fix point iteration along

the characteristic curves (see [Gugat and Ulbrich, 2018]). Further, in [Gugat and Wintergerst, 2018], the authors give analytic traveling wave solutions for (qISO) for real gas.

In [Bressan et al., 2014] the authors give an overview about general hyperbolic flow problems on networks. The existence, uniqueness and stability of general networked hyperbolic systems was studied in [Gugat et al., 2012]. Different coupling conditions for (qISO) without source term are analyzed in [Banda et al., 2006a]. These coupling conditions can be used similarly for general flow problems.

The existence and uniqueness of a global smooth solution for the fluid flow in a pipe (similar to (qISO)) with sufficiently small C^1 -initial data has been shown in [Luskin, 1981]. Moreover in [Lowe and Clarke, 2002] and [Lowe and Clarke, 2003] the authors stated analytical solutions with shocks for the full Euler equations with source term.

The model (qISO) was also considered in [Colombo and Marcellini, 2009] without friction, i.e. without source term. Conservation laws for traffic flow problems without source term are analyzed in [Garavello and Piccoli, 2009]. Further, in [Chen, 2012] the author analyzes the existence and uniqueness of Lipschitz solutions of 1-D first order quasilinear hyperbolic systems without source term.

In [Colombo et al., 2009], the authors study optimal control problems related to networks of pipes and canals. Further, results on the optimal compressor control for (qISO) are stated in [Banda and Herty, 2011]. The existence of an optimal control of systems governed by nonlinear systems of conservation laws, motivated by gas compressors in pipelines and open channels management, has been shown in [Colombo et al., 2008].

In [Gerster and Herty, 2019], the authors analyze the stability of flow problems with feedback control on networks, described by hyperbolic balance laws. Further they give sufficient conditions for exponential stability. The existence of classical solutions for (qISOR) on two pipes, connected by a compressor station, has been shown in [Gugat and Herty, 2011]. Further the authors analyze the feedback stabilization for the flow. In [Gugat et al., 2011] this has been done for more complex networks.

Surely there is much more literature on gas transportation, hyperbolic conservation laws, compressor control and stabilization which cannot all be mentioned here but this survey touches all these topics and states some important contributions related to gas transportation.

3. Probabilistic Constrained Optimization on Stationary Gas Networks

In this chapter, we consider gas network equilibria, i.e. stationary states in gas networks. Parts of this section are already published in [Schuster et al., 2021, Section 2] and [Gugat and Schuster, 2018]. In this work we go more into detail and we focus more on the optimization part as it is done in [Schuster et al., 2021] and [Gugat and Schuster, 2018]. For a detailed model introduction, we need a few definitions.

Definition 3.1. *Let a connected, directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a vertex set $\mathcal{V} = \{v_0, \dots, v_n\}$ and a set of edges $\mathcal{E} = \{e_1, \dots, e_m\} \subseteq \mathcal{V} \times \mathcal{V}$ be given ($n, m \in \mathbb{N}$).*

- (i) *For an edge $e_k = (v_i, v_j) \in \mathcal{E}$, $h(e_k) = v_j$ denotes the head of the edge e_k and $f(e_k) = v_i$ denotes the foot of the edge e_k ($i, j \in \{0, \dots, n\}$, $k \in \{1, \dots, m\}$).*
- (ii) *The set $\mathcal{E}_+(v_k) := \{e \in \mathcal{E} \mid f(e) = v_k\}$ denotes all edges with foot in v_k ($k \in \{0, \dots, n\}$).*
- (iii) *The set $\mathcal{E}_-(v_k) := \{e \in \mathcal{E} \mid h(e) = v_k\}$ denotes all edges with head in v_k ($k \in \{0, \dots, n\}$).*
- (iv) *The set $\mathcal{E}_0(v_k) := \mathcal{E}_+ \cup \mathcal{E}_-$ denotes all edges, which are connected to v_k ($k \in \{0, \dots, n\}$).*
- (v) *The set $\mathcal{V}_0 := \{v \in \mathcal{V} \mid |\mathcal{E}_-(v)| = \emptyset\}$ denotes all inflow nodes.*
- (vi) *The matrix $A^+ \in \mathbb{R}^{(n+1) \times m}$ with $A_{i,j}^+ = \sigma(v_i, e_j)$,*

$$\sigma(v, e) := \begin{cases} -1 & \text{if } e \in \mathcal{E}_+(v), \\ 1 & \text{if } e \in \mathcal{E}_-(v), \\ 0 & \text{else,} \end{cases}$$
is called incidence matrix of the graph \mathcal{G} .

Some of the results of this thesis mainly depend on the topology of the graph, so we introduce different structures for a graph.

Definition 3.2. Let a connected, directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a vertex set $\mathcal{V} = \{v_0, \dots, v_n\}$ and a set of edges $\mathcal{E} = \{e_1, \dots, e_m\} \subseteq \mathcal{V} \times \mathcal{V}$ be given ($n, m \in \mathbb{N}$).

- (i) The graph \mathcal{G} is called linear, if $|\mathcal{E}_+(v)| \leq 1$ and $|\mathcal{E}_-(v)| \leq 1$ for all $v \in \mathcal{V}$, i.e. the graph does not contain junctions.
- (ii) The graph \mathcal{G} is called tree-structured, if $|\mathcal{E}_-(v)| \leq 1$ for all $v \in \mathcal{V}$, i.e. the graph does not contain cycles.

In the next sections we introduce a model for the stationary gas flow in pipeline networks. We use both introduced methods, the SRD and the KDE (see *Chapter 1*) to compute the probability for random gas outflow to satisfy certain constraints. Then we use the developed theory to get some results on probabilistic constrained optimization problems in the context of gas transport. Last in this chapter we discuss general advantages and disadvantages of the SRD and the KDE applied to our model and similar settings. Throughout this chapter we use numerical examples to illustrate the results.

3.1. Mathematical Modelling and Problem Description

In this section we introduce a mathematical model for the stationary gas transport in a pipeline network. We focus on tree-structured networks but our analysis can be extended to networks with cycles as we will see later in *Section 3.7*. We mention here, that [Ahuja et al., 1993] gives a great introduction to several classes of flow problems on networks.

Consider a connected, directed, tree-structured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the vertex set $\mathcal{V} = \{v_0, \dots, v_n\}$ and the set of edges $\mathcal{E} = \{e_1, \dots, e_n\}$ with $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Due to the structure, the graph has a unique root v_0 (see *Definition 3.2*). From the root the graph is numbered using breadth-first search or depth-first search, a detailed explanation of these methods can be found in [Cormen et al., 2009, Chapter 22.2 and 22.3]. Every edge $e \in \mathcal{E}$ represents a pipe with (positive) length L^e . For $x \in [0, L^e]$ we consider the stationary semi-linear isothermal Euler equations for horizontal pipes and ideal gases

$$\begin{cases} q_x^e(x) = 0, \\ (c^e)^2 p_x^e(x) = -\frac{\lambda^e}{2D^e} (R_S T)^2 \frac{q^e(x) |q^e(x)|}{p^e(x)}. \end{cases} \quad (\text{sISOstat})$$

Here, p is the pressure, q is the flow and $c^e, \lambda^e, D^e \in \mathbb{R}_{>0}$ denote the speed of sound in the gas, the pipe friction coefficient and the pipe diameter of edge $e \in \mathcal{E}$ respectively. For simplicity we assume that these parameters are space independent. $R_S, T \in \mathbb{R}_{>0}$ are specific gas constant and (constant) temperature, both come from the ideal gas equation. Further p^e and q^e represent the restriction of the pressure and the flow defined over the network to a single edge $e \in \mathcal{E}$.

Remark 3.3. For a single pipe and boundary conditions $p(0) = p_0, q(L) = b$, the solution of (sISOstat) is given by

$$\begin{cases} q(x) = b, \\ p(x) = \sqrt{p_0^2 - \frac{\lambda}{c^2 D} (R_S T)^2 q(x) |q(x)| x}. \end{cases} \quad (3.1)$$

The second equation in (3.1) is called *Weymouth equation* (see e.g. [Gugat et al., 2018], [Domschke et al., 2017], [Gugat and Schuster, 2018]). For the gas network, we need to state coupling conditions at the nodes to guarantee well-posedness. We assume gas outflow at every node except the (unique) root of the graph. Let a vector $b \in \mathbb{R}_{\geq 0}^n$ be given. The i -th component of b represents the i -th consumers gas demand, i.e. the outflow at node $v_i \in \mathcal{V}$. In general, we state a positive sign of b_i if gas leaves the network and a negative sign for b_i if gas enters the network ($i = 1, \dots, n$). The vector $b \in \mathbb{R}_{\geq 0}^n$ is called load vector.

Note that q^e is constant on every edge. The sign of the flow q^e is positive if gas flows along the orientation of the edge $e \in \mathcal{E}$ and it is negative if gas flows against the orientation of the edge $e \in \mathcal{E}$.

Now we state coupling conditions on the graph. We assume conservation of mass at every node $v \in \mathcal{V} \setminus \mathcal{V}_0$. That means that the gas inflow at every node must be equal to the sum of gas outflow and load of this node, i.e.

$$\sum_{e \in \mathcal{E}_-(v)} q^e \left(\frac{D^e}{2} \right)^2 \pi = b^v + \sum_{e \in \mathcal{E}_+(v)} q^e \left(\frac{D^e}{2} \right)^2 \pi \quad \forall v \in \mathcal{V} \setminus \mathcal{V}_0.$$

The term $(D^e/2)^2 \pi$ is the cross section area of pipe $e \in \mathcal{E}$. We assume that every pipe has the same diameter D . Thus for the conservation of mass it follows

$$\sum_{e \in \mathcal{E}_-(v)} q^e = \frac{4}{D^2 \pi} b^v + \sum_{e \in \mathcal{E}_+(v)} q^e \quad \forall v \in \mathcal{V} \setminus \mathcal{V}_0. \quad (3.2)$$

As mentioned before, the graph has a unique root v_0 (i.e. $|\mathcal{V}_0| = 1$), which is the only node with gas inflow in the graph. The inflow is explicitly given by (3.2), s.t.

$$b^{v_0} = - \sum_{v \in \mathcal{V} \setminus \mathcal{V}_0} b^v. \quad (3.3)$$

Let the pressure at the inflow node v_0 be given, that is

$$p^e(0) = p_0 \in \mathbb{R}_{>0} \quad \forall e \in \mathcal{E}_+(v_0).$$

Further we assume continuity in pressure at every inner node, i.e. for all $v \in \mathcal{V}$ with $\mathcal{E}_-(v) \neq \emptyset$ and $\mathcal{E}_+ \neq \emptyset$ holds

$$p^{e_1}(L^{e_1}) = p^{e_2}(0) \quad \forall e_1 \in \mathcal{E}_-(v), e_2 \in \mathcal{E}_+(v). \quad (3.4)$$

Other coupling conditions for the isothermal Euler equations were discussed in [Banda et al., 2006a]. For isentropic flow coupling conditions were discussed recently in [Holle et al., 2020]. Before we state the complete model in this chapter, we define the constants

$$\phi^e := \frac{\lambda^e}{D(c^e)^2} (R_S T)^2,$$

for every edge $e \in \mathcal{E}$. So for every edge $e \in \mathcal{E}$, the full model is given by

$$\left\{ \begin{array}{l} q_x^e(x) = 0 \\ p_x^e(x) = -\frac{\phi^e q^e(x) |q^e(x)|}{2 p^e(x)} \\ q^e(L^e) = b^{h(e)} + \sum_{\kappa \in \mathcal{E}_+(h(e))} q^\kappa(0) \\ p^e(0) = p_0 \quad e \in \mathcal{E}_+(v_0) \\ p^e(0) = p^\kappa(L^\kappa) \quad \kappa \in \mathcal{E}_-(f(e)), e \in \mathcal{E} \setminus \mathcal{E}_+(v_0) \end{array} \right. \quad (\text{statModel})$$

The existence of a solution follows from [Gugat et al., 2015] but due to the pressure drop in the pipes caused by friction we need to assume that the pipes are sufficiently short, i.e. for an edge $e \in \mathcal{E}$ the critical length is given by

$$L_{\text{crit}}^e = \frac{(p^e(0))^2}{\phi^e q^e |q^e|}.$$

Note that the flow is constant in space, so there is no dependency in x . For $L^e > L_{\text{crit}}^e$ the pressure at the end of the edge would be negative. Usually the pipe friction coefficient is quite small, even close to zero, s.t. sufficiently short does not automatically mean that the pipe is really short. For tree-structured graphs, an explicit solution can directly be constructed by the solution of the

stationary semi-linear isothermal Euler equations. We use an explicit formulation of the solution later in *Section 3.2*.

Due to the fact that the solution of the isothermal Euler equations is constant in flow, we define the vector $q \in \mathbb{R}^n$, which denotes all flows in the network, where q_i denotes the flow at the edge $e_i \in \mathcal{E}$ ($i = 1, \dots, n$). Further, we define the vector $p \in \mathbb{R}^n$, which denotes the pressures at the nodes (except the node v_0), where p_k denote the pressure at node $v_k \in \mathcal{V}$ ($k = 1, \dots, n$). Note that the pressure at node v_0 is given by p_0 .

Our aim is now to find outflow vectors $b = (b^v)_{v \in \mathcal{V} \setminus \{v_0\}} \in \mathbb{R}^n$, s.t. the solution of system (statModel) satisfies box constraints for the pressures. The motivation of this model is application driven. The outflow represents the consumers gas demand and the inlet pressure is given by the gas producers. The equations in (statModel) represent the physics of stationary gas transport in a pipeline network. The box constraints for the pressures are a restriction for the solution of (statModel). Since the solution of the isothermal Euler equations is monotonously decreasing (for non-negative loads $b \geq 0$), it is sufficient to assume box constraints for the pressures at the nodes. These bounds can be interpreted as restrictions on the pipe material in the sense that too much pressure would damage the pipes and a pressure that is too low would lead to almost no gas flow.

The question that we want to answer is:

Q_{det}: For a given inlet pressure, can we guarantee, that all consumers receive their demanded gas, s.t. the gas pressure in the network is neither too high nor too low?

We are interested in the pressures at the nodes. Let $p \in \mathbb{R}^n$ be the vector of pressures at the outflow nodes, i.e. p_i is the pressure at node v_i ($i = 1, \dots, n$). For lower pressure bounds $p_i^{\min} > 0$ and upper pressure bounds $p_i^{\max} \geq p_i^{\min}$, we demand

$$p_i \in [p_i^{\min}, p_i^{\max}] \quad i = 1, \dots, n.$$

For convenience we write

$$p \in [p^{\max}, p^{\min}],$$

which has to be understood component-by-component. We define the set of feasible loads as follows:

Definition 3.4. For pressure bounds $p^{\min}, p^{\max} \in \mathbb{R}_{\geq 0}^n$ with $0 < p_i^{\min} \leq p_i^{\max}$ ($i = 1, \dots, n$), the set

$$M := \left\{ b \in \mathbb{R}_{\geq 0}^n \mid \begin{array}{l} \text{The solution } (p, q) \in \mathbb{R}^n \times \mathbb{R}^n \text{ of } (\text{statModel}) \\ \text{satisfies the box constraints } p \in [p^{\min}, p^{\max}]. \end{array} \right\} \quad (3.5)$$

is called the set of feasible loads.

Surely the pressures and the flows depend on the loads but we neglect this dependence here for better readability.

The gas market in European countries fundamentally changed in the beginning of the 21st century. On the one hand the aim was to counteract the monopoly of a few companies and on the other hand the gas consumers should be able to choose from which company they buy gas. Since 2005 the *Federal Network Agency* (Bundesnetzagentur, BNetzA) observes the German gas network and demands supply guarantee (among other things) from the gas network owners. But from a gas network owners point of view the consumers gas demand cannot be known a priori, it depends on various factors like e.g. temperature and gas price. Mathematically that means that the gas demand can be seen as random variable. For a mean value $\mu \in \mathbb{R}_{\geq 0}^n$ and a positive definite covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$, we define a Gaussian distributed random variable

$$\xi_b \sim \mathcal{N}(\mu, \Sigma),$$

on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We identify the load vector $b \in \mathbb{R}^n$ with the image $\xi_b(\omega)$ for $\omega \in \Omega$ on this probability space. In [Gotzes et al., 2016] and [Koch et al., 2015, Chapter 13], the authors explain, that a multivariate Gaussian distribution is a good choice for the random load vector.

So we cannot answer the question Q_{det} exactly, but we can answer it in a probabilistic sense. Our aim now is to answer the following question:

***Q_{prob}:** For a given inlet pressure, can we guarantee, that every consumer receives their demanded gas, s.t. the gas pressure in the network is neither too high nor too low, in at least $\alpha\%$ of all scenarios?*

Since we are in a stationary setting in this chapter, the scenarios are time independent. We want to compute the probability for a random load vector to be feasible, i.e.

$$\mathbb{P}(b \in M). \quad (3.6)$$

Usually the aim is to choose the box constraints for the pressure (or at least the upper pressure bound) such that the probability is almost 1. Then the network is robust in a probabilistic sense against perturbations in the outflow.

Later we consider probabilistic constrained optimization problems where we minimize the upper pressure bound, s.t. the probability (3.6) is sufficiently large. We mention again that we mean robustness in a probabilistic sense here, not in a classical sense. The classical robust optimization deals with the optimization of worst-case scenarios and this often leads to over-conservative results, which is not necessary in the context of gas networks. Surely there are applications, where it is indispensable that every scenario satisfies all given restrictions (e.g. in the context of nuclear power generation) but in the context of gas transport, it is not a disaster if a (improbable) scenario does not satisfy all needs.

In the next sections, we use the two introduced methods, the SRD (see *Section 1.2*) and the KDE (see *Section 1.1*) to compute this probability. The difference is the following:

The load vector b is considered to be random, it has a well-known distribution, but we have no information about the structure of the set of feasible loads M . The pressures at the nodes have to satisfy box constraints. We define the n -dimensional cuboid

$$P_{\min}^{\max} := \otimes_{i=1}^n [p_i^{\min}, p_i^{\max}], \quad (3.7)$$

as the set of feasible pressures. Due to the randomness of the load, the pressures at the nodes p are also random with an unknown distribution, but the set of feasible pressures is a well-known cuboid. It is

$$\mathbb{P}(b \in M) = \mathbb{P}(b \geq 0 \text{ and } p \in P_{\min}^{\max}). \quad (3.8)$$

So on the one hand, we characterize the set of feasible loads and use the SRD to compute the desired probability, on the other hand we use a KDE approach to estimate the probability density of the pressure and integrate this density over the set of feasible pressures to compute the desired probability. This scheme is shown in *Figure 3.1*.

3.2. Application of the SRD

In this section, we apply the SRD (introduced in *Section 1.2*) to the stationary model that was introduced in *Section 3.1*. The aim is to compute the probability for a random load vector to be feasible, so we basically answer the question Q_{prob} .

Consider the feasible set defined in (3.5). As it is mentioned in the last section, we first characterize the set of feasible loads. For the further analysis, we

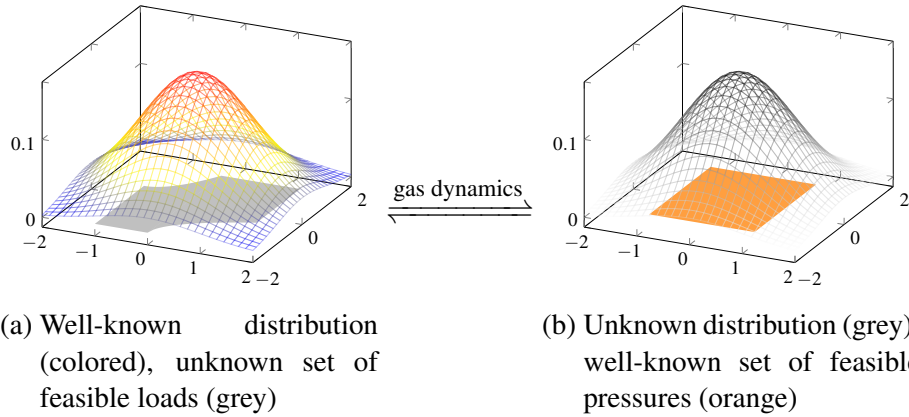


Figure 3.1.: The two different ways to compute the probability for a random load vector to be feasible (3.8)

need the incidence matrix of the graph \mathcal{G} (see *Definition 3.1*). Let A^+ be the incidence matrix for the graph \mathcal{G} . We define the incidence matrix for all outflow nodes $A \in \mathbb{R}^{n \times n}$ as $A^+ \in \mathbb{R}^{(n+1) \times n}$ without the first row (corresponding to the root of the graph), i.e.

$$A^+ = \begin{bmatrix} A_0 \\ A \end{bmatrix} \in \begin{bmatrix} \mathbb{R}^{1 \times n} \\ \mathbb{R}^{n \times n} \end{bmatrix}. \quad (3.9)$$

With the flow vector $q \in \mathbb{R}^n$ and the load vector $b \in \mathbb{R}^n$ we can state the conservation of mass for all outflow nodes of the graph as

$$Aq = b.$$

The corresponding inflow at the node v_0 (due to constant flow along the edges) can be computed by (3.3). Because A is a square matrix with full rank, we can directly compute the flow vector by

$$q = A^{-1}b. \quad (3.10)$$

If the graph would not be tree-structured, the matrix A would not be square and the vector q could not be computed that easy.

Remark 3.5. *Due to the breadth-first search resp. the depth-first search numbering of the graph, the incidence matrix A and its inverse A^{-1} are upper triangular with diagonal $\text{diag}(A) = \text{diag}(A^{-1}) = \mathbb{1}_n$, where $\mathbb{1}_n$ is the vector of ones in \mathbb{R}^n . Further for the j -th column of A^{-1} ($j = 1, \dots, n$) we have $A_{i,j}^{-1} = 1$ iff the node v_i is on the (unique) path from v_0 to v_j and we have $A_{i,j}^{-1} = 0$ otherwise (for all $i = 1, \dots, n$).*

For the continuity in pressure we define a diagonal matrix

$$\Phi \in \mathbb{R}^{n \times n}, \quad \Phi_{i,i} = \phi_i := \frac{\lambda^{e_i}}{(c^{e_i})^2 D^{e_i}} (R_{ST})^2 L^{e_i}.$$

Further we use (3.10) and we define a function

$$g: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad g(b) = (A^\top)^{-1} \Phi |A^{-1}b| (A^{-1}b), \quad (3.11)$$

which gives us the pressure loss in the graph with respect to the root of the graph (the only inflow node). So $g_i(b)$ states the pressure loss from the root to the node v_i ($i = 1, \dots, n$). The following result is stated in [Gotzes et al., 2016] and [Gugat and Schuster, 2018].

Corollary 3.6. *Consider the system (statModel). The pressure at the nodes $p \in \mathbb{R}^n$ is explicitly given by*

$$p^2 = \mathbb{1}_n p_0^2 - g(b),$$

in which the square has to be understood component-by-component.

With the explicit representation of $p \in \mathbb{R}^n$ we can get a characterization of the feasible set M :

Lemma 3.7. *A load vector $b \in \mathbb{R}_{\geq 0}^n$ is feasible, i.e. $b \in M$, iff the following system of inequalities hold:*

$$\begin{aligned} p_0^2 &\leq \min_{k=1, \dots, n} [(p_k^{\max})^2 + g_k(b)], \\ p_0^2 &\geq \max_{k=1, \dots, n} [(p_k^{\min})^2 + g_k(b)]. \end{aligned}$$

The fact that a load vector with negative components cannot be feasible comes from the application. A gas consumer cannot deliver gas to the network, the consumer can either consume gas (value in the load vector is positive) or not consume gas (value in the load vector is 0).

Proof of Lemma 3.7.

„ \Rightarrow “: Consider a feasible load vector $b \in \mathbb{R}_{\geq 0}^n$, i.e. $b \in M$. Then the solution $(p, q) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfies the pressure bounds $p \in [p^{\min}, p^{\max}]$. From Corollary 3.6 it follows

$$p^2 = p_0^2 - g(b),$$

and thus

$$p_0^2 = p_k^2 + g_k(b).$$

Since the pressure bounds hold we can estimate p_k^2 from above by $(p_k^{\max})^2$ and from below by $(p_k^{\min})^2$, so the inequalities in *Lemma 3.7* directly follow. „ \Leftarrow “: Consider $b \in \mathbb{R}_{\geq 0}^n$, that satisfies the inequalities in *Lemma 3.7*. Then for all $k \in \{1, \dots, n\}$ it follows

$$p_0^2 \leq (p_k^{\max})^2 + g_k(b), \quad p_0^2 \geq (p_k^{\min})^2 + g_k(b),$$

and this is equal to

$$p_k^2 \leq (p_k^{\max})^2 \quad \text{and} \quad p_k^2 \geq (p_k^{\min})^2.$$

Thus the pressure bounds are satisfied and we have $b \in M$. \square

With *Lemma 3.7* we have a characterization of the feasible set in the sense that for a given load vector, we only have to check two inequalities. So for a random load vector we now have information about both, the distribution of the random load vector and the set of feasible loads (cf. *Figure 3.1 (a)*).

Consider the n -dimensional Gaussian distributed random variable

$$\xi_b \sim \mathcal{N}(\mu, \Sigma),$$

with mean value $\mu \in \mathbb{R}_{\geq 0}^n$ and positive definite covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$ on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We identify the load vector $b \in \mathbb{R}^n$ with the image $\xi_b(\omega)$ for $\omega \in \Omega$ on this probability space. Let $L \in \mathbb{R}^{n \times n}$ be a Cholesky decomposition of Σ , i.e.

$$\Sigma = LL^\top,$$

and let $v \in \mathbb{S}^{n-1}$ be a point on the unit sphere. We set

$$b_v(r) := rLv + \mu.$$

We need to find the range of r , s.t. $b_v(r)$ is in M . The feasible set M has the additional restriction, that negative loads are not feasible. So we define the regular range

$$R_{v,\text{reg}} := \{r \geq 0 \mid b_v(r) \geq 0\}.$$

Now we need to find the feasible parts of the regular range, which we denote M_v . *Lemma 3.7* gives us a characterization of the feasible set and we can use the inequalities (3.7) to compute the feasible parts of the regular range. For $r \in R_{v,\text{reg}}$, $b_v(r)$ is not negative and thus, the absolute value in $g(b_v(\cdot))$ vanishes. It follows

$$g(b(r)) = (A^\top)^{-1} \Phi(A^{-1}b_v(r))^2.$$

Thus we can compute the one dimensional feasible set M_v by

$$M_v = \left\{ r \in R_{v,\text{reg}} \left| \begin{array}{l} p_0^2 \leq \min_{k=1,\dots,n} [(p_k^{\max})^2 + g_k(b(r))] \\ p_0^2 \geq \max_{k=1,\dots,n} [(p_k^{\min})^2 + g_k(b(r))] \end{array} \right. \right\}. \quad (3.12)$$

Since the $g_k(b(r))$ are quadratic functions in r , both inequalities are also quadratic in r and thus both inequalities provide at most two disjoint intervals for r to be feasible. Now the intersection of the intervals leads to a representation of M_v as union of $\ell_v \in \mathbb{N}$ disjoint intervals, i.e.

$$M_v = \bigcup_{j=1}^{\ell_v} I_{v,j}, \quad (3.13)$$

with intervals $I_{v,j} = [\underline{a}_{v,j}, \bar{a}_{v,j}]$ and interval bounds $\underline{a}_{v,j}, \bar{a}_{v,j} \in \mathbb{R}_{\geq 0}$, $\underline{a}_{v,j} < \bar{a}_{v,j}$ ($j = 1, \dots, \ell_v$).

Remark 3.8. *The interval bounds cannot be negative due to the restriction $r \geq 0$ in the regular range. Further, we can write the inequalities in (3.12) for $k \in \{1, \dots, n\}$ as*

$$\begin{array}{rcl} 0 & \leq & (p_k^{\max})^2 + g_k(b(r)) - p_0^2, \\ 0 & \leq & p_0^2 - g_k(b(r)) - (p_k^{\min})^2. \end{array}$$

One can see that the quadratic term has a positive sign in the first inequality and a negative sign in the second inequality and thus it follows

$$\bar{a}_{v,j} < \infty,$$

for all $j = 1, \dots, \ell_v$, $v \in \mathbb{S}^{n-1}$.

With the representation of the one-dimensional set M_v as union of disjoint intervals (3.13), it follows

$$\mu_\chi(M_v) = \mu_\chi \left(\bigcup_{j=1}^{\ell_v} I_{v,j} \right) = \sum_{j=1}^{\ell_v} F_\chi(\bar{a}_{v,j}) - F_\chi(\underline{a}_{v,j}),$$

where F_χ is the cumulative distribution function of the χ -distribution with n degrees of freedom. The probability for a random load vector to be feasible then is

$$\mathbb{P}(b \in M) = \int_{\mathbb{S}^{n-1}} \sum_{j=1}^{\ell_v} F_\chi(\bar{a}_{v,j}) - F_\chi(\underline{a}_{v,j}) d\mu_v(v).$$

For the numerical computations we use *Algorithm 1.13*. We call the resulting probability \mathbb{P}_{SRD} . For $N_{\text{SRD}} \in \mathbb{N}$ let

$$\mathcal{S} := \{v_1, \dots, v_{N_{\text{SRD}}}\} \subseteq \mathbb{S}^{n-1},$$

be a uniformly distributed sampling on the sphere \mathbb{S}^{n-1} . Due to *Algorithm 1.13* we have

$$\mathbb{P}_{\text{SRD}}(b \in M) = \frac{1}{N_{\text{SRD}}} \sum_{v \in \mathcal{S}} \sum_{j=1}^{\ell_v} F_{\chi}(\bar{a}_{v,j}) - F_{\chi}(\underline{a}_{v,j}). \quad (3.14)$$

As it is mentioned in *Section 1.2*, it follows

$$\mathbb{P}_{\text{SRD}}(b \in M) \xrightarrow{N_{\text{SRD}} \rightarrow \infty} \mathbb{P}(b \in M).$$

Thus with (3.14), we have an efficient and well-working way to compute the probability for a random load vector to be feasible resp. to answer the question $\mathcal{Q}_{\text{prob}}$.

3.3. Application of the KDE

In this section, we use the KDE (introduced in *Section 1.1*) to compute the probability for a random load vector to be feasible where the feasible set is defined in (3.5).

Consider the n -dimensional Gaussian distributed random variable

$$\xi_b \sim \mathcal{N}(\mu, \Sigma),$$

with mean value $\mu \in \mathbb{R}_{\geq 0}^n$ and positive definite covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$ on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We identify the load vector $b \in \mathbb{R}^n$ with the image $\xi_b(\omega)$ for $\omega \in \Omega$ on this probability space. As it is mentioned in (3.8), instead of directly computing the probability for a random load vector to be feasible (like we did in *Section 3.2*), we can compute the probability for the resulting random pressures to satisfy the box constraints (see *Figure 3.1 (b)*). The pressure bounds define a n -dimensional cuboid, but the distribution of the pressure is not known. We use the KDE introduced in *Section 1.1* to estimate the probability density function of the pressures at the nodes. Then we can compute the probability for the random pressures to satisfy the box constraints by integrating the estimated probability density over the n -dimensional cuboid defined by the pressure bounds.

Due to the fact that the load vector is random, the vector of pressures at the nodes $p \in \mathbb{R}^n$ is also random. We assume that p has an absolutely continuous distribution function with probability density function ϱ_p . Further we assume that μ and Σ are chosen s.t. b is almost surely non negative, i.e $\mathbb{P}(b \geq 0) \approx 1$. This assumption is motivated by the application. A gas consumer can neither produce gas nor simply return the demanded gas to the network, so the load for this consumer cannot be negative. This assumption is also necessary for our computation as we will see later.

The probability that the random pressures at the nodes satisfy the box constraints (which is equal to the probability for a random load vector to be feasible) can be computed by

$$\mathbb{P}(p \in P_{\min}^{\max}) = \int_{P_{\min}^{\max}} \varrho_p(z) dz.$$

Since ϱ_p is unknown, we approximate it by using the KDE approach introduced in *Section 1.1*.

For $N_{\text{KDE}} \in \mathbb{N}$ let $\mathcal{B} = \{b^{S,1}, \dots, b^{S,N_{\text{KDE}}}\} \subseteq \mathbb{R}_{\geq 0}^n$ be independent and identically distributed (positive) samples of the random variable ξ_b . Further let $\mathcal{P}_{\mathcal{B}} = \{p(b^{S,1}), \dots, p(b^{S,N_{\text{KDE}}})\} \subseteq \mathbb{R}^n$ be the corresponding pressures at the nodes for the different loads $b^{S,i} \in \mathcal{B}$ ($i = 1, \dots, N_{\text{KDE}}$). Then the $p(b^{S,i})$ are also independent and identically distributed. Let $\Sigma_{N_{\text{KDE}}}$ be the covariance matrix of the sampling $\mathcal{P}_{\mathcal{B}}$. We use the KDE introduced in (1.6) with Gaussian kernel and diagonal bandwidth matrix $H_{i,i} = h^2(\Sigma_{N_{\text{KDE}}})_{i,i}$ (h defined in (1.7)) to get an approximation of the probability density function $\varrho_{p,N_{\text{KDE}}}$. We have

$$\varrho_{p,N_{\text{KDE}}}(z) = \frac{1}{N_{\text{KDE}} \det(H)^{\frac{1}{2}}} \sum_{i=1}^{N_{\text{KDE}}} \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2} \left\| H^{-\frac{1}{2}} (z - p(b^{S,i})) \right\|_2^2\right).$$

Here $\|\cdot\|_2$ is the Euklclidean norm. We set $h_i^2 := H_{i,i}$. Due to the diagonality of H it follows

$$\varrho_{p,N_{\text{KDE}}}(z) = \frac{1}{N_{\text{KDE}} \prod_{j=1}^n h_j} \sum_{i=1}^{N_{\text{KDE}}} \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2} \sum_{j=1}^n \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right),$$

which is equal to the KDE with Gaussian product kernel

$$\varrho_{p,N_{\text{KDE}}}(z) = \frac{1}{N_{\text{KDE}} \prod_{j=1}^n h_j} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right). \quad (3.15)$$

From Section 1.1 we know that

$$\varrho_{p, N_{\text{KDE}}} \xrightarrow{N_{\text{KDE}} \rightarrow \infty} \varrho_p \quad \mathbb{P} - \text{almost surely.}$$

Now we integrate the approximated probability density over the set of feasible pressures. It follows

$$\mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) = \int_{P_{\min}^{\max}} \varrho_{p, N_{\text{KDE}}}(z) dz, \quad (3.16)$$

and since we assumed $\mathbb{P}(b \geq 0) \approx 1$, we have

$$\mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) \xrightarrow{N_{\text{KDE}} \rightarrow \infty} \mathbb{P}(b \in M) \quad \mathbb{P} - \text{almost surely.}$$

Remark 3.9. *If we allow negative values in \mathcal{B} , problems in the theory can occur. Then the computed probability $\mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max})$ only guarantees that the pressure bounds hold but not that the load is non negative. Due to the definition of the feasible set we need to guarantee that too. When we only take the positive elements of \mathcal{B} and compute the probability via a KDE approach, then this leads to the conditional probability*

$$\mathbb{P}_{b \geq 0}(p \in P_{\min}^{\max}).$$

Then the probability that a random load vector is be feasible can be computed by

$$\mathbb{P}(b \geq 0 \text{ and } p \in P_{\min}^{\max}) = \mathbb{P}_{b \geq 0}(p \in P_{\min}^{\max}) \mathbb{P}(b \geq 0).$$

The stated method for computing the probability for a random load vector to be feasible via a KDE approach is much more general than using the SRD, it also holds for other distributions like a truncated Gaussian distribution or a log-normal distribution. Both do not allow negative values and both are also proposed in [Koch et al., 2015] (like the Gaussian distribution) for modelling the consumers gas demand.

For N_{KDE} sufficiently large, (3.16) gives a sufficiently good approximation for a random load vector to be feasible. We now give an efficient way how to compute the integral in (3.16). We have

$$\begin{aligned} & \mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) = \\ & = \int_{P_{\min}^{\max}} \frac{1}{N_{\text{KDE}} \prod_{j=1}^n h_j} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right) dz. \end{aligned}$$

This is equal to

$$\begin{aligned} & \mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) = \\ &= \frac{1}{N_{\text{KDE}} \prod_{j=1}^n h_j} \sum_{i=1}^{N_{\text{KDE}}} \int_{p_{\min}^{\max}} \prod_{j=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right) dz. \end{aligned}$$

Due to the continuity of $\varrho_{p, N_{\text{KDE}}}$ and the compactness of P_{\min}^{\max} , $\varrho_{p, N_{\text{KDE}}}$ is integrable and we can apply *Fubini's theorem* (see [Elstrodt, 2011]). Thus we have

$$\begin{aligned} & \mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) = \\ &= \frac{1}{N_{\text{KDE}} \prod_{j=1}^n h_j} \sum_{i=1}^{N_{\text{KDE}}} \int_{p_1^{\min}}^{p_1^{\max}} \cdots \int_{p_n^{\min}}^{p_n^{\max}} \prod_{j=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right) dz_n \cdots dz_1. \end{aligned}$$

Since every factor of the product in $\varrho_{p, N_{\text{KDE}}}$ depends only on one dimension, we can swap the integral and the product. It follows

$$\begin{aligned} & \mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) = \\ &= \frac{1}{N_{\text{KDE}} \prod_{j=1}^n h_j} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right) dz_j. \end{aligned}$$

We define functions

$$\varphi_{i,j} : \mathbb{R} \rightarrow \mathbb{R}, \quad \varphi_{i,j} : x \mapsto \left(\frac{x - p_j(b^{S,i})}{\sqrt{2} h_j}\right),$$

so we have

$$\begin{aligned} & \mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) = \\ &= \frac{1}{N_{\text{KDE}} \prod_{j=1}^n h_j} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp(-\varphi_{i,j}^2(z_j)) dz_j. \end{aligned}$$

We set $\tau_{i,j} := \varphi_{i,j}(z_j)$ and use integration by substitution. Then with

$$\frac{d}{dx} \varphi_{i,j}(x) = \frac{1}{\sqrt{2} h_j},$$

it follows

$$\begin{aligned} & \mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) = \\ &= \frac{1}{N_{\text{KDE}} \prod_{j=1}^n h_j} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^n \int_{\varphi_{i,j}(p_j^{\min})}^{\varphi_{i,j}(p_j^{\max})} \frac{1}{\sqrt{2\pi}} \exp(-\tau_{i,j}^2) \sqrt{2} h_j d\tau_{i,j}, \end{aligned}$$

which is equal to

$$\mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) = \frac{1}{N_{\text{KDE}}} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^n \int_{\varphi_{i,j}(p_j^{\min})}^{\varphi_{i,j}(p_j^{\max})} \frac{1}{\sqrt{\pi}} \exp(-\tau_{i,j}^2) d\tau_{i,j}. \quad (3.17)$$

The integral in (3.17) is closely related to the *Gauss error function* (see e.g. [Andrews, 1998]):

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt. \quad (3.18)$$

We have

$$\int_{\varphi_{i,j}(p_j^{\min})}^{\varphi_{i,j}(p_j^{\max})} \frac{1}{\sqrt{\pi}} \exp(-\tau_{i,j}^2) d\tau_{i,j} = \frac{1}{2} \left[\text{erf}(\varphi_{i,j}(p_j^{\max})) - \text{erf}(\varphi_{i,j}(p_j^{\min})) \right],$$

and thus it follows

$$\begin{aligned} \mathbb{P}_{\text{KDE}}(p \in P_{\min}^{\max}) &= \\ &= \frac{1}{N_{\text{KDE}} 2^n} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^n \left[\text{erf}(\varphi_{i,j}(p_j^{\max})) - \text{erf}(\varphi_{i,j}(p_j^{\min})) \right]. \end{aligned} \quad (3.19)$$

The Gauss error function (3.18) is well established in mathematics and physics for decades, so we make use of its theory later in optimization. Furthermore almost all programming languages contain precise approximations of the Gauss error function, which increases the speed of computation enormously, especially in the numerical examples in the next section. Thus for N_{KDE} sufficiently large, (3.19) provides an efficient and fast way for computing the probability that a random load vector is be feasible.

3.4. Numerical Examples

In this section we illustrate the results of the last sections on three easy example graphs. Throughout this section we assume that all variables and constants are non-dimensionalized. All implementations are done in *MATLAB*[®] 2019a but they also work with *MATLAB*[®] 2015a.

3.4.1. A Single Edge

Consider a graph that consists in a single edge, i.e. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with

$$\mathcal{V} = \{v_0, v_1\},$$

$$\text{and } \mathcal{E} = \{e_1\} = \{(v_0, v_1)\},$$

which is shown in *Figure 3.2*. All constants and values (without units) are stated in *Table 3.1*.

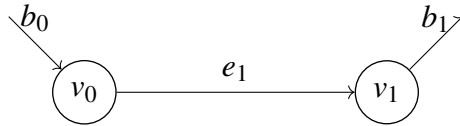


Figure 3.2.: Graph with a single edge

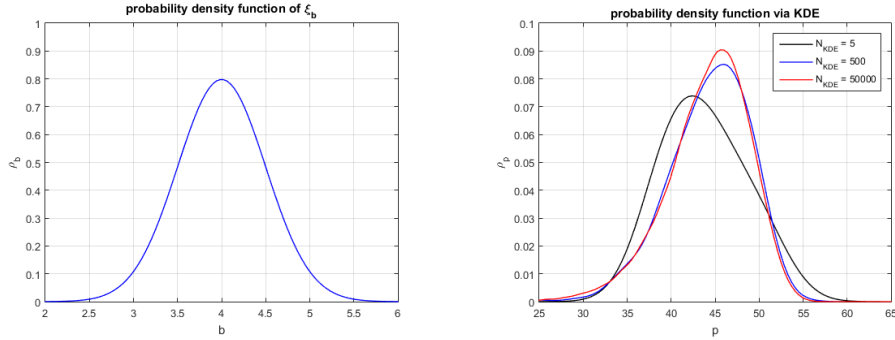
p_0	p^{\min}	p^{\max}	μ	σ	ϕ
60	40	60	4	0.5	100

Table 3.1.: Values for the example with one edge

The probability density function of the random variable $\xi_b \sim \mathcal{N}(\mu, \sigma)$ is shown in *Figure 3.3 (a)*. Further the probability density function of the pressure (computed with (3.15)) is shown in *Figure 3.3 (b)* for different sampling lengths ($N_{\text{KDE}} = 5$ (black), $N_{\text{KDE}} = 500$ (blue) and $N_{\text{KDE}} = 50000$ (red)). One can see, that the probability density functions $\varrho_{p, N_{\text{KDE}}}$ are similar for the different N_{KDE} . At least for $N_{\text{KDE}} = 500$ and for $N_{\text{KDE}} = 50000$ there is almost no difference.

We first compute the probability for a random vector to be feasible using the SRD as it is explained in *Section 3.2*. For $b \in \mathbb{R}_{\geq 0}$ and $A = A^\top = A^{-1} = [1]$, *Lemma 3.7* implies

$$0 \leq b \leq \sqrt{20}.$$



(a) probability density function of ξ_b (plotted with 100 points) (b) probability density function ρ_p for different values of N_{KDE}

Figure 3.3.: Probability density functions of the random variable ξ_b and the corresponding random pressure ρ_p for different sampling length

Since we have only one outflow, the random variable is one dimensional and the $(n - 1)$ -dimensional sphere is given by

$$\mathbb{S}^{n-1} = \{-1, 1\}.$$

For $b(r) = r\sigma v + \mu$, from (3.12) we get

$$r \in \left[0, \sqrt{80} + 8\right] \quad \text{for } v = -1,$$

and

$$r \in \left[0, \sqrt{80} - 8\right] \quad \text{for } v = 1.$$

Then due to (3.14) it follows

$$\mathbb{P}_{\text{SRD}}(b \in M) = 82,75\%.$$

Since the $(n - 1)$ -dimensional sphere contains only two values, we choose the sphere itself as sampling and thus, the probability computed via SRD is exact (apart from numerical errors due to the evaluation of the χ -distribution).

Next let $\mathcal{B} = \{b^{S,1}, \dots, b^{S,N_{\text{KDE}}}\} \subseteq \mathbb{R}$ be a sampling of

$$N_{\text{KDE}} \in \{5, 500, 50000\},$$

independent and identically $\mathcal{N}(\mu, \sigma)$ -distributed values. Let $\mathcal{P}_{\mathcal{B}}$ be the sampling of the corresponding pressures. We compare the (exact) probability computed via SRD with both, the probability computed via KDE using (3.19) and the probability computed by a classical Monte-Carlo (MC) method.

The MC probability is the percentage of pressures in $\mathcal{P}_{\mathcal{B}}$ that satisfies the pressure bounds. For both, the MC computation and the KDE computation, we use the same sampling \mathcal{B} . All results are listed in *Table 3.2*.

Method	N_{KDE}	Test 1	Test 2	Test 3	Test 4	Test 5	\varnothing
MC	5	80,00%	80,00%	80,00%	80,00%	60,00%	76,00%
	500	87,00%	83,60%	83,60%	81,20%	80,60%	83,20%
	50000	82,93%	82,89%	83,00%	82,78%	82,76%	82,87%
KDE	5	80,93%	79,94%	79,20%	81,06%	79,35%	80,10%
	500	82,72%	81,06%	81,14%	83,29%	80,11%	82,66%
	50000	82,70%	82,72%	82,70%	82,79%	82,64%	82,71%
SRD		82,75%					

Table 3.2.: Probabilities of the example with one edge

One can see that in all three cases, that the mean KDE probability is a bit closer to the (exact) SRD probability. The variances of the 5 tests in the MC computation are $80.0000 \cdot 10^{-4}$ for 5 samples, $6.3800 \cdot 10^{-4}$ for 500 samples and $0.0103 \cdot 10^{-4}$ for 50000 samples. For the KDE computation, the variances are $0.7521 \cdot 10^{-4}$ for 5 samples, $1.7033 \cdot 10^{-4}$ for 500 samples and $0.0029 \cdot 10^{-4}$ for 50000 samples. Again the KDE probability leads to a smaller variance between the tests. The computation time is negligible. For all 15 MC probabilities the computation time was 0.0025s and for the KDE probabilities it was 0.0054, twice as much. But for small computation times like this we cannot make a statement about the speed of computation because a simple computer background process can influence these small computation times. One can only see that both methods are quite fast.

3.4.2. A Network with two parallel edges

Consider a tree-structured graph with two parallel edges, i.e. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with

$$\mathcal{V} = \{v_0, v_1, v_2\},$$

$$\text{and } \mathcal{E} = \{e_1, e_2\} = \{(v_0, v_1), (v_0, v_2)\}.$$

The graph is shown in *Figure 3.4*. All constants and values for this example are given in *Table 3.3*. As in the last subsection, we assume that all values are non-dimensionalized.

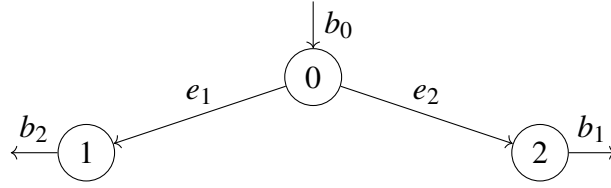
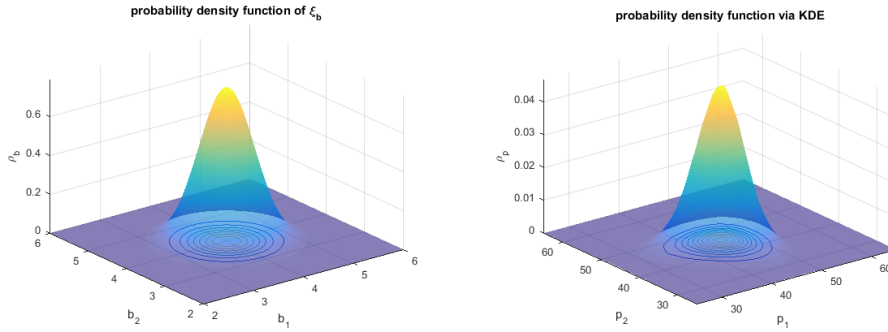


Figure 3.4.: Tree-structured graph with two parallel edges

p_0	p^{\min}	p^{\max}	μ	Σ	ϕ
60	$\begin{pmatrix} 40 \\ 30 \end{pmatrix}$	$\begin{pmatrix} 60 \\ 50 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0 \\ 0 & 0.2 \end{pmatrix}$	$\begin{pmatrix} 100 \\ 100 \end{pmatrix}$

Table 3.3.: Values for the example with two parallel edges

The probability density function of the random variable ξ_b is shown in *Figure 3.5 (a)*. The estimated probability density function of the pressure (computed with (3.15)) is shown in *Figure 3.5 (b)* for $5 \cdot 10^5$ samples.



(a) probability density function of ξ_b (plotted with 100×100 points) (b) probability density function ϱ_p for $N_{\text{KDE}} = 5 \cdot 10^5$

Figure 3.5.: Probability density functions of the random variable ξ_b and the corresponding random pressure ϱ_p for $5 \cdot 10^5$ samples

We compute the probability for a random vector to be feasible using the SRD approach. From *Lemma 3.7* we have

$$\begin{bmatrix} 0 \\ \sqrt{11} \end{bmatrix} \leq \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \leq \begin{bmatrix} \sqrt{20} \\ \sqrt{27} \end{bmatrix}.$$

The unit sphere \mathcal{S}^1 is not finite in this example, so we compute the probability via SRD for the different sampling lengths $N_{\text{SRD}} \in \{5 \cdot 10^1, 5 \cdot 10^3, 5 \cdot 10^5\}$.

Let $\mathcal{S} = \{v_1, \dots, v_{N_{\text{SRD}}}\}$ be a sampling of N_{SRD} uniformly distributed points on the sphere \mathbb{S}^1 . We use (3.14) to compute the desired probability. Let $\mathcal{B} = \{b^{S,1}, \dots, b^{S,N_{\text{KDE}}}\} \subseteq \mathbb{R}^2$ be a sampling of $N_{\text{KDE}} \in \{5 \cdot 10^1, 5 \cdot 10^3, 5 \cdot 10^5\}$ independent and identically $\mathcal{N}(\mu, \Sigma)$ -distributed points. Let $\mathcal{P}_{\mathcal{B}}$ be the sampling of the corresponding pressures. We compute the desired probability via KDE with (3.19). Further we compare the probabilities computed via KDE and SRD with the probabilities computed by a classical Monte-Carlo (MC) method. For the MC method we also use the samplings \mathcal{B} for the different N_{KDE} . All results are shown in *Table 3.4*.

Method	N_{KDE}	Test 1	Test 2	Test 3	Test 4	Test 5	\varnothing
MC	$5 \cdot 10^1$	82,00%	80,00%	80,00%	82,00%	76,00%	80,00%
	$5 \cdot 10^3$	79,36%	80,00%	79,86%	80,36%	81,04%	80,12%
	$5 \cdot 10^5$	79,61%	79,71%	79,81%	79,80%	79,83%	79,75%
KDE	$5 \cdot 10^1$	74,09%	75,41%	77,34%	71,69%	69,54%	73,61%
	$5 \cdot 10^3$	78,21%	78,76%	78,46%	79,04%	79,22%	78,74%
	$5 \cdot 10^5$	79,32%	79,42%	79,50%	79,51%	79,50%	79,45%
SRD	$5 \cdot 10^1$	79.71%	79.71%	79.73%	79.73%	79.70%	79.72%
	$5 \cdot 10^3$	79,72%					
	$5 \cdot 10^5$	79,72%					

Table 3.4.: Probabilities of the example with two parallel edges

The reference value in *Table 3.4* is the SRD probability for $5 \cdot 10^5$ sampling points. Since the results in all SRD tests for $5 \cdot 10^3$ and $5 \cdot 10^5$ were equal (rounded to 4 decimal places) we only wrote the result once.

Here (in contrary to the example with one edge of the previous section) the MC probability is a bit closer to the (almost exact) SRD probability. The variances of the 5 tests in the MC computation are $0.6000 \cdot 10^{-3}$ for $5 \cdot 10^1$ samples, $0.0391 \cdot 10^{-3}$ for $5 \cdot 10^3$ samples and $0.0008 \cdot 10^{-3}$ for $5 \cdot 10^5$ samples. For the KDE computation, the variances are $0.9413 \cdot 10^{-3}$ for $5 \cdot 10^1$ samples, $0.0171 \cdot 10^{-3}$ for $5 \cdot 10^3$ samples and $0.0007 \cdot 10^{-3}$ for $5 \cdot 10^5$ samples. The variances of the SRD computations are negligible.

We mention again that we only use a rule-of-thumb formula for the bandwidth, with the exact bandwidth or even more samples, the result could be even better. The computation time for all 15 MC probabilities was 0.0331s

and for the KDE probabilities it was $0.0971s$ while the computation time for the 15 SRD probabilities was almost two minutes. The implementation can surely be improved. Although the SRD provides good results for much less sampling points, its computation needs much more time.

3.4.3. A Network with two linear edges

Consider a linear graph with two edges, i.e. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with

$$\mathcal{V} = \{v_0, v_1, v_2\},$$

$$\text{and } \mathcal{E} = \{e_1, e_2\} = \{(v_0, v_1), (v_1, v_2)\}.$$

The graph is shown in *Figure 3.6*. All (non-dimensionalized) constants and values for this example are given in *Table 3.5*.

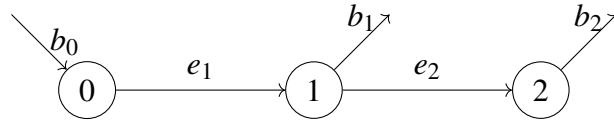


Figure 3.6.: Tree-structured graph with two linear edges

p_0	p^{\min}	p^{\max}	μ	Σ	ϕ
60	$\begin{pmatrix} 40 \\ 30 \end{pmatrix}$	$\begin{pmatrix} 60 \\ 50 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 1.5 \end{pmatrix}$	$\begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$	$\begin{pmatrix} 100 \\ 100 \end{pmatrix}$

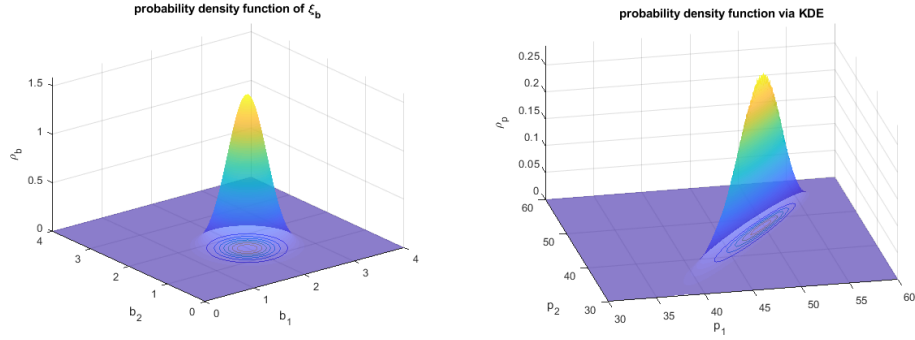
Table 3.5.: Values for the example with two linear edges

The probability density function of the Gaussian distributed random variable ξ_b is shown in *Figure 3.7 (a)*. The estimated probability density function of the pressure (computed with (3.15)) is shown in *Figure 3.7 (b)* for $5 \cdot 10^5$ samples.

We first compute the probability for a random load vector to be feasible using the SRD approach. From *Lemma 3.7* we have

$$b_2 \geq \frac{1}{2} \left(\sqrt{22 - b_1^2} - b_1 \right),$$

$$\text{and } b_2 \leq \min \left\{ \sqrt{20} - b_1, \frac{1}{2} \left(\sqrt{54 - b_1^2} - b_1 \right) \right\}.$$



(a) probability density function of ξ_b (plotted with 100×100 points) (b) probability density function ϱ_p for $N_{\text{KDE}} = 5 \cdot 10^5$

Figure 3.7.: Probability density functions of the random variable ξ_b and the corresponding random pressure ϱ_p for $5 \cdot 10^5$ samples

We compute the probability via SRD for the different sampling lengths $N_{\text{SRD}} \in \{5 \cdot 10^1, 5 \cdot 10^3, 5 \cdot 10^5\}$. Let $\mathcal{S} = \{v_1, \dots, v_{N_{\text{SRD}}}\}$ be a sampling of N_{SRD} uniformly distributed points on the sphere \mathbb{S}^1 . We use (3.14) to compute the desired probability. Let $\mathcal{B} = \{b^{S,1}, \dots, b^{S,N_{\text{KDE}}}\} \subseteq \mathbb{R}^2$ be a sampling of $N_{\text{KDE}} \in \{5 \cdot 10^1, 5 \cdot 10^3, 5 \cdot 10^5\}$ independent and identically $\mathcal{N}(\mu, \Sigma)$ -distributed points. Let $\mathcal{P}_{\mathcal{B}}$ be the sampling of the corresponding pressures. We compute the desired probability via KDE approach (using (3.19)). Further we compare the probabilities computed via KDE and SRD with the probability computed by a classical Monte-Carlo (MC) method. For the MC method we also use the samplings \mathcal{B} for the different N_{KDE} . All results are shown in *Table 3.6*.

The reference value in *Table 3.6* is the SRD probability whereat the SRD probabilities are equal in at least 4 digits for all sampling lengths. The results for the KDE and for the MC method show the same as in the previous example: Both methods provide quite good results, where the MC probability is a bit closer to the reference value (the SRD probability). The variances of the 5 tests in the MC computation are $19.2000 \cdot 10^{-4}$ for $5 \cdot 10^1$ samples, $0.4170 \cdot 10^{-4}$ for $5 \cdot 10^3$ samples and $0.0069 \cdot 10^{-4}$ for $5 \cdot 10^5$ samples. For the KDE computation the variances are $10.0438 \cdot 10^{-4}$ for $5 \cdot 10^1$ samples, $0.3706 \cdot 10^{-4}$ for $5 \cdot 10^3$ samples and $0.0083 \cdot 10^{-4}$ for $5 \cdot 10^5$ samples. For the SRD computation the variances are negligible.

The computation time for all 15 MC probabilities was $0.0084s$ and for the KDE probabilities it was $0.0883s$, while the computation time for

Method	N_{KDE}	Test 1	Test 2	Test 3	Test 4	Test 5	\varnothing
MC	$5 \cdot 10^1$	88,00%	84,00%	78,00%	78,00%	78,00%	81,20%
	$5 \cdot 10^3$	82,52%	81,78%	82,40%	82,48%	82,64%	82,36%
	$5 \cdot 10^5$	82,15%	82,21%	82,10%	82,25%	82,25%	82,19%
KDE	$5 \cdot 10^1$	80,13%	82,26%	78,59%	75,55%	75,32%	78,37%
	$5 \cdot 10^3$	81,37%	80,65%	81,49%	81,34%	81,54%	81,28%
	$5 \cdot 10^5$	81,93%	81,99%	81,90%	82,02%	82,03%	81,97%
SRD	$5 \cdot 10^1$	82,21%					
	$5 \cdot 10^3$	82,21%					
	$5 \cdot 10^5$	82,21%					

Table 3.6.: Probabilities of the example with two linear edges

the 15 SRD probabilities was about two minutes. Surely the implementation can be improved but the differences in the computation times are enormous even if the SRD probability for $5 \cdot 10^1$ samples is almost the same as the SRD probability for $5 \cdot 10^5$ samples.

3.5. A Model Extension: Unknown Inlet Pressure and Compressor Control

In this section we extend the model from *Section 3.1* by an unknown inlet pressure and by compressor control. We show why the KDE approach from *Section 3.3* cannot be used for this model but the SRD used in *Section 3.2* can be applied here. The results of this section are already published in *Mathematical Problems in Engineering* (see [Gugat and Schuster, 2018]).

Consider a connected, directed, tree-structured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the vertex set $\mathcal{V} = \{v_0, \dots, v_n\}$ and the set of edges $\mathcal{E} = \{e_1, \dots, e_n\}$ with $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Due to the structure, the graph has a unique root v_0 . An edge can either be a flux edge, so the pressure decreases along the edge (caused by friction) or a compressor edge, so the pressure increases along the edge (caused by pressure compression). We define \mathcal{E}_F as the set of all flux edges and \mathcal{E}_C as the set of all compressor edges, where $\mathcal{E} = \mathcal{E}_F \cup \mathcal{E}_C$ and $\mathcal{E}_F \cap \mathcal{E}_C = \emptyset$. We set $|\mathcal{E}_F| = n_F$ and $|\mathcal{E}_C| = n_C$, where $n_F + n_C = n$.

On every flux edge $e \in \mathcal{E}_F$ the stationary semi-linear isothermal Euler equations for horizontal pipes and ideal gases (sISOstat) hold. As in *Section 3.1* we will later consider pressure bounds at the nodes of the network graph. Since the solution of (sISOstat) is constant in flow and monotonously decreasing in pressure we are only interested in the pressures at the nodes which leads to the so-called *Weymouth* equation (see e.g. [Gugat et al., 2018])

$$(p^e(0))^2 - (p^e(L^e))^2 = \phi^e |q^e| q^e \quad \text{with} \quad \phi^e = (R_S T)^2 \frac{\lambda^e L^e}{D^e (c^e)^2}, e \in \mathcal{E}_F. \quad (3.20)$$

Here, R_S is the specific gas constant, T is the temperature of the gas, λ^e is the pipe friction coefficient, D^e is the pipe diameter, c^e is the speed of sound in the gas and L^e is the length of pipe $e \in \mathcal{E}_F$.

The compressor stations counteract the pressure drop along the pipes caused by friction. Modelling a compressor station in detail is not trivial, one can fill a whole thesis with this topic. In this thesis we use a simple multiplicative compressor model that is suggested in [Domschke et al., 2017, Section 7.4.7] and [Rose et al., 2016, Section 2.3]:

$$\frac{\tilde{c}}{\gamma} \left(\left(\frac{p_{\text{out}}}{p_{\text{in}}} \right)^\gamma - 1 \right) = \tilde{u}.$$

Here p_{out} and p_{in} are the pressures at the end resp. at the beginning of the pipe, \tilde{c} and γ are constants and \tilde{u} , which is the control in our setting, is the specific change in adiabatic enthalpy. More information about the parameters can be found in [Rose et al., 2016]. We also refer to [Koch et al., 2015, Section 2.3.5] for a general overview about compressor machines and drives. So for every compressor edge $e \in \mathcal{E}_C$ we assume

$$\left(\frac{p^e(L^e)}{p^e(0)} \right)^2 = u^e \quad \text{with} \quad u^e = \left(\frac{\gamma}{\tilde{c}} \tilde{u} + 1 \right)^\frac{2}{\gamma}. \quad (3.21)$$

It is $u^e \geq 1$ where $u^e = 1$ means that the compressor station is switched off.

Another prevalent way for modelling a compressor station is that one assumes a fixed parameter u^e and an integer variable decides whether the compressor is switched on or off. But this would lead to *Mixed Integer Problems* and the *Mixed Integer Programming* theory is not part of this thesis.

Let $b \in \mathbb{R}_{\geq 0}^n$ be the load vector, i.e. the vector of gas outflow corresponding to the vertices v_1, \dots, v_n . Further let $p^+ \in \mathbb{R}^{n+1}$ denote the vector of pressures at the nodes v_0, \dots, v_n . We define pressure bounds $p^{+, \min}, p^{+, \max} \in \mathbb{R}_{\geq 0}^{n+1}$

with $p^{+, \min} \leq p^{+, \max}$ component-by-component. Last let $q \in \mathbb{R}^n$ be the vector of the (constant) flows along the edges e_1, \dots, e_n .

We assume conservation of mass at every node $v \in \mathcal{V} \setminus \mathcal{V}_0$ which is stated in (3.2) and continuity in pressure at every node, i.e. (3.4) holds for all $v \in \mathcal{V}$.

Now we define the set of feasible loads. A load vector $b \in \mathbb{R}^n$ is feasible if we can find a corresponding flow vector $q \in \mathbb{R}^n$ and a corresponding pressure vector $p^+ \in \mathbb{R}^{n+1}$, s.t. the stationary isothermal Euler equations (sISOstat) hold on every flow edge $e \in \mathcal{E}_F$, the compressor property (3.21) holds on every compressor edge $e \in \mathcal{E}_C$, the coupling conditions (conservation on mass and continuity in pressure) hold at the nodes and the bounds for the pressure are satisfied. Thus we define the feasible set M as

$$M := \left\{ b \in \mathbb{R}_{\geq 0}^n \left| \begin{array}{l} \text{There exist } (p^+, q) \in \mathbb{R}^{n+1} \times \mathbb{R}^n \text{ s.t.} \\ (3.2), (3.4), (3.20), (3.21) \text{ are satisfied} \\ \text{and } p^+ \in [p^{+, \min}, p^{+, \max}] \end{array} \right. \right\}. \quad (3.22)$$

The difference of the feasible set M in this section (defined in (3.22)) and the set of feasible loads in *Section 3.1* (defined in (3.5)) is, that we cannot state an explicit solution here due to the unknown inlet pressure at the node v_0 . The question we want to answer in this section is basically the same as Q_{prob} but without the additional phrase „For a given inlet pressure“. So for a given load vector $b \in \mathbb{R}^n$, we first have to find a flow vector $q \in \mathbb{R}^n$ and a pressure vector $p^+ \in \mathbb{R}^{n+1}$ s.t. the desired properties are satisfied.

As it is explained in *Section 3.1*, the load vector (which is the consumers gas demand) is assumed to be random. So for a given mean value $\mu \in \mathbb{R}^n$ and a given positive definite covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$, we define a Gaussian distributed random variable $\xi_b \sim \mathcal{N}(\mu, \Sigma)$ on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We identify the load vector $b \in \mathbb{R}^n$ with the image $\xi_b(\omega)$ for $\omega \in \Omega$ on this probability space. Our aim in this section is to compute the probability, that a random load vector is feasible, i.e.

$$\mathbb{P}(b \in M).$$

Since the pressure at the node v_0 is not given a priori here we cannot apply the KDE as we did in *Section 3.3* because we neither can state an explicit solution nor compute a numerical solution without fixing at least one pressure at some node.

Nevertheless we can apply the SRD here if we slightly reformulate the problem. The clou in *Section 3.2* is *Lemma 3.7*, which characterizes the set of feasible loads. With this characterization we can simply check whether a load vector is feasible or not without computing the flows and the pressures first. We need a similar result in this section to compute the desired probability. The idea for the characterization of the feasible set in this section comes from [Gotzes et al., 2016]. For a tree-structured graph without compressor edges, i.e. $\mathcal{E}_C = \emptyset$, we can use the incidence matrix $A^+ \in \mathbb{R}^{(n+1) \times n}$ to get an expression for the pressure loss (cf. (3.20)) for the whole graph in matrix form, which is

$$(A^+)^{\top} (p^+)^2 = -\Phi |q|q. \quad (3.23)$$

The matrix $\Phi \in \mathbb{R}^{n \times n}$ is diagonal with diagonal entries $\Phi_{i,i} = \phi_i$ where $\phi_i = \phi^{e_i}$ defined in (3.20). The $(p^+)^2$ as well as the product $|q|q$ have to be understood component-by-component.

Due to the fact that $\mathcal{E}_C \neq \emptyset$ in our setting, we cannot directly use (3.23) here. The idea how to apply (3.23) to our setting is, roughly speaking, we remove the compressor edges from the graph. That means we split the graph at the compressor edges and treat the flows through the compressor edges as loads. The remaining graph is not connected anymore, it consists in $n_C + 1$ disjoint subgraphs. This procedure is shown in *Figure 3.8*. Mathematically that means we slightly change the formulation for the conservation of mass (3.2). For the notation we state $b_i = b^{v_i}$ and $p_i^+ = p^{v_i}$ for a node $v \in \mathcal{V} \setminus \mathcal{V}_0$. So for every node $v \in \mathcal{V} \setminus \mathcal{V}_0$ we have

$$\sum_{e \in \mathcal{E}_-(v)} q^e - \sum_{e \in \mathcal{E}_+(v)} q^e = b^v,$$

which is equal to

$$\sum_{e \in \mathcal{E}_-(v) \cap \mathcal{E}_F} q^e - \sum_{e \in \mathcal{E}_+(v) \cap \mathcal{E}_F} q^e = b^v - \sum_{e \in \mathcal{E}_-(v) \cap \mathcal{E}_C} q^e + \sum_{e \in \mathcal{E}_+(v) \cap \mathcal{E}_C} q^e. \quad (3.24)$$

We now identify a new load vector \tilde{b} with the right hand side of (3.24), i.e. for every $v \in \mathcal{V} \setminus \mathcal{V}_0$ the components of \tilde{b} are given by

$$\tilde{b}^v := \begin{cases} b^v + q^e & \text{for all } e \in \mathcal{E}_+ \cup \mathcal{E}_C, \\ b^v - q^e & \text{for all } e \in \mathcal{E}_- \cup \mathcal{E}_C, \\ b^v & \text{else.} \end{cases} \quad (3.25)$$

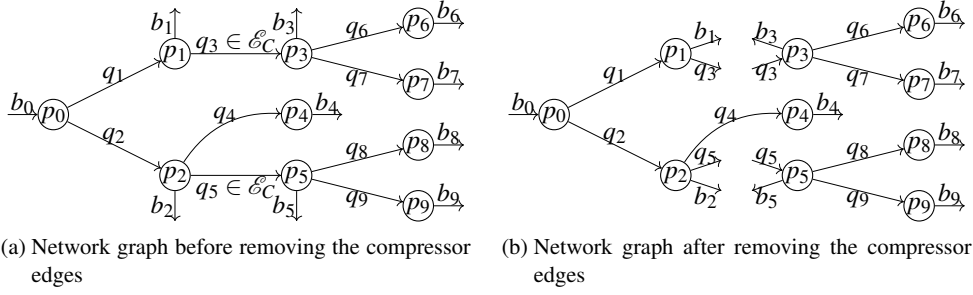


Figure 3.8.: Network graph with two compressor edges (left) and new network graph with three connected subgraphs after removing the compressor edges (right)¹

The resulting graph that corresponds to \tilde{b} is not connected anymore, it consists in $n_C + 1$ disjoint subgraphs which interact due to the compressor property (3.21). For this new structure, we need an additional notation. Let \mathcal{G}_i be the i -th subgraph of \mathcal{G} , which is connected, directed and tree-structured. Let n_i denote the number of outflow nodes and edges in the i -th subgraph. The vectors $\tilde{b}_i \in \mathbb{R}^{n_i}$, $q_i \in \mathbb{R}^{n_i}$ and $p_i^+ \in \mathbb{R}^{n_i+1}$ denote the load vector, the flow vector and the pressure vector of subgraph G_i . Then $\tilde{b}_{i,j}$, $q_{i,j}$ and $p_{i,j}^+$ is the j -th component of the load vector, the flow vector and the pressure vector of the subgraph \mathcal{G}_i . Further, A_i^+ resp. A_i (see (3.9)) denotes the incidence matrix of the i -th subgraph and Φ_i is the diagonal matrix of (3.23) for the i -th subgraph. We mention here that the original pressure and flow vectors $p^+ \in \mathbb{R}^{n+1}$ and $q \in \mathbb{R}^n$ consists of the values in p_i^+ and q_i ($i = 1, \dots, n_C + 1$), depending on the numbering of the graph.

Remark 3.10. *The new load vector $\tilde{b} \in \mathbb{R}^n$ is not non-negative anymore. Since the graph \mathcal{G} is separated in $n_C + 1$ subgraphs \mathcal{G}_i , the new load vector \tilde{b} contains n_C non-positive values corresponding to the inflow nodes of the resulting subgraphs (without the first subgraph \mathcal{G}_1 , whose inflow node is neither contained in b nor in \tilde{b}). The load vectors $\tilde{b}_i \in \mathbb{R}_{\geq 0}^{n_i}$ contain the outflows of the subgraph \mathcal{G}_i , thus the values are non-negative. For Figure 3.8 this means that the values of \tilde{b} corresponding to node v_3 and v_5 are non-positive.*

¹Figure already published in [Gugat and Schuster, 2018, page 5]

We now define a new feasible set \tilde{M} for the graph without the compressor edges:

$$\tilde{M} := \left\{ \tilde{b} \in \mathbb{R}_{\geq 0}^n \left| \begin{array}{l} \exists (p^+, q) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n_F} \text{ s.t.} \\ 1) \quad A_i q_i = \tilde{b}_i \quad \forall i = 1, \dots, n_C + 1 \\ 2) \quad (A_i^+)^{\top} (p_i^+)^2 = \Phi_i |q_i| q_i \quad \forall i = 1, \dots, n_C + 1 \\ 3) \quad \left(\frac{p^{h(e)}}{p^{f(e)}} \right)^2 = u_e \quad \forall e \in \mathcal{E}_C \\ 4) \quad p^+ \in [p^{+, \min}, p^{+, \max}] \end{array} \right. \right\}. \quad (3.26)$$

So for a given load vector $\tilde{b} \in \mathbb{R}_{\geq 0}^n$ we need to find a pressure vector $p^+ \in \mathbb{R}^{n+1}$ and a flow vector $q \in \mathbb{R}^n$, s.t. conservation of mass (3.26, 1)) holds for every subgraph, continuity in pressure (3.26, 2)) holds for every subgraph, the compressor property (3.26, 3)), which is the only interaction between the subgraphs, holds for every compressor edge and the pressure vector has to satisfy the pressure bounds (3.26, 4)).

Lemma 3.11. *The feasible sets defined in (3.22) and (3.26) are equivalent in the following sense:*

$$b \in M \quad \Leftrightarrow \quad \tilde{b} \in \tilde{M}.$$

Proof.

„ \Rightarrow “: Let a load vector $b \in M$ be given. Consider the load vector \tilde{b} defined in (3.25). Then 3) and 4) in (3.26) directly follow because these conditions equally occur in the definitions of M and \tilde{M} . Further 1) in (3.26) follows directly from (3.24). Last due to the structure of the incidence matrices A_i ($i = 1, \dots, n_C + 1$), every line in (3.26, 2)) is the Weymouth equation for one edge $e \in \mathcal{E}_F$ and since the flow for all flux edges $e \in \mathcal{E}_F$ does not change in (3.25), (3.26, 2)) is equal to (3.20) for all edges $e \in \mathcal{E}_F$. Thus it follows $\tilde{b} \in \tilde{M}$.

„ \Leftarrow “: Let a load vector $\tilde{b} \in \tilde{M}$ be given and consider the load vector b by (3.25). Again, the compressor property (3.21) and the pressure bounds equivalently occur in the definition of M and \tilde{M} , so they directly follow. Moreover (3.2) directly follows from (3.24). Since the lines in (3.26, 2)) state the Weymouth equation for an edge $e \in \mathcal{E}_F$ and since the flows in the flux edges $e \in \mathcal{E}_F$ do not change in (3.25), (3.20) holds. Last we only consider a vector of pressures at the nodes and since the change in pressure in the edges is given by (3.20) for all flow edges and by (3.21) for all compressor edges, (3.4) also holds and thus it follows $b \in M$. \square

Since every subgraph is tree-structured, the incidence matrices A_i are square with full rank and it follows

$$q_i = A_i^{-1} \tilde{b}_i,$$

so for $b \in M$ resp. $\tilde{b} \in \tilde{M}$ we can directly compute the flow vector $q \in \mathbb{R}^n$ resp. $q \in \mathbb{R}^{n_F}$. Let $g_i(\cdot)$ denote the pressure loss function defined in (3.11) for the subgraph G_i . Consequently $g_{i,j}(\cdot)$ denotes the j -th component of $g_i(\cdot)$.

As mentioned before, the aim is to get a characterization of the feasible set \tilde{M} similar to *Lemma 3.7* because then for a given load vector \tilde{b} we need not to find a flow vector and a pressure vector anymore, s.t. the conditions of (3.26) hold, but we can simply check whether a number of inequalities is satisfied. For a characterization of this form we need a few definitions:

Definition 3.12. *Let a connected, directed, tree-structured graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertex set $\mathcal{V} = \{v_0, \dots, v_n\}$ and a set of edges $\mathcal{E} = \{e_1, \dots, e_m\} \subseteq \mathcal{V} \times \mathcal{V}$ be given ($n, m \in \mathbb{N}$).*

(i) For $v \in \mathcal{V}$, we define

$$\Pi(v) := \{v\} \cup \bigcup_{\{w \in \mathcal{V} \mid \exists e \in \mathcal{E}: f(e)=w \wedge h(e)=v\}} \Pi(w)$$

as the (unique) directed path from the root to node v .

(ii) For $v, w \in \mathcal{V}$, we define

$$\Pi_w(v) := \{w\} \cup \Pi(v) \setminus \Pi(w)$$

as the (unique) directed path from node w to node v .

It is $\Pi_w(v) := \emptyset$ if there exists no directed path from w to v .

(iii) For $i, j \in \{1, \dots, n_C\}$, we define

$$\mathcal{C}_{i,j} := \left\{ k \in \{1, \dots, n_C\} \mid \begin{array}{l} f(e_k) \in \Pi_{v_{i,0}}(v_{j,0}), \\ h(e_k) \in \Pi_{v_{i,0}}(v_{j,0}), \\ e_k \in \mathcal{E}_C \end{array} \right\}$$

as the index set of the compressor edges on the path from node $v_{i,0}$ to node $v_{j,0}$.

(iv) For $i, j \in \{1, \dots, n_C + 1\}$, we define

$$\mathfrak{I}_{i,j} := \{k \in \{1, \dots, n_C + 1\} \setminus \{i, j\} \mid v_{k,0} \in \Pi_{v_{i,0}}(v_{j,0})\}$$

as the index set of all subgraphs between \mathcal{G}_i and \mathcal{G}_j .

(v) For $i, j \in \{1, \dots, n_C + 1\}$, we define

$$k_{i,j}^* := \max \{k \in \{1, \dots, n_C + 1\} \mid v_{k,0} \in \Pi(v_{i,0}) \wedge v_{k,0} \in \Pi(v_{j,0})\}$$

as the largest index of all subgraphs, the paths from the root to $v_{i,0}$ and from the root to $v_{j,0}$ pass through.

For the notation it is very important, that the nodes, the subgraphs, the flow edges and the compressor edges are numbered consistently. Otherwise we need to change the notation and this would make everything even harder to read. The most common ways for numbering the nodes, the subgraphs and the compressor edges of a tree-structured graph is to use *breadth-first search* or *depth-first search* (as we did in Section 3.1). Then the edge $e \in \mathcal{E}_F$ with $h(e) = v_i$ gets the number i ($i = 1, \dots, n$). An example graph for illustrating the notation is shown in Figure 3.9.

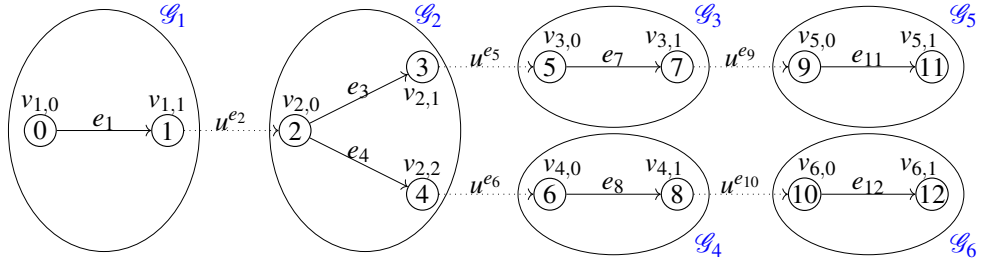


Figure 3.9.: Example graph numbered with breadth-first search for illustrating the introduced notation¹

Roughly speaking with the notation explained above and shown in Figure 3.9 we can follow the pressure loss along the paths. For example if we want to know the change in pressure between node v_0 resp. $v_{1,0}$ and v_4 resp. $v_{2,2}$ we first add the pressure loss function $g_{1,1}(\tilde{b}_1)$ to $p_{1,0}^2$, then we multiply it with the control u^{e_2} and last we add the pressure loss function $g_{2,2}(\tilde{b}_2)$. Then we now how the pressure changes along the path from node $v_{1,0}$ and node $v_{2,2}$. Before we state the characterization theorem, we define the sum

$$\Sigma_{i,j}(\tilde{b}) := \sum_{k=1}^{|\mathcal{J}_{i,j}|} \frac{1}{\prod_{\ell \in \mathcal{C}_{i,j}(k)} u^{e_\ell}} g_{\mathcal{J}_{i,j}(k), f(e_{\mathcal{C}_{i,j}(k+1)})}(\tilde{b}_{\mathcal{J}_{i,j}(k)}) + g_{i, f(e_{\mathcal{C}_{i,j}(1)})}(\tilde{b}_i), \quad (3.27)$$

and the product

$$\Pi_{i,j}(u) = \frac{1}{\prod_{\ell \in \mathcal{C}_{i,j}} u^{e_\ell}}, \quad (3.28)$$

where $u \in \mathbb{R}^{nc}$ is the vector of controls.

¹Similar figure already published in [Gugat and Schuster, 2018, page 7]

At first glance the notation used in the sum (3.27) and the product (3.28) seems to be very confusing but it will be clear as soon as we give a short example. We want to know the change in pressure between node v_0 resp. $v_{1,0}$ and node v_{12} resp. $v_{6,1}$ (see *Figure 3.9*). So we follow the change in pressure along the path from $v_{1,0}$ to $v_{6,1}$ by adding pressure loss functions and multiplying controls. It follows

$$\underbrace{\left[\underbrace{\left[\underbrace{p_{1,0}^2 - g_{1,1}(\tilde{b}_1)}_{=p_{2,0}^2} \right] u^{e_2} - g_{2,2}(\tilde{b}_2)}_{=p_{4,0}^2} \right] u^{e_6} - g_{4,1}(\tilde{b}_4)}_{=p_{6,0}^2} u^{e_{10}} + g_{6,1}(\tilde{b}_6) = p_{6,1}^2.$$

This can be equally written as

$$p_{1,0}^2 = \underbrace{\frac{1}{u^{e_2 u^{e_6} u^{e_{10}}}}}_{=\Pi_{1,6}(u)} (p_{6,1}^2 - g_{6,1}(\tilde{b}_6)) + \underbrace{\frac{1}{u^{e_2 u^{e_6}}} g_{4,1}(\tilde{b}_4) + \frac{1}{u^{e_2}} g_{2,2}(\tilde{b}_2) + g_{1,1}(\tilde{b}_1)}_{=\Sigma_{1,6}(\tilde{b})}.$$

The other way around, for $i = 1$ and $j = 6$, we have

$$\mathfrak{J}_{1,6} = \{2, 4\},$$

and

$$\mathcal{C}_{1,6} = \{2, 6, 10\}, \quad \mathcal{C}_{1,4} = \{2, 6\}, \quad \mathcal{C}_{1,2} = \{2\}.$$

It follows

$$\Pi_{1,6} = \frac{1}{\prod_{\ell \in \mathcal{C}_{1,6}} u^{e_\ell}} = \frac{1}{u^{e_2 u^{e_6} u^{e_{10}}}},$$

and

$$\begin{aligned} \Sigma_{1,6}(\tilde{b}) &= \sum_{k=1}^2 \frac{1}{\prod_{\ell \in \mathcal{C}_{1,\mathfrak{J}_{1,6}(k)}} u^{e_\ell}} g_{\mathfrak{J}_{1,6}(k), f(e_{\mathcal{C}_{1,6}(k+1)})}(\tilde{b}_{\mathfrak{J}_{1,6}(k)}) + g_{1, f(e_2)}(\tilde{b}_1) \\ &= \underbrace{\frac{1}{\prod_{\ell \in \{2\}} u^{e_\ell}} g_{\mathfrak{J}_{1,6}(1), f(e_6)}(\tilde{b}_{\mathfrak{J}_{1,6}(1)})}_{k=1} + \underbrace{\frac{1}{\prod_{\ell \in \{2,6\}} u^{e_\ell}} g_{\mathfrak{J}_{1,6}(2), f(e_{10})}(\tilde{b}_{\mathfrak{J}_{1,6}(2)})}_{k=2}} \\ &\quad + g_{1,1}(\tilde{b}_1) \\ &= \frac{1}{u^{e_2}} g_{2,2}(\tilde{b}_2) + \frac{1}{u^{e_2 u^{e_6}}} g_{4,1}(\tilde{b}_4) + g_{1,1}(\tilde{b}_1). \end{aligned}$$

The index $f(e_6)$ in $g_{2,f(e_6)}(\tilde{b}_2)$ states the second component of $g_2(\tilde{b}_2)$ because it is $f(e_6) = v_{2,2}$, i.e. the node with number 2 in the numbering of \mathcal{G}_2 . Analogously this holds for $g_{4,f(e_{10})}(\tilde{b}_4)$. Thus we can write the change in pressure between node $v_{1,0}$ and node $v_{6,1}$ as

$$p_{1,0}^2 = \Pi_{1,6}(u) (p_{6,1}^2 - g_{6,1}(\tilde{b}_6)) + \Sigma_{1,6}(\tilde{b}).$$

If we want to know the change in pressure between e.g. node $v_{5,1}$ and node $v_{6,1}$, we need to know the change in pressure between node $v_{2,0}$ and node $v_{5,1}$ and the change in pressure between node $v_{2,0}$ and node $v_{6,1}$. Then we need the index $k_{i,j}^*$ from *Definition 3.12*, it is $k_{5,6}^* = 2$. For better readability we write k^* instead of $k_{i,j}^*$ from now on.

The following characterization theorem is a main result of [Gugat and Schuster, 2018] with a slightly changed and better explained notation. It holds for a general tree-structured graph with a finite number of compressor edges.

Theorem 3.13. *Let pressure bounds $p^{+,min}, p^{+,max} \in \mathbb{R}^{n+1}$ and controls $u^e \in \mathbb{R}$ ($e \in \mathcal{E}_C$) be given.*

A vector $\tilde{b} \in \mathbb{R}^n$ is feasible, i.e. $\tilde{b} \in \tilde{M}$, if and only if the following inequalities hold:

For all $i = 1, \dots, n_C + 1$, the inequalities

$$(p_{i,0}^{min})^2 \leq \min_{k=1, \dots, n_i} \left[(p_{i,k}^{max})^2 + g_{i,k}(\tilde{b}_i) \right], \quad (3.29)$$

$$(p_{i,0}^{max})^2 \geq \max_{k=1, \dots, n_i} \left[(p_{i,k}^{min})^2 + g_{i,k}(\tilde{b}_i) \right], \quad (3.30)$$

$$\max_{k=1, \dots, n_i} \left[(p_{i,k}^{min})^2 + g_{i,k}(\tilde{b}_i) \right] \leq \min_{k=1, \dots, n_i} \left[(p_{i,k}^{max})^2 + g_{i,k}(\tilde{b}_i) \right], \quad (3.31)$$

hold (feasibility inside the subgraphs). For all $i, j = 1, \dots, n_C + 1$, $i < j$, the inequalities

$$\Pi_{k^*,i}(u) (p_{i,0}^{min})^2 + \Sigma_{k^*,i}(\tilde{b}) \leq \Pi_{k^*,j}(u) (p_{j,0}^{max})^2 + \Sigma_{k^*,j}(\tilde{b}), \quad (3.32)$$

$$\Pi_{k^*,i}(u) (p_{i,0}^{max})^2 + \Sigma_{k^*,i}(\tilde{b}) \geq \Pi_{k^*,j}(u) (p_{j,0}^{min})^2 + \Sigma_{k^*,j}(\tilde{b}), \quad (3.33)$$

hold (feasibility between the roots of subgraph \mathcal{G}_i and \mathcal{G}_j).

For all $i, j = 1, \dots, n_C + 1, i < j$, the inequalities

$$\begin{aligned} \Pi_{k^*,i}(u) (p_{i,0}^{min})^2 + \Sigma_{k^*,i}(\tilde{\mathbf{b}}) &\leq \\ &\leq \Pi_{k^*,j}(u) \min_{k=1,\dots,n_j} \left[(p_{j,k}^{max})^2 + g_{j,k}(\tilde{\mathbf{b}}_j) \right] + \Sigma_{k^*,j}(\tilde{\mathbf{b}}), \end{aligned} \quad (3.34)$$

$$\begin{aligned} \Pi_{k^*,i}(u) (p_{i,0}^{max})^2 + \Sigma_{k^*,i}(\tilde{\mathbf{b}}) &\geq \\ &\geq \Pi_{k^*,j}(u) \max_{k=1,\dots,n_j} \left[(p_{j,k}^{min})^2 + g_{j,k}(\tilde{\mathbf{b}}_j) \right] + \Sigma_{k^*,j}(\tilde{\mathbf{b}}), \end{aligned} \quad (3.35)$$

hold (feasibility between the root of subgraph \mathcal{G}_i and the outflows of subgraph \mathcal{G}_j). For all $i, j = 1, \dots, n_C + 1, i < j$, the inequalities

$$\begin{aligned} \Pi_{k^*,i}(u) \max_{k=1,\dots,n_i} \left[(p_{i,k}^{min})^2 + g_{i,k}(\tilde{\mathbf{b}}_i) \right] + \Sigma_{k^*,i}(\tilde{\mathbf{b}}) &\leq \\ &\leq \Pi_{k^*,j}(u) (p_{j,0}^{max})^2 + \Sigma_{k^*,j}(\tilde{\mathbf{b}}), \end{aligned} \quad (3.36)$$

$$\begin{aligned} \Pi_{k^*,i}(u) \min_{k=1,\dots,n_i} \left[(p_{i,k}^{max})^2 + g_{i,k}(\tilde{\mathbf{b}}_i) \right] + \Sigma_{k^*,i}(\tilde{\mathbf{b}}) &\geq \\ &\geq \Pi_{k^*,j}(u) (p_{j,0}^{min})^2 + \Sigma_{k^*,j}(\tilde{\mathbf{b}}), \end{aligned} \quad (3.37)$$

hold (feasibility between the outflows of subgraph \mathcal{G}_i and the root of subgraph \mathcal{G}_j). For all $i, j = 1, \dots, n_C + 1, i < j$, the inequalities

$$\begin{aligned} \Pi_{k^*,i}(u) \max_{k=1,\dots,n_i} \left[(p_{i,k}^{min})^2 + g_{i,k}(\tilde{\mathbf{b}}_i) \right] + \Sigma_{k^*,i}(\tilde{\mathbf{b}}) &\leq \\ &\leq \Pi_{k^*,j}(u) \min_{k=1,\dots,n_j} \left[(p_{j,k}^{max})^2 + g_{j,k}(\tilde{\mathbf{b}}_j) \right] + \Sigma_{k^*,j}(\tilde{\mathbf{b}}), \end{aligned} \quad (3.38)$$

$$\begin{aligned} \Pi_{k^*,i}(u) \min_{k=1,\dots,n_i} \left[(p_{i,k}^{max})^2 + g_{i,k}(\tilde{\mathbf{b}}_i) \right] + \Sigma_{k^*,i}(\tilde{\mathbf{b}}) &\geq \\ &\geq \Pi_{k^*,j}(u) \max_{k=1,\dots,n_j} \left[(p_{j,k}^{min})^2 + g_{j,k}(\tilde{\mathbf{b}}_j) \right] + \Sigma_{k^*,j}(\tilde{\mathbf{b}}), \end{aligned} \quad (3.39)$$

hold (feasibility between the outflows of subgraph \mathcal{G}_i and \mathcal{G}_j).

The result of *Theorem 3.13* is the following: For a given load vector $\tilde{\mathbf{b}} \in \mathbb{R}^n$ we need not to find a $(p^+, q) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n_F}$ anymore to show that $\tilde{\mathbf{b}} \in \tilde{M}$, it is sufficient to check the inequalities (3.29) - (3.39). These inequalities compare the pressure bounds among each other to guarantee that a pressure vector exists, but we need not to compute it. The inequalities (3.29) - (3.31) compare the pressure bounds inside the subgraphs. The other inequalities compare the pressure bounds between the subgraphs.

Remark 3.14. For the inequalities (3.32) - (3.39) we assumed

$$i, j = 1, \dots, n_C + 1 \quad \text{with} \quad i < j.$$

For $i = j$ the inequalities hold according to the requirement $p^{+,min} \leq p^{+,max}$. For $i > j$ the inequalities also hold since we can swap the indices, e.g. for $i > j$ inequality (3.33) becomes (3.32) and vice versa.

Proof of Theorem 3.13.

„ \Rightarrow “: Consider a feasible load vector $\tilde{b} \in \tilde{M}$. The equation for the pressure loss (3.23) for the subgraph \mathcal{G}_i ($i = 1, \dots, n_C + 1$) is

$$(A_i^+)^{\top} (p_i^+)^2 = -\Phi_i |q_i| q_i.$$

We split the incidence matrix A_i^+ in its first row $A_{i,0}^+$ and the remaining part A_i . It follows

$$(A_{i,0}^+)^{\top} p_{i,0}^2 + A_i^{\top} p_i^2 = -\Phi_i |q_i| q_i. \quad (3.40)$$

Since the columns in A_i^+ sum up to zero, it holds

$$-\mathbb{1}_{n_i}^{\top} A_i = A_{i,0}^+, \quad (3.41)$$

where $\mathbb{1}_{n_i}$ is the all-ones vector in \mathbb{R}^{n_i} . Thus we have

$$-A_i^{\top} \mathbb{1}_{n_i} p_{i,0}^2 + A_i^{\top} p_i^2 = -\Phi_i |q_i| q_i. \quad (3.42)$$

Next we insert (3.10) into (3.42) and multiply the equation by $(A_i^{\top})^{-1}$ from the left. Then we have

$$\mathbb{1}_{n_i} p_{i,0}^2 = p_i^2 + (A_i^{\top})^{-1} \Phi_i |A_i^{-1} \tilde{b}_i| A_i^{-1} \tilde{b}_i,$$

which is equal to

$$\mathbb{1}_{n_i} p_{i,0}^2 = p_i^2 + g_i(\tilde{b}_i), \quad (3.43)$$

by the definition of the pressure loss function g for the subgraph \mathcal{G}_i (see (3.11)). Note that p_i and $g_i(\cdot)$ are pressure vector and pressure loss function for the subgraph \mathcal{G}_i , not the i -th component of the vector p resp. the function $g(\cdot)$. The inequalities (3.29) - (3.31) follow directly from (3.43) by estimating the pressures by its bounds $p_i^{+,min}, p_i^{+,max} \in \mathbb{R}^{n_i+1}$.

We fix $i, j \in \{1, \dots, n_C + 1\}$ with $i < j$. Remember that $k_{i,j}^*$ is the largest index of all subgraphs which both paths pass through, the path from the root

to $v_{i,0}$ and the path from the root to $v_{j,0}$. We mentioned before that we write k^* instead of $k_{i,j}^*$. Then from (3.43) it follows

$$p_{k^*,0}^2 = p_{k^*,k}^2 + g_{k^*,k}(\tilde{b}_{k^*}) \quad \text{for all } k = 1, \dots, n_{k^*}. \quad (3.44)$$

The indices of the compressor edges on the paths from $v_{k^*,0}$ to $v_{i,0}$ and $v_{j,0}$ are given in $\mathcal{C}_{k^*,i}$ and $\mathcal{C}_{k^*,j}$. The indices of the subgraphs between \mathcal{G}_{k^*} and \mathcal{G}_i resp. \mathcal{G}_j are given in $\mathfrak{J}_{k^*,i}$ resp. $\mathfrak{J}_{k^*,j}$. We follow the path from $v_{k^*,0}$ to $v_{i,0}$. For better readability we only write \mathcal{C} instead of $\mathcal{C}_{k^*,i}$ and \mathfrak{J} instead of $\mathfrak{J}_{k^*,i}$. We use the compressor property for the first compressor edge on the path from $v_{k^*,0}$ to $v_{i,0}$, i.e. $e_{\mathcal{C}(1)}$, which is

$$\frac{p_{h(e_{\mathcal{C}(1)})}^2}{p_{f(e_{\mathcal{C}(1)})}^2} = u^{e_{\mathcal{C}(1)}},$$

and insert it into (3.44) for $p_{h(e_{\mathcal{C}(1)})} = p_{\mathfrak{J}(1),0}$ and $p_{f(e_{\mathcal{C}(1)})} = p_{k^*,k}$. It follows

$$p_{k^*,0}^2 = \frac{1}{u^{e_{\mathcal{C}(1)}}} p_{\mathfrak{J}(1),0}^2 + g_{k^*,f(e_{\mathcal{C}(1)})}(\tilde{b}_{k^*}). \quad (3.45)$$

If $n_{\mathcal{C}} = 1$, we are done. Otherwise we follow the path from $v_{k^*,0}$ to $v_{i,0}$ to the next compressor edge (which is $\mathcal{C}(2)$) and insert (3.44) for $v_{k^*,k} = f(e_{\mathcal{C}(2)})$ into (3.45), thus we have

$$p_{k^*,0}^2 = \frac{1}{u^{e_{\mathcal{C}(1)}}} \left[p_{\mathfrak{J}(1),f(e_{\mathcal{C}(2)})}^2 + g_{\mathfrak{J}(1),f(e_{\mathcal{C}(2)})}(\tilde{b}_{\mathfrak{J}(1)}) \right] + g_{k^*,f(e_{\mathcal{C}(1)})}(\tilde{b}_{k^*}). \quad (3.46)$$

Roughly speaking we arrived at the end of the first subgraph between \mathcal{G}_{k^*} and \mathcal{G}_i . We need to continue following the path from $v_{k^*,0}$ to $v_{i,0}$ by inserting the compressor property for $\mathcal{C}(2)$ in (3.46). It follows

$$p_{k^*,0}^2 = \frac{1}{u^{e_{\mathcal{C}(1)}}} \left[\frac{1}{u^{e_{\mathcal{C}(2)}}} p_{\mathfrak{J}(2),0}^2 + g_{\mathfrak{J}(1),f(e_{\mathcal{C}(2)})}(\tilde{b}_{\mathfrak{J}(1)}) \right] + g_{k^*,f(e_{\mathcal{C}(1)})}(\tilde{b}_{k^*}). \quad (3.47)$$

We repeat the procedure of (3.46) and (3.47) until we reach \mathcal{G}_i . To avoid introducing new variables we write $\mathcal{C}(\text{end})$ resp. $\mathfrak{J}(\text{end})$ for the last element of \mathcal{C} resp. \mathfrak{J} . Then we have

$$p_{k^*,0}^2 = \frac{1}{u^{e_{\mathcal{C}(1)}}} \left[\frac{1}{u^{e_{\mathcal{C}(2)}}} \left[\dots \left[\frac{1}{u^{e_{\mathcal{C}(\text{end})}}} p_{i,0}^2 + g_{\mathfrak{J}(\text{end}),f(e_{\mathcal{C}(\text{end})})}(\tilde{b}_{\mathfrak{J}(\text{end})}) \right] \dots \right] + g_{\mathfrak{J}(1),f(e_{\mathcal{C}(2)})}(\tilde{b}_{\mathfrak{J}(1)}) \right] + g_{k^*,f(e_{\mathcal{C}(1)})}(\tilde{b}_{k^*}). \quad (3.48)$$

We expand (3.48) and since we have $|\mathcal{J}| = |\mathcal{C}| - 1$, it follows

$$\begin{aligned} p_{k^*,0}^2 &= \frac{1}{u^{e_{\mathcal{C}}(1)} u^{e_{\mathcal{C}}(2)} \dots u^{e_{\mathcal{C}}(\text{end}-1)} u^{e_{\mathcal{C}}(\text{end})}} p_{i,0}^2 \\ &+ \frac{1}{u^{e_{\mathcal{C}}(1)} u^{e_{\mathcal{C}}(2)} \dots u^{e_{\mathcal{C}}(\text{end}-1)}} g_{\mathcal{J}(\text{end}), f(e_{\mathcal{C}}(\text{end}))}(\tilde{\mathbf{b}}_{\mathcal{J}(\text{end})}) \\ &+ \dots \\ &+ \frac{1}{u^{e_{\mathcal{C}}(1)}} g_{\mathcal{J}(1), f(e_{\mathcal{C}}(2))}(\tilde{\mathbf{b}}_{\mathcal{J}(1)}) \\ &+ g_{k^*, f(e_{\mathcal{C}}(1))}(\tilde{\mathbf{b}}_{k^*}), \end{aligned}$$

which is equal to

$$p_{k^*,0}^2 = \Pi_{k^*,i}(u) p_{i,0}^2 + \Sigma_{k^*,i}(\tilde{\mathbf{b}}), \quad (3.49)$$

by the definitions (3.27) and (3.28). Analogously we get

$$p_{k^*,0}^2 = \Pi_{k^*,j}(u) p_{j,0}^2 + \Sigma_{k^*,j}(\tilde{\mathbf{b}}). \quad (3.50)$$

We equalize (3.49) with (3.50) and we have

$$\Pi_{k^*,i}(u) p_{i,0}^2 + \Sigma_{k^*,i}(\tilde{\mathbf{b}}) = \Pi_{k^*,j}(u) p_{j,0}^2 + \Sigma_{k^*,j}(\tilde{\mathbf{b}}). \quad (3.51)$$

We estimate $p_{i,0}$ and $p_{j,0}$ by its pressure bounds $p_i^{+, \min}, p_i^{+, \max} \in \mathbb{R}^{n_i+1}$ and $p_j^{+, \min}, p_j^{+, \max} \in \mathbb{R}^{n_j+1}$ and the inequalities (3.32), (3.33) directly follow.

From (3.43) it follows

$$p_{j,0}^2 = p_{j,k}^2 + g_{j,k}(\tilde{\mathbf{b}}_j) \quad k = 1, \dots, n_j. \quad (3.52)$$

We insert this in (3.51) and by estimating the pressures by its bounds the inequalities (3.34), (3.35) follow. From (3.43) it also follows

$$p_{i,0}^2 = p_{i,k}^2 + g_{i,k}(\tilde{\mathbf{b}}_i) \quad k = 1, \dots, n_i. \quad (3.53)$$

We insert this in (3.51) and again by estimating the pressures by its bounds the inequalities (3.36), (3.37) follow.

Last we insert both, (3.52) and (3.53), in (3.51) and by estimating the pressures by its bounds the inequalities (3.38) and (3.39) hold.

In a nutshell, a given feasible load vector $\tilde{\mathbf{b}} \in \tilde{M}$ satisfies the inequalities (3.29) - (3.39) and the first part of the proof is complete.

„ \Leftarrow “: For this part of the proof, we consider a vector \tilde{b} , that satisfies the inequalities (3.29) - (3.39). We can easily find a flow vector $q \in \mathbb{R}^{n_F}$ via (3.10) for every subgraph, but the clue is to find a pressure vector $p^+ \in \mathbb{R}^{n+1}$, s.t. \tilde{b} is feasible, i.e. $\tilde{b} \in \tilde{M}$. The idea is to find a reasonable pressure at the node v_0 s.t. we can derive a pressure vector $p^+ \in \mathbb{R}^{n+1}$. For $i = 2, \dots, n_C + 1$ we define the following sets:

$$\begin{aligned}
P_{1,1} &:= \left[(p_{1,0}^{\min})^2, (p_{1,0}^{\max})^2 \right], \\
P_{2,1} &:= \left[\max_{k=1, \dots, n_i} \left[(p_{1,k}^{\min})^2 + g_{1,k}(\tilde{b}_1) \right], \min_{k=1, \dots, n_i} \left[(p_{1,k}^{\max})^2 + g_{1,k}(\tilde{b}_1) \right] \right], \\
P_{1,i} &:= \left[\Pi_{1,i}(u)(p_{i,0}^{\min})^2 + \Sigma_{1,i}(\tilde{b}), \Pi_{1,i}(u)(p_{i,0}^{\max})^2 + \Sigma_{1,i}(\tilde{b}) \right], \\
P_{2,i} &:= \left[\Pi_{1,i}(u) \max_{k=1, \dots, n_i} \left[(p_{i,k}^{\min})^2 + g_{i,k}(\tilde{b}_i) \right], \Pi_{1,i}(u) \min_{k=1, \dots, n_i} \left[(p_{i,k}^{\max})^2 + g_{i,k}(\tilde{b}_i) \right] \right].
\end{aligned}$$

Further we define

$$P_0 := \bigcap_{k=1}^{n_C+1} P_{1,k} \cap \bigcap_{k=1}^{n_C+1} P_{2,k}.$$

We first show, that P_0 is nonempty. Mention that for $i = 1, \dots, n_C + 1$ it is

$$\begin{cases} P_{1,i} \neq \emptyset & \text{according to requirement,} \\ P_{2,i} \neq \emptyset & \text{due to (3.31).} \end{cases}$$

It is sufficient to show, that $P_{i,j} \cap P_{k,\ell}$ ($i, k \in \{1, 2\}$, $j, \ell \in \{1, \dots, n_C + 1\}$) is nonempty, then it follows that P_0 is nonempty, too. This can easily be shown by the inequalities (3.29) - (3.39). It is $P_{1,i} \cap P_{2,i} \neq \emptyset$ due to (3.29) and (3.30). Further for $i, j \in \{1, \dots, n_C + 1\}$, $i < j$ we have

$$\begin{cases} P_{1,i} \cap P_{1,j} \neq \emptyset & \text{due to (3.32) and (3.33),} \\ P_{1,i} \cap P_{2,j} \neq \emptyset & \text{due to (3.34) and (3.35),} \\ P_{2,i} \cap P_{1,j} \neq \emptyset & \text{due to (3.36) and (3.37),} \\ P_{2,i} \cap P_{2,j} \neq \emptyset & \text{due to (3.38) and (3.39).} \end{cases}$$

For $i, j \in \{1, \dots, n_C + 1\}$, $i > j$ the intersections are also nonempty (see *Remark 3.14*). Thus it follows

$$P_0 \neq \emptyset.$$

We fix a value $p_0 \in P_0$ and define the values

$$\begin{aligned} p_{1,0}^2 &:= p_0, \\ p_{i,0}^2 &:= \frac{1}{\Pi_{1,i}(u)} (p_{1,0}^2 - \Sigma_{1,i}(\tilde{\mathbf{b}})) \quad i = 2, \dots, n_C + 1, \\ p_{i,k}^2 &:= p_{i,0}^2 - g_{i,k}(\tilde{\mathbf{b}}_i) \quad i = 1, \dots, n_C + 1, \quad k = 1, \dots, n_i, \\ q_i &:= A_i^{-1} \tilde{\mathbf{b}}_i \quad i = 1, \dots, n_C + 1. \end{aligned}$$

We need to show, that the pressure vector $p^+ \in \mathbb{R}^{n+1}$ (which consists of the vectors $p_i^+ \in \mathbb{R}^{n_i+1}$) and the flow vector $q \in \mathbb{R}^{n_F}$ (which consists of the vectors $q_i \in \mathbb{R}^{n_i}$) satisfy the conditions in (3.26). Since $p_{1,0} \in P_{1,1}$, it follows

$$p_{1,0} \in \left[p_{1,0}^{\min}, p_{1,0}^{\max} \right].$$

Further since $p_{i,0}^2 \in P_{1,i}$ for all $i = 2, \dots, n_C + 1$, it holds

$$\Pi_{1,i}(u) p_{i,0}^2 + \Sigma_{1,i}(\tilde{\mathbf{b}}) \leq \Pi_{1,i}(u) (p_{i,0}^{\max})^2 + \Sigma_{1,i}(\tilde{\mathbf{b}}),$$

and

$$\Pi_{1,i}(u) p_{i,0}^2 + \Sigma_{1,i}(\tilde{\mathbf{b}}) \geq \Pi_{1,i}(u) (p_{i,0}^{\min})^2 + \Sigma_{1,i}(\tilde{\mathbf{b}}).$$

Now we insert $p_{i,0}^2 = p_{i,k}^2 + g_{i,k}(\tilde{\mathbf{b}}_i)$ in these estimations and since $p_0 \in P_{2,i}$, for all $i = 1, \dots, n_C + 1$ it follows

$$\Pi_{1,i}(u) (p_{i,k}^2 + g_{i,k}(\tilde{\mathbf{b}}_i)) + \Sigma_{1,i} \leq \Pi_{1,i}(u) \min_{k=1, \dots, n_i} \left[(p_{i,k}^{\max})^2 + g_{i,k}(\tilde{\mathbf{b}}_i) \right] + \Sigma_{1,i}(\tilde{\mathbf{b}}),$$

and

$$\Pi_{1,i}(u) (p_{i,k}^2 + g_{i,k}(\tilde{\mathbf{b}}_i)) + \Sigma_{1,i} \geq \Pi_{1,i}(u) \min_{k=1, \dots, n_i} \left[(p_{i,k}^{\min})^2 + g_{i,k}(\tilde{\mathbf{b}}_i) \right] + \Sigma_{1,i}(\tilde{\mathbf{b}}).$$

Thus we have $p_{i,k} \in [p_{i,k}^{\min}, p_{i,k}^{\max}]$ for all $k = 1, \dots, n_i$, which is equivalent to

$$p^+ \in \left[p^{+, \min}, p^{+, \max} \right],$$

so the pressure vector $p^+ \in \mathbb{R}^{n+1}$ satisfies the pressure bounds (3.26, 4)).

The equation of mass conservation for every subgraph (3.26, 1)) directly follows from the definition of the $q_i \in \mathbb{R}^{n_i}$ for $i = 1, \dots, n_C + 1$.

From the definition of $p_{i,k}^2$ it follows

$$p_i^2 = \mathbb{1}_{n_i} p_{i,0}^2 - g_i(\tilde{b}_i).$$

We multiply this equation by A_i^\top from the left and use (3.41). Then it follows

$$(A_i^+)^{\top} (p_i^+)^2 = -\Phi_i |q_i| q_i,$$

which is (3.26, 2)).

The last equation we have to show is the compressor property. Consider a compressor edge $e_k \in \mathcal{E}_C$, which connects the subgraphs \mathcal{G}_i with \mathcal{G}_j ($i < j$). By the definition of $p_{j,0}^2$ we have

$$p_{j,0}^2 = \frac{1}{\prod_{1,j}(u)} (p_{1,0}^2 - \Sigma_{1,j}(\tilde{b})).$$

We split the sum $\Sigma_{1,j}(\tilde{b})$ (defined in (3.27)) in the part for $k = |\mathcal{J}_{1,j}|$ and the remaining part (for $k < |\mathcal{J}_{1,j}|$) which is equal to $\Sigma_{1,i}(\tilde{b})$. We have

$$\Sigma_{1,j}(\tilde{b}) = \Sigma_{1,i}(\tilde{b}) + \frac{1}{\prod_{\ell \in \mathcal{C}_{1, \mathcal{J}_{1,j}(\text{end})}} u^{e_\ell}} g_{\mathcal{J}_{1,j}(\text{end}), f(e_{\mathcal{C}_{1, \mathcal{J}_{1,j}(\text{end})}})}(\tilde{b}_{\mathcal{J}_{1,j}(\text{end})}).$$

By our choice of i and j , it is $\mathcal{J}_{1,j}(\text{end}) = i$ and it is $e_{\mathcal{C}_{1, \mathcal{J}_{1,j}(\text{end})}} = e_k$. So we have

$$\Sigma_{1,j}(\tilde{b}) = \Sigma_{1,i}(\tilde{b}) + \frac{1}{\prod_{\ell \in \mathcal{C}_{1,i}} u^{e_\ell}} g_{i, f(e_k)}(\tilde{b}_i),$$

and thus

$$p_{j,0}^2 = \frac{1}{\prod_{1,j}(u)} \left(p_{1,0}^2 - \Sigma_{1,i}(\tilde{b}) - \frac{1}{\prod_{\ell \in \mathcal{C}_{1,i}} u^{e_\ell}} g_{i, f(e_k)}(\tilde{b}_i) \right).$$

We use (3.28) and it follows

$$p_{j,0}^2 = \frac{1}{\prod_{1,j}(u)} (p_{1,0}^2 - \Sigma_{1,i}(\tilde{b}) - \prod_{1,i}(u) g_{i, f(e_k)}(\tilde{b}_i)).$$

Further we can split the product $\prod_{1,j}(u)$ in the last control on the path from $v_{1,0}$ to $v_{j,0}$ and the remaining controls. We have

$$\prod_{1,j}(u) = \prod_{1,i}(u) \frac{1}{u^{e_k}}.$$

Then it follows

$$p_{j,0}^2 = u^{e_k} \frac{1}{\prod_{1,i}(u)} (p_{1,0}^2 - \Sigma_{1,i}(\tilde{b}) - \prod_{1,i}(u) g_{i,f(e_k)}(\tilde{b}_i)),$$

which is equal to

$$p_{j,0}^2 = u^{e_k} \left(\frac{1}{\prod_{1,i}(u)} (p_{1,0}^2 - \Sigma_{1,i}(\tilde{b})) - g_{i,f(e_k)}(\tilde{b}_i) \right).$$

With the definition of $p_{i,0}^2$ it follows

$$p_{j,0}^2 = u^{e_k} (p_{i,k}^2 - g_{i,f(e_k)}(\tilde{b}_i)),$$

and with the definition of $p_{i,k}^2$ it follows

$$p_{j,0}^2 = u^{e_k} p_{i,f(e_k)}^2,$$

which is the compressor property (3.26, 3)) for the compressor edge $e_k \in \mathcal{E}_C$.

So we have shown that all conditions in (3.26) are satisfied and \tilde{b} is feasible, i.e. $\tilde{b} \in \tilde{M}$. Hence this part of the proof is complete as well as the whole proof. □

Remark 3.15. *The proof of Theorem 3.13 seems to be very difficult but it becomes quite clear as soon as one understand the notation introduced above Theorem 3.13. The clou consists in following the change of pressure along the edges in the graph.*

The characterization stated in *Theorem 3.13* provides the announced advantage: For a given load vector $\tilde{b} \in \mathbb{R}^n$ we need not to find a pressure vector $p^+ \in \mathbb{R}^{n+1}$ and a flow vector $q \in \mathbb{R}^{n_F}$, we only need to check the inequalities (3.29) - (3.39).

Now we introduce the uncertainty and use the SRD to compute the probability for a random vector to be feasible. Consider the Gaussian distributed n -dimensional random variable

$$\xi_b \sim \mathcal{N}(\mu, \Sigma),$$

with mean value $\mu \in \mathbb{R}_{\geq 0}^n$ and positive definite covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$ on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We identify the load vector $b \in \mathbb{R}^n$

with the image $\xi_b(\omega)$ for $\omega \in \Omega$ on this probability space. Let $L \in \mathbb{R}^{n \times n}$ be the Cholesky decomposition of Σ , i.e.

$$\Sigma = LL^\top,$$

which exists due to the positive definiteness of Σ . Let $v \in \mathbb{S}^{n-1}$ be a point on the unit sphere. We define $\omega_v := Lv$ and set

$$b_v(r) = r\omega_v + \mu.$$

As in *Section 3.2* we need to find the range of r , s.t. $b_v(r)$ is in M . Due to the additional non-negativity restriction we define the regular range

$$R_{v,\text{reg}} := \{r \geq 0 | b_v(r) \geq 0\}.$$

Now we need to find the feasible parts of the regular range, which is slightly different to *Section 3.2*. By (3.10) we can compute the flow vector q (which depends on r now) and we define the new load vector $\tilde{b}(r)$ by using (3.25). Due to *Lemma 3.11* it holds

$$M_v = \{r \in R_{v,\text{reg}} | b(r) \in M\} = \{r \in R_{v,\text{reg}} | \tilde{b}(r) \in \tilde{M}\},$$

and due to *Theorem 3.13* it follows

$$M_v = \{r \in R_{v,\text{reg}} | \tilde{b}(r) \text{ satisfies (3.29) - (3.39)}\}.$$

Since the $g_{i,k}(\tilde{b}_i(r))$ are quadratic functions in r , the inequalities (3.29) - (3.39) are also quadratic in r and thus we can write the one dimensional feasible sets M_v as a union of $s \in \mathbb{N}$ disjoint intervals (cf. (3.13)). Since we can apply *Remark 3.8* analogously to the inequalities of *Theorem 3.13*, it follows that M_v is a union of finite, closed intervals.

Then for $N_{\text{SRD}} \in \mathbb{N}$ let

$$\mathcal{S} := \{v_1, \dots, v_{N_{\text{SRD}}}\} \subseteq \mathbb{S}^{n-1},$$

be a uniformly distributed sampling on the sphere \mathbb{S}^{n-1} . Then we can compute the probability for a random load vector to be feasible

$$\mathbb{P}_{\text{SRD}}(b \in M),$$

by using (3.14) and for $N_{\text{SRD}} \rightarrow \infty$, this probability converges to the exact probability.

In this section we cannot apply the KDE as we did in *Section 3.3* due to the complexity of the model. We would need to have at least one pressure by assumption to estimate the probability density of the pressures.

An application of this model to a real gas network can be found in [Gugat and Schuster, 2018, Section 5].

3.6. Optimization Problems with Probabilistic Constraints

In this section we consider two probabilistic constrained optimization problems that fit to the setting of this chapter and we analyze the existence of optimal solutions. We use the KDE approach that was introduced in *Section 1.1* and applied in *Section 3.3* to approximate the probabilistic constraints and we also analyze the existence of an optimal solution for the corresponding approximated probabilistic constrained optimization problems. Further we show that if the sampling size is sufficiently large, the solutions of the approximated problems are almost surely close to the solutions of the exact problems. In the end of this section we compute the derivatives of the approximated probabilistic constraints and we derive necessary optimality conditions for the approximated probabilistic constrained optimization problems. We mention R. Römisch and R. Schultz at this point, who published a row of papers about the field of stochastic programming in the 90th (see e.g. [Römisch and Schultz, 1991a], [Römisch and Schultz, 1991b] and [Römisch and Schultz, 1993]).

Let convex functions

$$f_1 : \mathbb{R}^n \rightarrow \mathbb{R}, \quad p^{\max} \mapsto f_1(p^{\max}),$$

and

$$f_2 : \mathbb{R} \rightarrow \mathbb{R}, \quad p_0 \mapsto f_2(p_0), \tag{3.54}$$

be given. For a probability level $\alpha \in (0, 1)$ consider the optimization problems

$$\begin{aligned} \min_{p^{\max} \in \mathbb{R}^n} \quad & f_1(p^{\max}) \\ \text{s.t.} \quad & \mathbb{P}(b \in M(p^{\max})) \geq \alpha, \\ & p^{\max} \geq p^{\min}, \end{aligned} \tag{3.55}$$

and

$$\begin{aligned} \min_{p_0 \in \mathbb{R}} \quad & f_2(p_0) \\ \text{s.t.} \quad & \mathbb{P}(b \in M(p_0)) \geq \alpha, \\ & p_0 \geq 0, \end{aligned} \tag{3.56}$$

where M is defined in (3.5). Usually α is chosen large, i.e. close to 1.

We first analyze the optimization problem (3.55). From (3.8) it follows that the probabilistic constraint in (3.55) is equal to

$$\int_{p_{\min}^{\max}} \varrho_p(z) dz \geq \alpha,$$

where ϱ_p is the probability density function of the pressure and p_{\min}^{\max} is defined in (3.7). The constraint $p^{\max} \geq p^{\min}$ comes from the application and it has to be understood component-by-component. In one dimension (i.e. for a single pipe) this constraint is redundant since a backward integration of a positive function is always negative and thus smaller than α . However in n dimensions this constraint can never be active. If we have $p_i^{\min} = p_i^{\max}$ for at least one index $i \in \{1, \dots, n\}$, then the probability in the first constraint is always 0 and the constraint cannot be satisfied. Thus the second constraint is always inactive.

For the existence of a solution we define the set of admissible pressures as

$$\mathcal{R} := \left\{ p^{\max} \in \mathbb{R}^n \mid p^{\max} \geq p^{\min} \text{ and } \int_{p_{\min}^{\max}} \varrho_p(z) dz \geq \alpha \right\}. \tag{3.57}$$

We need to assume that $p^{\min} \in \mathbb{R}^n$ is sufficiently small, s.t. the probabilistic constraint in (3.55) can even be satisfied. Since

$$\int_{\mathbb{R}^n} \varrho_p(z) dz = 1,$$

it could happen that the constraint cannot even reach α (an example is shown in *Figure 3.10*).

The next lemma states the existence of a solution of (3.55) under certain assumptions.

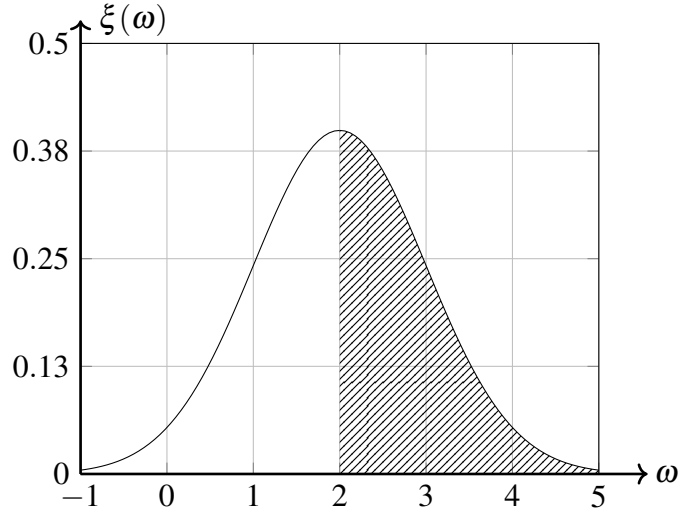


Figure 3.10.: Gaussian distribution with $\mu = 2$; It follows $|\int_2^x \xi(\omega) dt| \leq 0.5$ for all $x \in \mathbb{R}$

Lemma 3.16. *Let a probability level $\alpha \in (0, 1)$, an inlet pressure $p_0 \in \mathbb{R}_{\geq 0}$ and a lower pressure bound $p^{\min} \in \mathbb{R}_{\geq 0}^n$ be given. Assume that there exists $x \geq p^{\min}$ (component-by-component), s.t.*

$$\int_{p^{\min}}^x \varrho_p(z) dz > \alpha. \quad (3.58)$$

Let f_1 be strictly monotone increasing in the sense that for all positive directions $d \in \mathbb{R}_{\geq 0}^n$ (with $d > 0$ in at least one component) it follows

$$f_1(p^{\max}) < f_1(p^{\max} + \varepsilon d),$$

for all $\varepsilon > 0$. Then there exists a solution of (3.55).

Further let $p^{*,\max} \in \mathbb{R}^n$ be a solution of (3.55). Then the probabilistic constraint is always active, i.e. it is

$$\mathbb{P}(b \in M(p^{*,\max})) = \alpha.$$

The integral in (3.58) has to be understood as n dimensional integral over the set

$$\otimes_{j=1}^n [p_j^{\min}, x_j].$$

Remark 3.17. *Due to the convexity and the strict monotonicity of f_1 , for all $d \in \mathbb{R}_{\geq 0}^n$ (with $d > 0$ in at least one component) it follows*

$$\lim_{a \rightarrow \infty} f_1(p^{\max} + ad) = \infty. \quad (3.59)$$

with $a \in \mathbb{R}$.

Proof of Lemma 3.16. The assumption (3.58) implies that the set \mathcal{R} (defined in (3.57)) is nonempty. Due to the property (3.59) we can find $\bar{p} > p^{\min}$ (component-by-component) sufficiently large, s.t. for all $x \in \mathcal{R}$ with $x > \bar{p}$ in at least one component, there exists $y \in \mathcal{R}$, $y \leq \bar{p}$ (component-by-component) with

$$f_1(y) \leq f_1(x).$$

So if there exists a solution of (3.55), this solution can only be found in $\mathcal{R} \cap [p^{\min}, \bar{p}]$. Since the probabilistic constraint in (3.55) is not strict, the set $\mathcal{R} \cap [p^{\min}, \bar{p}]$ is compact and thus, a solution of (3.55) exists due to the *extreme value theorem* of Weierstrass.

Let $p^{*,\max}$ be a solution of (3.55) with

$$\mathbb{P}(b \in M(p^{*,\max})) > \alpha.$$

Then for all positive directions $d \in \mathbb{R}_{\geq 0}^n$ (with $d > 0$ in at least one component) there exists $\varepsilon > 0$ s.t.

$$\mathbb{P}(b \in M(p^{*,\max} - \varepsilon d)) = \alpha.$$

Due to the strict monotonicity this implies

$$f_1(p^{*,\max} - \varepsilon d) < f_1(p^{*,\max}),$$

so $p^{*,\max}$ cannot be a solution of (3.55). Thus, the probabilistic constraint has to be active for every solution of (3.55). \square

The assumption (3.58) is necessary to guarantee that the set of admissible upper pressure bounds \mathcal{R} is nonempty. The property (3.59) guarantees the existence of a solution. Just the convexity of f_1 is not sufficient, e.g. if we have

$$f_1(p^{\max}) = \exp(-p^{\max})$$

in one dimension, then the larger p^{\max} is, the smaller $f(p^{\max})$ gets. The strict monotonicity implies that the probabilistic constraint is active but it is not necessary. We could also assume the strict monotonicity of f_1 only on the set of admissible upper pressure bounds \mathcal{R} , that also implies that the probabilistic constraint is active. If we do not assume the strict monotonicity a solution can also exist inside of \mathcal{R} , e.g. for $x \in \mathcal{R}^\circ$ if we choose

$$f_1(p^{\max}) = \|p^{\max} - x\|,$$

then the solution of (3.55) is x (in the interior of \mathcal{R}) and the probabilistic constraint would be inactive.

Remark 3.18. *Motivated by the application, the strict monotonicity assumption is not very strong. The upper pressure bound p^{\max} can be seen as a measure for the costs of the pipes in the sense that the more pressure a pipe can take, the more material is needed to build the pipe, which increase the costs of the pipe.*

Next we analyze (3.56). As in (3.55) the probabilistic constraint can be written as

$$\int_{p_{\min}^{\max}} \varrho_{p(p_0)}(z) dz \geq \alpha,$$

where we write $\varrho_{p(p_0)}$ to explicitly state that the probability density function depends on the inlet pressure p_0 . The second constraint guarantees a non-negative inlet pressure, which of course is motivated by the application. This constraint can only be active if the pipes are friction free (i.e. $\lambda^e = 0$ for all $e \in \mathcal{E}$) and the lower pressure bound is equal to 0. Since this is never the case in reality we assume $\lambda^e > 0$ for all pipes $e \in \mathcal{E}$ and thus the second constraint in (3.56) cannot be active.

We assumed earlier, that the network graph is tree-structured, i.e. it has only one inflow node. So the probabilistic constrained optimization problem (3.56) is one-dimensional. We can state the following lemma:

Lemma 3.19. *Let pressure bounds $p^{\min}, p^{\max} \in \mathbb{R}^n$ with $p^{\max} > p^{\min}$ be given. Assume that there exists $x \geq 0$, s.t.*

$$\int_{p_{\min}^{\max}} \varrho_{p(x)}(z) dz \geq \alpha. \quad (3.60)$$

Then (3.56) has at least one solution. If f_2 is strictly monotone (increasing or decreasing), the solution is unique and the probabilistic constraint is always active.

Similar to (3.58) we need an assumption here to guarantee that the set of admissible inlet pressures is nonempty. In comparison to *Lemma 3.16* the convexity of the objective function is sufficient for the existence of a solution and the strict monotonicity guarantees both, that the solution is unique and that the probabilistic constraint is active.

Proof of Lemma 3.19. Since the pressure at the outflow nodes $p \in \mathbb{R}^n$ has a probability density function $\varrho_{p(p_0)}$, also the squared pressure $p^2 \in \mathbb{R}^n$ has a probability density function $\varrho_{p^2(p_0)}$ and due to the positivity of the probability density function and the pressure bounds it holds

$$\int_{p_{\min}^{\max}} \varrho_{p(p_0)}(z) dz = \int_{p_{\min^2}^{\max^2}} \varrho_{p^2(p_0)}(z) dz,$$

where $P_{\min}^{\max^2}$ is the cuboid of squared pressure bounds. As it is stated in *Corollary 3.6* the squared pressure at the nodes can be computed by

$$p^2 = \mathbb{1}_n p_0^2 - g(b),$$

One can see that only the pressure loss function g depends on the random vector b and the inlet pressure shifts the function g , so the shape of the probability density function of p^2 only depends on $g(b)$, not on p_0 . Due to assumption (3.60) we can find $\underline{p}, \bar{p} \in \mathbb{R}_{\geq 0}$, s.t.

$$\int_{P_{\min}^{\max^2}} \varrho_{p^2(\underline{p})}(z) dz = \int_{P_{\min}^{\max^2}} \varrho_{p^2(\bar{p})}(z) dz = \alpha,$$

and

$$\int_{P_{\min}^{\max^2}} \varrho_{p^2(x)}(z) dz < \alpha \quad \text{for all } x < \underline{p} \text{ or } x > \bar{p}.$$

So due to the *extreme value theorem* of Weierstrass the convex (and continuous) function f_2 has a minimum on the compact set $[\underline{p}, \bar{p}]$. An illustration of this idea in one dimension is shown in *Figure 3.11*. Further, if the function f_2 is strictly monotonous increasing, it takes its unique minimum in \underline{p} and if the function f_2 is strictly monotonous decreasing, it takes its unique minimum in \bar{p} , thus the probabilistic constraint is active. \square

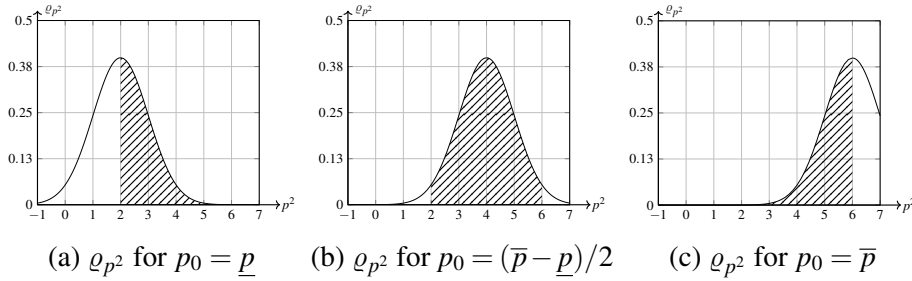


Figure 3.11.: One dimensional probability density function ϱ_{p^2} with $(p^{\min})^2 = 2$, $(p^{\max})^2 = 6$ and $\alpha = 0.5$ for different inlet pressures

The strict monotonicity assumption can be weakened in the sense that the statement of *Lemma 3.19* still holds if f_2 is only strictly monotone on the set of admissible pressures. If we assume f_2 to be strictly convex but not strictly monotone, we still can find a unique solution but the probabilistic constraint is not necessarily active.

Next in this section we consider probabilistic constrained optimization problems in which the probabilistic constraint is approximated by the KDE approach that was described in *Section 3.3*. For a probability level α consider the optimization problems

$$\begin{aligned} \min_{p^{\max} \in \mathbb{R}^n} \quad & f_1(p^{\max}) \\ \text{s.t.} \quad & \mathbb{P}_{\text{KDE}}(b \in M(p^{\max})) \geq \alpha, \\ & p^{\max} \geq p^{\min}, \end{aligned} \tag{3.61}$$

and

$$\begin{aligned} \min_{p_0 \in \mathbb{R}} \quad & f_2(p_0) \\ \text{s.t.} \quad & \mathbb{P}_{\text{KDE}}(b \in M(p_0)) \geq \alpha, \\ & p_0 \geq 0, \end{aligned} \tag{3.62}$$

where M is defined in (3.5). Both problems only differ from (3.55) and (3.56) in the approximated probabilistic constraint. Since the approximated probability density function is also positive, the proof for the existence of a solution is analogous.

Corollary 3.20. *Let a probability level $\alpha \in (0, 1)$, an inlet pressure $p_0 \in \mathbb{R}_{\geq 0}$ and a lower pressure bound $p^{\min} \in \mathbb{R}_{\geq 0}^n$ be given. Assume that there exists $x \geq p^{\min}$ (component-by-component), s.t.*

$$\int_{p^{\min}}^x \varrho_{p, N_{\text{KDE}}}(z) dz > \alpha. \tag{3.63}$$

Let f_1 be strictly monotone increasing in the sense that for all positive directions $d \in \mathbb{R}_{\geq 0}^n$ (with $d > 0$ in at least one component) it follows

$$f_1(p^{\max}) < f_1(p^{\max} + \varepsilon d),$$

for all $\varepsilon > 0$. Then there exists a solution of (3.61).

Further let $p^{*, \max} \in \mathbb{R}^n$ be a solution of (3.61). Then the probabilistic constraint is always active, i.e. it is

$$\mathbb{P}_{\text{KDE}}(b \in M(p^{\max})) = \alpha.$$

Corollary 3.21. *Let pressure bounds $p^{\min}, p^{\max} \in \mathbb{R}^n$ with $p^{\max} > p^{\min}$ be given. Assume that there exists $x \geq 0$, s.t.*

$$\int_{p_{\min}^{\max}}^x \varrho_{p(x), N_{\text{KDE}}}(z) dz \geq \alpha. \tag{3.64}$$

Then (3.62) has at least one solution. If f_2 is strictly monotone (increasing or decreasing), the solution is unique and the probabilistic constraint is always active.

We now know that the optimization problems (3.55) and (3.56) as well as (3.61) and (3.62) have a solution (under certain assumptions). We want to show that the solutions of the approximated problems (3.61) and (3.62) are useful in the sense that (under weak additional assumptions) they are close to the solutions of the exact problems (3.55) and (3.56) if the sampling size is sufficiently large. For $\alpha \in (0, 1)$ we define $g_1^\alpha, g_{1, N_{\text{KDE}}}^\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ with

$$\begin{aligned} g_1^\alpha : p^{\max} &\mapsto \alpha - \mathbb{P}(b \in M(p^{\max})), \\ g_{1, N_{\text{KDE}}}^\alpha : p^{\max} &\mapsto \alpha - \mathbb{P}_{N_{\text{KDE}}}(b \in M(p^{\max})), \end{aligned}$$

and $g_2^\alpha, g_{2, N_{\text{KDE}}}^\alpha : \mathbb{R} \rightarrow \mathbb{R}$ with

$$\begin{aligned} g_2^\alpha : p_0 &\mapsto \alpha - \mathbb{P}(b \in M(p_0)), \\ g_{2, N_{\text{KDE}}}^\alpha : p_0 &\mapsto \alpha - \mathbb{P}_{N_{\text{KDE}}}(b \in M(p_0)). \end{aligned}$$

Further we define

$$\begin{aligned} \mathcal{Z} &:= \{x \in \mathbb{R}^n \mid x \geq p^{\min} \text{ and } g_1^\alpha(x) = 0\}, \\ \mathcal{Z}_{N_{\text{KDE}}} &:= \{x \in \mathbb{R}^n \mid x \geq p^{\min} \text{ and } g_{1, N_{\text{KDE}}}^\alpha(x) = 0\}, \end{aligned}$$

as set of all points, for which the probabilistic constraint in (3.55) and the approximated probabilistic constraint in (3.61) are active, i.e. \mathcal{Z} contains the roots of g_1^α and $\mathcal{Z}_{N_{\text{KDE}}}$ contains the roots of $g_{1, N_{\text{KDE}}}^\alpha$. The following theorem is one of the key results of this thesis. It has not been published yet and it states that for sufficiently large N_{KDE} , the solution of the approximated problem (3.61) is (almost surely) close to the solution of the exact problem (3.55).

Theorem 3.22. *Let a probability level $\alpha \in (0, 1)$, an inlet pressure $p_0 \in \mathbb{R}_{\geq 0}$ and a lower pressure bound $p^{\min} \in \mathbb{R}_{\geq 0}^n$ be given. Assume that there exist $x \geq p^{\min}$ (component-by-component), s.t. (3.58) and (3.63) hold.*

Let f_1 be strictly monotone increasing in the sense of Lemma 3.16 and Corollary 3.20.

Assume that for all $x \in \mathcal{Z}$ there exist $d_1, d_2 \in \mathbb{R}^n \setminus \{0_n\}$, s.t. for all $\tau \in (0, 1)$ it holds

$$g_1^\alpha(x + \tau d_1) < 0 \quad \text{and} \quad g_1^\alpha(x + \tau d_2) > 0. \quad (3.65)$$

Let $p^{, \max}$ be a solution of (3.55). Assume that the solution is unique. Further assume that there exist $\delta, \varepsilon > 0$, s.t. for $p \in \mathcal{Z}$ with*

$$\|p^{*, \max} - p\| > \frac{\delta}{2},$$

it holds

$$|f(p^{*, \max}) - f(p)| > \varepsilon. \quad (3.66)$$

Then there exist a sufficiently large number N_{KDE} , s.t. the solution $p_{N_{\text{KDE}}}^{*,\max}$ of (3.61) is close to $p^{*,\max}$ in the sense that we have

$$\|p^{*,\max} - p_{N_{\text{KDE}}}^{*,\max}\| < \delta \quad \mathbb{P} - \text{almost surely.} \quad (3.67)$$

The statement in *Theorem 3.22* does not guarantee that the solution of (3.61) converges to the solution of (3.55) or it does not even guarantee convergence, but it guarantees that the solution of the approximated problem (3.61) is close to the solution of the exact problem (3.55). Note that $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^n .

Remark 3.23. *The assumption (3.65) guarantees, that no point can be both, a root of g_1^α and a minimum or maximum of g_1^α . This assumption is similar to the transversality assumption of the Mangasarian-Fromovitz constraint qualification (MFCQ).*

The uniqueness assumption is quite common in similar settings because otherwise it is not a priori clear which optimization algorithm would converge to which solution.

The last assumption (3.66) is slightly stronger than a uniqueness assumption but it guarantees that the objective function cannot get arbitrary close to its optimal value if its argument is not close to the optimal argument.

Proof of Lemma 3.7.

Since the proof is quite technical, we split the proof in three main parts. In the first part we show that for sufficiently large N_{KDE} every element in \mathcal{L} is close to an element in $\mathcal{L}_{N_{\text{KDE}}}$ and vice versa. In the second part we show that if elements of \mathcal{L} and $\mathcal{L}_{N_{\text{KDE}}}$ are close, also their objective function values are close and in the third part we prove (3.67) by contradiction. Since we are in a stochastic setting here, all results hold \mathbb{P} -almost surely. For the readers' convenience we do not explicitly mention that for every inequality. Let $\varepsilon, \delta > 0$ be given, s.t. (3.66) holds.

Part 1: In the first part of the proof we show that we can find N_{KDE} sufficiently large, s.t. for every root of g_1^α , i.e. for every $y \in \mathcal{L}$, there exists a root of $g_{1,N_{\text{KDE}}}^\alpha$, i.e. there exists $z \in \mathcal{L}_{N_{\text{KDE}}}$ with

$$\|y - z\| < \frac{\delta}{2}.$$

Let $y \in \mathcal{L}$ be a root of g_1^α . Then due to (3.65) there exist $d_1, d_2 \in \mathbb{R}^n \setminus \{0_n\}$ with $\|d_1\| = \|d_2\|$ and

$$g_1^\alpha(y + d_1) < 0 \quad \text{and} \quad g_1^\alpha(y + d_2) > 0.$$

We set $\tilde{\delta} := \min\{\frac{\delta}{2}, \|d_1\|\}$ and define

$$\begin{aligned}\varepsilon_- &:= g_1^\alpha\left(y + \tilde{\delta} \frac{d_1}{\|d_1\|}\right) \quad (< 0), \\ \text{and } \varepsilon_+ &:= g_1^\alpha\left(y + \tilde{\delta} \frac{d_2}{\|d_2\|}\right) \quad (> 0).\end{aligned}$$

We set $\tilde{\varepsilon} := \min\{\varepsilon_-, \varepsilon_+\}$. Then due to the almost surely convergence of $g_{1, N_{\text{KDE}}}^\alpha$ we can find $N_{\text{KDE},1}$ sufficiently large, s.t.

$$\begin{aligned}\left|g_1^\alpha\left(y + \tilde{\delta} \frac{d_1}{\|d_1\|}\right) - g_{1, N_{\text{KDE},1}}^\alpha\left(y + \tilde{\delta} \frac{d_1}{\|d_1\|}\right)\right| &\leq \frac{\tilde{\varepsilon}}{2}, \\ \left|g_1^\alpha\left(y + \tilde{\delta} \frac{d_2}{\|d_2\|}\right) - g_{1, N_{\text{KDE},1}}^\alpha\left(y + \tilde{\delta} \frac{d_2}{\|d_2\|}\right)\right| &\leq \frac{\tilde{\varepsilon}}{2}.\end{aligned}$$

This implies

$$g_{1, N_{\text{KDE},1}}^\alpha\left(y + \tilde{\delta} \frac{d_1}{\|d_1\|}\right) < 0 \quad \text{and} \quad g_{1, N_{\text{KDE},1}}^\alpha\left(y + \tilde{\delta} \frac{d_2}{\|d_2\|}\right) > 0,$$

and thus, due to the continuity of $g_{1, N_{\text{KDE}}}^\alpha$ there exists $z \in \mathcal{L}_{N_{\text{KDE}}}$ with

$$\|y - z\| < \frac{\delta}{2}. \quad (3.68)$$

Further since there exists no convergent sequence $(y_i)_{i \geq 1}$ in \mathbb{R}^n with limit $\lim_{i \rightarrow \infty} y_i \notin \mathcal{L}$ and $\lim_{i \rightarrow \infty} g_1^\alpha(y_i) = 0$ (i.e. apart from a neighborhood of \mathcal{L} , g_1^α cannot get arbitrary close to 0), we can find $N_{\text{KDE},2}$ sufficiently large, s.t. every root of $g_{1, N_{\text{KDE},2}}^\alpha$ is close to a root of g_1^α , i.e. for every $z \in \mathcal{L}_{N_{\text{KDE}}}$, there exists $y \in \mathcal{L}$ with

$$\|z - y\| < \frac{\delta}{2} \quad (3.69)$$

Thus the first part of the proof is done.

Part 2: In the second part of the proof we show, that if elements of \mathcal{L} and $\mathcal{L}_{N_{\text{KDE}}}$ are close, also their objective function values are close. Consider $y \in \mathcal{L}$ and $z \in \mathcal{L}_{N_{\text{KDE}}}$ with

$$z = \min_{x \in \mathcal{L}_{N_{\text{KDE}}}} \|y - x\|,$$

i.e. there exists no element in $\mathcal{L}_{N_{\text{KDE}}}$, that is closer to x . Since f is continuous, we can find $N_{\text{KDE},3}$ sufficiently large that for $\varepsilon > 0$ (given by (3.66)) we can find $\delta^* > 0$, s.t.

$$\|y - z\| < \delta^* \quad \text{and} \quad |f(y) - f(z)| \leq \frac{\varepsilon}{2}. \quad (3.70)$$

The other way around for $z \in \mathcal{Z}_{N_{\text{KDE}}}$ consider $y \in \mathcal{Z}$ with

$$y = \min_{x \in \mathcal{Z}} \|z - x\|,$$

as root of g_1^α with minimal distance to z . Due to (3.69) there exists $N_{\text{KDE},4}$ sufficiently large that for $\varepsilon > 0$ we can find $\delta^* > 0$, s.t.

$$\|z - y\| \leq \delta^* \quad \text{and} \quad |f(z) - f(y)| \leq \frac{\varepsilon}{2}, \quad (3.71)$$

and thus the second part of the proof is done.

Part 3: In the last part of the proof we show that (3.67) holds. For this part we use a proof by contradiction combined with a proof by cases. For

$$N_{\text{KDE}} := \max\{N_{\text{KDE},1}, N_{\text{KDE},2}, N_{\text{KDE},3}, N_{\text{KDE},4}\},$$

let $p_{N_{\text{KDE}}}^{*,\max}$ be the solution of (3.61). Assume that

$$\|p^{*,\max} - p_{N_{\text{KDE}}}^{*,\max}\| \geq \delta,$$

where δ comes from the assumption (3.66).

Part 3 i) Assume that

$$f(p_{N_{\text{KDE}}}^{*,\max}) \geq f(p^{*,\max}).$$

Then due to (3.68) and (3.70) there exists $\tilde{p}_{N_{\text{KDE}}} \in \mathcal{Z}_{N_{\text{KDE}}}$ with

$$\|\tilde{p}_{N_{\text{KDE}}} - p^{*,\max}\| < \frac{\delta}{2} \quad \text{and} \quad |f(\tilde{p}_{N_{\text{KDE}}}) - f(p^{*,\max})| < \frac{\varepsilon}{2}.$$

We know that $f(\tilde{p}_{N_{\text{KDE}}})$ must be larger than or equal to $f(p^{*,\max})$, because otherwise we would have

$$f(\tilde{p}_{N_{\text{KDE}}}) < f(p^{*,\max}) \leq f(p_{N_{\text{KDE}}}^{*,\max}),$$

and thus $p_{N_{\text{KDE}}}^{*,\max}$ cannot be a solution of (3.61). It follows

$$f(\tilde{p}_{N_{\text{KDE}}}) < f(p^{*,\max}) + \frac{\varepsilon}{2}. \quad (3.72)$$

Further due to (3.69) and (3.71) there exists $\tilde{p} \in \mathcal{Z}$ with

$$\|\tilde{p} - p_{N_{\text{KDE}}}^{*,\max}\| < \frac{\delta}{2} \quad \text{and} \quad |f(\tilde{p}) - f(p_{N_{\text{KDE}}}^{*,\max})| < \frac{\varepsilon}{2}.$$

Since we want to estimate $f(\tilde{p})$ from above, we have

$$f(\tilde{p}) < f(p_{N_{\text{KDE}}}^{*,\max}) + \frac{\varepsilon}{2}. \quad (3.73)$$

Then since $\|\tilde{p} - p^{*,\max}\| > \frac{\delta}{2}$, with (3.66) and $f(p^{*,\max}) < f(\tilde{p})$ (otherwise $p^{*,\max}$ would not be a solution of (3.55)) we have

$$f(p^{*,\max}) < f(\tilde{p}) - \varepsilon. \quad (3.74)$$

Now we can combine the derived inequalities where we start with (3.72), we first insert (3.74) and then we insert (3.73). It follows

$$\begin{aligned} f(\tilde{p}_{N_{\text{KDE}}}) &< f(p^{*,\max}) + \frac{\varepsilon}{2} \\ &< f(\tilde{p}) - \varepsilon + \frac{\varepsilon}{2} \\ &< f(p_{N_{\text{KDE}}}^{*,\max}) + \frac{\varepsilon}{2} - \varepsilon + \frac{\varepsilon}{2}. \end{aligned}$$

Thus we have $f(\tilde{p}_{N_{\text{KDE}}}) < f(p_{N_{\text{KDE}}}^{*,\max})$, which is a contradiction to the fact that $p_{N_{\text{KDE}}}^{*,\max}$ is a solution of (3.61). So (3.67) follows for this part.

Part 3 ii) In the end of *Part 3* we assume

$$f(p_{N_{\text{KDE}}}^{*,\max}) < f(p^{*,\max}).$$

Due to (3.68) and (3.70) there exists $\tilde{p}_{N_{\text{KDE}}} \in \mathcal{L}_{N_{\text{KDE}}}$ with

$$\|\tilde{p}_{N_{\text{KDE}}} - p^{*,\max}\| < \frac{\delta}{2} \quad \text{and} \quad |f(\tilde{p}_{N_{\text{KDE}}}) - f(p^{*,\max})| < \frac{\varepsilon}{2}.$$

Since we want to estimate $f(p^{*,\max})$ from below, we have

$$f(p^{*,\max}) > f(\tilde{p}_{N_{\text{KDE}}}) - \frac{\varepsilon}{2}. \quad (3.75)$$

Further due to (3.69) and (3.71) there exists $\tilde{p} \in \mathcal{L}$ with

$$\|\tilde{p} - p_{N_{\text{KDE}}}^{*,\max}\| < \frac{\delta}{2} \quad \text{and} \quad |f(\tilde{p}) - f(p_{N_{\text{KDE}}}^{*,\max})| < \frac{\varepsilon}{2}.$$

We know that $f(\tilde{p})$ must be larger than or equal to $f(p_{N_{\text{KDE}}}^{*,\max})$, because otherwise we would have

$$f(\tilde{p}) < f(p_{N_{\text{KDE}}}^{*,\max}) < f(p^{*,\max}),$$

and thus, $p^{*,\max}$ cannot be a solution of (3.55). It follows

$$f(p_{N_{\text{KDE}}}^{*,\max}) > f(\tilde{p}) - \frac{\varepsilon}{2}. \quad (3.76)$$

Then since $\|\tilde{p} - p^{*,\max}\| > \frac{\delta}{2}$, with (3.66) and $f(p^{*,\max}) < f(\tilde{p})$ (otherwise $p^{*,\max}$ would not be a solution of (3.55)) we have

$$f(\tilde{p}) > f(p^{*,\max}) + \varepsilon. \quad (3.77)$$

We again combine the derived inequalities where we start with (3.76), we first insert (3.77) and then we insert (3.75). It follows

$$\begin{aligned} f(p_{N_{\text{KDE}}}^{*,\max}) &> f(\tilde{p}) - \frac{\varepsilon}{2} \\ &> f(p^{*,\max}) + \varepsilon - \frac{\varepsilon}{2} \\ &> f(\tilde{p}_{N_{\text{KDE}}}) - \frac{\varepsilon}{2} + \varepsilon - \frac{\varepsilon}{2}. \end{aligned}$$

Thus we have $f(p_{N_{\text{KDE}}}^{*,\max}) > f(\tilde{p}_{N_{\text{KDE}}})$, which is a contradiction to the fact that $p_{N_{\text{KDE}}}^{*,\max}$ is a solution of (3.61).

So for $N_{\text{KDE}} := \max\{N_{\text{KDE},1}, N_{\text{KDE},2}, N_{\text{KDE},3}, N_{\text{KDE},4}\}$ it follows

$$\|p^{*,\max} - p_{N_{\text{KDE}}}^{*,\max}\| < \delta \quad \mathbb{P}\text{-almost surely,}$$

and the proof is complete. \square

For the solutions of the optimization problems (3.56) and (3.62) we can get a similar result as *Theorem 3.22* but since the solutions are one dimensional, the proof is slightly easier.

Theorem 3.24. *Consider $\delta > 0$ and let pressure bounds $p^{\min}, p^{\max} \in \mathbb{R}^n$ with $p^{\max} > p^{\min}$ be given. Assume that there exists $x \geq 0$, s.t. (3.60) and (3.64) hold. Assume that f_2 is strictly monotone (increasing or decreasing). Let p_0^* be the solution of (3.56) and let $p_{0,N_{\text{KDE}}}^*$ be the solution of (3.62). If for all $x \geq 0$ with $g_2^\alpha(x) = 0$ there exists $\varepsilon > 0$, s.t.*

$$\text{sgn}(g_2^\alpha(x - \tau)) = -\text{sgn}(g_2^\alpha(x + \tau)) \quad \forall \tau \in (0, \varepsilon),$$

where sgn is the sign function, then for N_{KDE} sufficiently large it follows

$$p_{0,N_{\text{KDE}}}^* \in (p_0^* - \delta, p_0^* + \delta).$$

Proof. Let $\delta > 0$ be given. We first show that we can find N_{KDE} sufficiently large s.t. $g_{2,N_{\text{KDE}}}^\alpha$ has a root close to every root of g_2^α . Let y be a root of g_2^α . We define

$$\begin{aligned} \varepsilon_- &:= g_2^\alpha\left(y - \frac{\delta}{2}\right), \\ \text{and } \varepsilon_+ &:= g_2^\alpha\left(y + \frac{\delta}{2}\right). \end{aligned}$$

We set $\tilde{\varepsilon} := \min\{\varepsilon_-, \varepsilon_+\}$ and choose $N_{\text{KDE},1}$ sufficiently large, s.t.

$$\left| g_2^\alpha - g_{2,N_{\text{KDE},1}}^\alpha \right| \leq \frac{\tilde{\varepsilon}}{2} \quad \mathbb{P} - \text{almost surely.}$$

Thus we have

$$\left| g_2^\alpha \left(y - \frac{\delta}{2} \right) - g_{2,N_{\text{KDE},1}}^\alpha \left(y - \frac{\delta}{2} \right) \right| \leq \frac{\tilde{\varepsilon}}{2} \quad \mathbb{P} - \text{almost surely,}$$

which implies

$$g_{2,N_{\text{KDE},1}}^\alpha \left(y - \frac{\delta}{2} \right) \in \left[\varepsilon_- - \frac{\tilde{\varepsilon}}{2}, \varepsilon_- + \frac{\tilde{\varepsilon}}{2} \right] \subseteq \mathbb{R}_- \quad \mathbb{P} - \text{almost surely.}$$

Equivalently we have

$$\left| g_2^\alpha \left(y + \frac{\delta}{2} \right) - g_{2,N_{\text{KDE},1}}^\alpha \left(y + \frac{\delta}{2} \right) \right| \leq \frac{\tilde{\varepsilon}}{2} \quad \mathbb{P} - \text{almost surely,}$$

which implies

$$g_{2,N_{\text{KDE},1}}^\alpha \left(y + \frac{\delta}{2} \right) \in \left[\varepsilon_+ - \frac{\tilde{\varepsilon}}{2}, \varepsilon_+ + \frac{\tilde{\varepsilon}}{2} \right] \subseteq \mathbb{R}_+ \quad \mathbb{P} - \text{almost surely.}$$

Thus $g_{2,N_{\text{KDE}}}^\alpha$ must have a root in $\left(y - \frac{\delta}{2}, y + \frac{\delta}{2} \right)$ (due to the continuity). Further we can choose $N_{\text{KDE},2}$ sufficiently large, s.t. $g_{2,N_{\text{KDE},2}}^\alpha$ has the same number of roots as g_2^α (\mathbb{P} -almost surely). Last we can choose $N_{\text{KDE},3}$ sufficiently large, s.t. for every root \tilde{y} of $g_{2,N_{\text{KDE},3}}^\alpha$, the constraint g_2^α has also a root in $\left(\tilde{y} - \frac{\delta}{2}, \tilde{y} + \frac{\delta}{2} \right)$ (almost surely).

Let p_0^* be the solution of (3.56) and let $p_{0,N_{\text{KDE}}}^*$ be the solution of (3.62) for

$$N_{\text{KDE}} = \max\{N_{\text{KDE},1}, N_{\text{KDE},2}, N_{\text{KDE},3}\}.$$

Due to the strict monotonicity of f_2 we have

$$g_2^\alpha(p_0^*) = 0 = g_{2,N_{\text{KDE}}}^\alpha(p_{0,N_{\text{KDE}}}^*).$$

We assume, that $p_{0,N_{\text{KDE}}}^* \notin (p_0^* - \delta, p_0^* + \delta)$. If $p_{0,N_{\text{KDE}}}^* < p_0^*$ then g_2^α has a root $z_1 \in \left(p_{0,N_{\text{KDE}}}^* - \frac{\delta}{2}, p_{0,N_{\text{KDE}}}^* + \frac{\delta}{2} \right)$ with $f_2(z_1) < f_2(p_0^*)$. That implies that p_0^* cannot be the solution of (3.56).

Otherwise if $p_{0,N_{\text{KDE}}}^* > p_0^*$ then $g_{2,N_{\text{KDE}}}^\alpha$ has a root $z_2 \in \left(p_0^* - \frac{\delta}{2}, p_0^* + \frac{\delta}{2} \right)$ with $f_2(z_2) < f_2(p_{0,N_{\text{KDE}}}^*)$. That implies that $p_{0,N_{\text{KDE}}}^*$ cannot be the solution of (3.62). Consequently we have

$$p_{0,N_{\text{KDE}}}^* \in (p_0^* - \delta, p_0^* + \delta) \quad \text{almost surely,}$$

and the theorem is proven. \square

Remark 3.25. *It could happen that the solution of the approximated problem (3.61) resp. (3.62) is outside the admissible set of the exact problem (3.55) resp. (3.56) but due to Theorem 3.22 resp. Theorem 3.24 the solution of the approximated problem is close to the admissible set and close to the solution of the exact problem.*

Remark 3.26. *If $p^{*,max}$ is the solution of (3.55) and both constraints are inactive in $p^{*,max}$ (this can happen e.g. when f_1 is not strictly monotone), then for N_{KDE} sufficiently large $p^{*,max}$ is also the solution of (3.61). The same holds for (3.56) and (3.62).*

Last in this section we derive necessary optimality conditions for the approximated probabilistic constrained optimization problems (3.61) and (3.62). Hence we compute the derivative of the probabilistic constraint in (3.61) and (3.62). We have

$$\begin{aligned} \mathbb{P}_{N_{KDE}}(b \in M(p^{\max})) &= \\ &= \frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \prod_{j=1}^n \left[\operatorname{erf}(\varphi_{i,j}(p_j^{\max})) - \operatorname{erf}(\varphi_{i,j}(p_j^{\min})) \right], \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}_{N_{KDE}}(b \in M(p_0)) &= \\ &= \frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \prod_{j=1}^n \left[\operatorname{erf}(\varphi_{i,j}(p_j^{\max}, p_0)) - \operatorname{erf}(\varphi_{i,j}(p_j^{\min}, p_0)) \right]. \end{aligned}$$

We explicitly state the dependence of p_0 where it is necessary. Now we compute the partial derivative of g_1^α with respect to p_k^{\max} ($k \in \{1, \dots, n\}$). We have

$$\begin{aligned} \frac{\partial}{\partial p_j^{\max}} g_1^\alpha(p^{\max}) &= -\frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \left(\prod_{j=1, j \neq k}^n \left[\operatorname{erf}(\varphi_{i,j}(p_j^{\max})) - \operatorname{erf}(\varphi_{i,j}(p_j^{\min})) \right] \right) \\ &\quad \cdot \frac{d}{dp_k^{\max}} \operatorname{erf}(\varphi_{i,k}(p_k^{\max})). \end{aligned}$$

The derivative of the Gauss error function is given by

$$\frac{d}{dx} \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \exp(-x^2),$$

and the derivative of $\varphi_{i,k}$ with respect to p_k^{\max} is given by

$$\frac{d}{dp_k^{\max}} \varphi_{i,k}(p_k^{\max}) = \frac{1}{\sqrt{2} h_k}.$$

Thus it follows

$$\begin{aligned} \frac{\partial}{\partial p_j^{\max}} g_1^\alpha(p^{\max}) &= -\frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \left(\prod_{j=1, j \neq k}^n \left[\operatorname{erf}(\varphi_{i,j}(p_j^{\max})) - \operatorname{erf}(\varphi_{i,j}(p_j^{\min})) \right] \right. \\ &\quad \left. \cdot \frac{\sqrt{2}}{\sqrt{\pi} h_k} \exp(-\varphi_{i,k}^2(p_k^{\max})) \right). \end{aligned} \quad (3.78)$$

Then the k -th component of the gradient $\nabla g_1^\alpha(p^{\max}) \in \mathbb{R}^n$ is given by (3.78).

Next we compute the derivative of g_2^α with respect to p_0 . We have

$$\frac{d}{dp_0} g_2^\alpha(p_0) = -\frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \frac{d}{dp_0} \prod_{j=1}^n \left[\operatorname{erf}(\varphi_{i,j}(p_j^{\max}, p_0)) - \operatorname{erf}(\varphi_{i,j}(p_j^{\min}, p_0)) \right].$$

Since $\varphi_{i,j}(\cdot)$ depends on p_0 for all $j = 1, \dots, n$, the derivative can be computed by the product rule. It follows:

$$\begin{aligned} \frac{d}{dp_0} g_2^\alpha(p_0) &= -\frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \sum_{k=1}^n \prod_{j=1, j \neq k}^n \left[\operatorname{erf}(\varphi_{i,j}(p_j^{\max}, p_0)) - \operatorname{erf}(\varphi_{i,j}(p_j^{\min}, p_0)) \right] \\ &\quad \cdot \frac{d}{dp_0} \left[\operatorname{erf}(\varphi_{i,k}(p_k^{\max}, p_0)) - \operatorname{erf}(\varphi_{i,k}(p_k^{\min}, p_0)) \right]. \end{aligned}$$

The derivative of $\varphi_{i,k}(\cdot)$ with respect to p_0 is given by

$$\frac{\partial}{\partial p_0} \varphi_{i,k}(\cdot, p_0) = -\frac{1}{\sqrt{2} h_k} \frac{d}{dp_0} p_k(b^{S,i}).$$

Hence we have

$$\begin{aligned} \frac{d}{dp_0} g_2^\alpha(p_0) &= -\frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \sum_{k=1}^n \prod_{j=1, j \neq k}^n \left[\operatorname{erf}(\varphi_{i,j}(p_j^{\max}, p_0)) - \operatorname{erf}(\varphi_{i,j}(p_j^{\min}, p_0)) \right] \\ &\quad \cdot \frac{\sqrt{2}}{\sqrt{\pi} h_k} \left[-\exp(-\varphi_{i,k}^2(p_k^{\max}, p_0)) + \exp(-\varphi_{i,k}^2(p_k^{\min}, p_0)) \right] \frac{d}{dp_0} p_k(b^{S,i}, p_0). \end{aligned} \quad (3.79)$$

Our aim now is to formulate necessary optimality conditions for the probabilistic constrained optimization problems (3.61) and (3.62).

Remark 3.27. *Since the constraint $p^{\max} \geq p^{\min}$ in (3.61) and (3.62) can never be active, the probabilistic constraint $\mathbb{P}(b \in M(p^{\max})) \geq \alpha$ resp. $\mathbb{P}(b \in M(p_0))$ is the only constraint that can be active, thus the linear independence constraint qualification (LICQ) holds for both problems (3.61) and (3.62).*

With this, we can formulate the necessary optimality conditions for the approximated probabilistic constrained optimization problems (3.61) and (3.62):

Corollary 3.28. *Let $p^{*,max} \in \mathbb{R}^n$ be a (local) optimal solution of (3.61). Since the LICQ holds in $p^{*,max}$ (cf. Remark 3.27), there exists a multiplier $\mu^* \geq 0$, s.t.*

$$\begin{aligned}\nabla f_1(p^{*,max}) + \mu^* \nabla g_1^\alpha(p^{*,max}) &= 0, \\ g_1^\alpha(p^{*,max}) &\leq 0, \\ \mu^* g_1^\alpha(p^{*,max}) &= 0,\end{aligned}$$

where the components of the gradient of g_1^α are stated in (3.78). Thus $(p^{*,max}, \mu^*) \in \mathbb{R}^n \times \mathbb{R}$ is a Karush-Kuhn-Tucker point.

Corollary 3.29. *Let $p_0^* \in \mathbb{R}$ be a (local) optimal solution of (3.62). Since the LICQ holds in p_0^* (cf. Remark 3.27), there exists a multiplier $\mu^* \geq 0$, s.t.*

$$\begin{aligned}f'(p_0^*) + \mu^* (g_2^\alpha)'(p_0^*) &= 0, \\ g_2^\alpha(p_0^*) &\leq 0, \\ \mu^* g_2^\alpha(p_0^*) &= 0,\end{aligned}$$

where the derivative of g_2^α is stated in (3.79). Thus $(p_0^*, \mu^*) \in \mathbb{R} \times \mathbb{R}$ is a Karush-Kuhn-Tucker point.

The necessary optimality conditions *Corollary 3.28* and *Corollary 3.29* can be used to characterize the solution of the approximated problems (3.61) and (3.62). Since the set of optimal solutions is a subset of all points that satisfy necessary optimality conditions, the points that satisfy the conditions in *Corollary 3.28* and *Corollary 3.29* are good candidates for the solution of (3.61) and (3.62). If the set of admissible pressures \mathcal{R} (defined in (3.57)) is convex and f_1 resp. f_2 are strictly convex, the necessary optimality conditions are also sufficient which means that the set of optimal solutions is identical to the set of points that satisfy the necessary optimality conditions. A scheme of this is shown in *Figure 3.12*.

As a perspective for the future it would make sense to analyze the relation between the lower two nodes in *Figure 3.12*, i.e. to analyze whether all points that satisfy the necessary optimality conditions for the approximated problems (3.61) and (3.62) are close to the points that satisfy the necessary optimality conditions for the original problems (3.55) and (3.56).

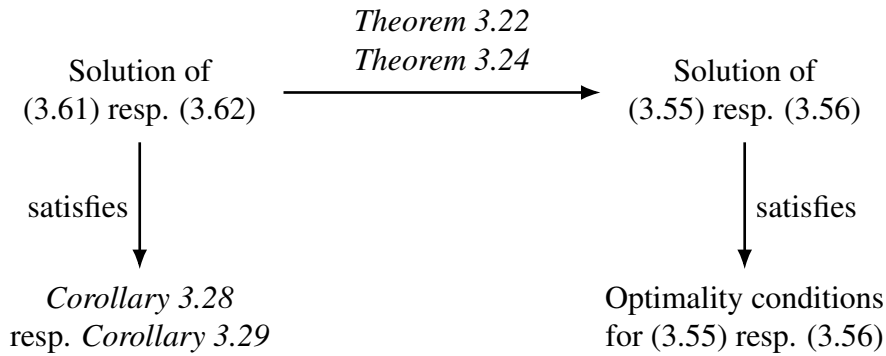


Figure 3.12.: Scheme on the relation of the necessary optimality conditions

3.7. Summary, Advantages and Disadvantages of the SRD and the KDE

In this chapter we derived two different approaches to compute the probability that the solution of a stationary hyperbolic balance law satisfies box constraints. Both approaches, the KDE approach and the SRD approach were introduced in *Chapter 1* and applied to gas networks in this chapter. Surely both methods can easily be used for a larger class of partial differential equations with uncertain boundary data, if the probabilistic constraint is time independent but we focused on stationary gas transport in pipeline networks here. In the next chapter we focus on dynamic partial differential equations and time dependent probabilistic constraints.

Here, in the last section of this chapter we shortly discuss the advantages and disadvantages of both methods, beginning with the SRD approach. The SRD approach was based on a characterization of the feasible set in terms of inequalities (see *Lemma 3.7* and *Theorem 3.13*). These inequalities allow us to simply check whether a boundary vector is feasible or not. The SRD approach decreases the dimension of integration by one and due to precise numerical integration methods the quadrature error can be kept small. A big disadvantage of the SRD approach is, that an explicit solution of the system is needed since this solution is the basis for the characterization. That means for more complex systems, where we might not get an explicit solution, we could get in trouble using the SRD. However if an explicit solution is known, the SRD is a quite powerful tool in this context. In *Section 3.5* we applied the SRD to a more complex setting where one boundary condition (the inlet pressure) was not explicitly stated, we only defined box constraints for the

inlet pressure. So feasibility not only meant „Does the solution satisfy certain box constraints?“, but „When can we find a solution that satisfies the box constraints?“. So for random boundary data (the load vector) we showed how to compute the probability that a solution exists which satisfies the box constraints. This is where problems occur when we use the KDE approach.

The KDE approach was applied to our setting as an alternative to the SRD approach. The biggest advantage of the KDE approach is, that an explicit solution is not necessarily required, a numerical solver is sufficient for the KDE. Since there are many good numerical methods for solving partial differential equations accurately (e.g. finite element methods or finite volume methods), this is not a problem. We defined an independent and identically distributed sampling of the random variable, computed the solution of the system for every sampled point and collected the computed points in another sampling. With this sampling we estimated the probability density function of the state that is required to satisfy the box constraints. With wisely chosen kernel functions the integral of the estimated probability density function over the box constraints was easy to compute since the box constraints define a n dimensional cuboid. A slight disadvantage of the KDE approach (compared to the SRD approach) is that the convergence theory only holds \mathbb{P} -almost surely. That means that the sampling of the random variable could be bad in the sense that it does not represent the random variable properly, but this is very unlikely to happen for a large number of samples. Another important point is that we could not apply the KDE to the setting of *Section 3.5* because since one boundary condition was not defined, we were not able to compute the sampling of the state variable that was required to satisfy the box constraints (pressure bounds).

In *Section 3.6* we analyzed certain probabilistic constrained optimization problems and their corresponding approximated optimization problems, where the probability was approximated by the KDE approach. One of the key results of this thesis is, that we proved we can find a sufficiently large sampling size, s.t. the optimal solution of the approximated problem is \mathbb{P} -almost surely close to the exact solution. Further due to the explicit representation of the approximated probability density function we could compute the derivative of the approximated probabilistic constraint, which is always an advantage for numerical solvers. Differentiability of probabilistic constraints that are reformulated with the SRD approach was analyzed e.g. in [van Ackooij and Henrion, 2014a], [van Ackooij and Henrion, 2014b], [van Ackooij et al., 2018] and [van Ackooij et al., 2020]. Since the KDE approach was not used in

the context of probabilistic constrained optimization with partial differential equations yet we focus on dynamic systems and time dependent probabilistic constraints in the next chapter. In [Caillaud et al., 2018] the authors use the KDE approach for probabilistic constrained optimization problems with ordinary differential equations but an analysis of the preciseness of the optimal solutions has not been done before.

4. Probabilistic Constrained Optimization on Dynamic Flow Networks

In this chapter we consider gas network dynamics. We introduce a model for the dynamic gas flow in pipeline networks and we discuss the existence of a solution for this model. We first discuss time dependent probabilistic constraints and time dependent uncertainty. Then we use the KDE approach (see *Section 1.1*) to compute the probability that random gas outflows satisfy certain constraints in the time interval $[0, T]$. Further we use the results from the stationary case of *Chapter 3* to derive necessary optimality conditions for a probabilistic constrained optimization problem in the context of dynamic gas transport in pipeline networks. Last in this chapter we present a numerical example for the gas flow in a single pipe.

4.1. Time dependent Probabilistic Constraints

In this section, we discuss probabilistic constraints in a dynamic context. This topic has also been discussed in [Schuster et al., 2021, Section 3.1].

There is a slight difference to the stationary case since time dependencies occur, so e.g. the feasible set can be time dependent. If we fix a point in time $t^* \in [0, T]$, a similar procedure as in the stationary case can be used to compute probabilistic constraints, but if we do not fix a point in time, the question is how probabilistic constraints can be understood for a time period $[0, T]$. Let a probability level $\alpha \in (0, 1)$ be given. Usually α is chosen large, close to 1. For a time dependent uncertain boundary function $b(t)$ and a (time dependent) feasible set $M(t)$, the formulation of time dependent probabilistic constraints that we will use later is

$$\mathbb{P}(b \in M(t) \forall t \in [0, T]) \geq \alpha. \quad (4.1)$$

This is a strong condition. It means, we want to guarantee, that a percentage α of all possible random boundary functions (in an appropriate probability

space $(\Omega, \mathcal{A}, \mathbb{P})$) is feasible in every point in time $t \in [0, T]$. So we do not allow any violation of the feasibility for most of the random boundary functions. In the context of gas transport this is a reasonable condition since the pressure bounds are fix and do not allow violations. In fact the formulation (4.1) is a so-called probust constraint, i.e. a constraint that contains both, a probabilistic and a robust part. The class of probust constraints has been developed recently. Probust constraints are currently of great interest in research (see e.g. [González-Gradón et al., 2017], [van Ackooij et al., 2020] and [Adelhütte et al., 2020]).

Another possible formulation for time dependent probabilistic constraints is

$$\mathbb{P}(b \in M(t)) \geq \alpha \quad \forall t \in [0, T]. \quad (4.2)$$

Compared to (4.1) this means that we want all possible scenarios to be feasible with a percentage of α for all times. So we allow a small violation of the feasibility in every point in time. For the application this would mean that we have a relatively good working system in the complete time period. But for our applications this is not enough: Roughly speaking we want a perfect working system under “normal“ conditions, we only „allow“ problems when something really unpredictable happens, which means that problems should only occur under unlikely conditions. This is given by (4.1). Constraints of the form (4.2) have been discussed e.g. in [van Ackooij et al., 2016] and [Roald et al., 2020].

A third possible formulation for time dependent probabilistic constraints is

$$\frac{1}{T} \int_0^T \mathbb{P}(b \in M(t)) dt \geq \alpha.$$

This is an average value condition for time dependent probabilistic constraints. This condition guarantee that the average scenario is feasible with a percentage α over the complete time interval. So for short time intervals inside $[0, T]$, the scenario can be better than α but for some other intervals, it can be even worse. Similar to (4.2) this formulation is not very appropriate for our applications, this fits more to e.g. finance mathematics.

4.2. Uncertain Time dependent Functions

In this section we discuss time dependent uncertainty. We first illustrate difficulties in this area in general. An idea for uncertain time dependent functions is the following:

Let a function

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f : t \mapsto f(t),$$

be given and consider a family of random variables

$$(\xi_t)_{t \in [0, T]},$$

on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then for $\omega \in \Omega$ we define the random function

$$f^\omega(t) = \xi_t(\omega)f(t).$$

For an appropriate choice of $(\xi_t)_{t \in [0, T]}$, f^ω is a slight variation of f but this variation implies that f^ω has no regularity properties anymore. Even if f is continuously differentiable, f^ω is not even continuous almost surely. Since we later use random functions as uncertain boundary data for a hyperbolic balance law, problems occur due to the lack of regularity.

Another idea was recently used in [Farshbaf-Shaker et al., 2020]. There the authors use random Fourier series to create random scenarios. This idea is not new, in the 1930s *R. Paley and A. Zygmund* studied random Fourier series of the form

$$\sum_{n=0}^{\infty} c_n \varepsilon_n \exp(inx), \tag{4.3}$$

where $|c_n|$ are real numbers and $\{\varepsilon_n\}$ is a sequence of independent *Rademacher* distributed random variables (see [Paley and Zygmund, 1930] and [Paley and Zygmund, 1932]). They proved the almost surely uniform convergence of (4.3) under weak assumptions. Further they showed that if $\sum_{n=0}^{\infty} c_n \exp(inx)$ is in L^2 , then (4.3) is in L^p (for $2 \leq p < \infty$) almost surely. The consequential question was, whether the results of Paley and Zygmund also hold for a larger class of distributions and the answer is yes. Let $(\xi_n)_{n \geq 0}$ be a sequence of independent random variables and consider the random Fourier series

$$\sum_{n=0}^{\infty} c_n \xi_n \exp(inx). \tag{4.4}$$

In [Marcus and Pisier, 1981] the authors claim the almost surely convergence of (4.4) for every sequence of independent centered real-valued random variables. Further in [Imekraz et al., 2016] the authors consider independent and

identically distributed random variables and they show that (4.4) converges almost surely in L^p iff (4.3) converges almost surely in L^p for $p \in [2, \infty)$. A detailed introduction and a wider overview about the topic of random Fourier series can be found in [Marcus and Pisier, 1981].

As mentioned before, our approach comes from [Farshbaf-Shaker et al., 2020]. Let a function $f \in L^2([0, T])$ with $f(0) = 0$ be given. For $m = 0, 1, 2, \dots$, we define the orthonormal series

$$\psi_m(t) := \frac{\sqrt{2}}{\sqrt{T}} \sin\left(\left(\frac{\pi}{2} + m\pi\right) \frac{t}{T}\right), \quad (4.5)$$

and the coefficients

$$\alpha_m^0(f) := \int_0^T f(t) \psi_m(t) dt. \quad (4.6)$$

Then the Fourier series representation of $f(t)$ is given by

$$f(t) = \sum_{m=0}^{\infty} \alpha_m^0(f) \psi_m(t). \quad (4.7)$$

Since it is $\psi_m(0) = 0$ and thus,

$$\sum_{m=0}^{\infty} \alpha_m^0(f) \psi_m(0) = 0,$$

we also assume $f(0) = 0$. For the implementation we truncate the Fourier series after $N_F \in \mathbb{N}$ terms, i.e.

$$f_{N_F}(t) := \sum_{m=0}^{N_F} \alpha_m^0 \psi_m(t). \quad (4.8)$$

Remark 4.1. *The accuracy of the truncated Fourier series strongly depends on the data $f(t)$ and on the number N_F . Usually N_F is chosen s.t. the truncation error is small. One criterion to find a sufficiently large number N_F to guarantee the desired accuracy is to state a bound for the L^2 truncation error. For $\vartheta \in (0, 1)$ in [Schuster et al., 2021] the authors require that the L^2 truncation error should depend on the L^2 -norm of f , i.e. that N_F is chosen sufficiently large, s.t.*

$$\|f(t) - f_{N_F}(t)\|_{L^2}^2 \leq \vartheta \|f(t)\|_{L^2}^2.$$

Surely the choice of ϑ depends on the application. Another criteria to find a sufficient large number N_F to guarantee the desired accuracy is to state a

bound for the L^∞ truncation error. For $\vartheta \in (0, 1)$ in [Schuster et al., 2021] the authors require that the L^∞ truncation error should depend on the maximal amplitude of f , i.e. that N_F is chosen sufficiently large, s.t.

$$\|f(t) - f_{N_F}(t)\|_{L^\infty} \leq \vartheta \left[\sup_{\tau \in [0, T]} f(\tau) - \inf_{\tau \in [0, T]} f(\tau) \right].$$

The L^∞ -bound for the truncation error might cause problems due to the Gibbs phenomenon for data f that contains discontinuities.

If f has better regularity properties we can get better statements for the accuracy and also stronger convergence results:

Lemma 4.2. *Let f be at least two times continuously differentiable, i.e. let $f \in C^k(0, T)$ with $k \geq 2$. Further define*

$$\kappa_{\alpha_m^0} := |f'(T)| + \int_0^T |f''(t)| dt.$$

Then $\sum_{m=0}^{\infty} \alpha_m^0(f) \psi_m(t)$ converges absolutely to $f(t)$ and we have

$$|\alpha_m^0(f)| \leq \kappa_{\alpha_m^0} \frac{T^2}{\left(\frac{\pi}{2} + m\pi\right)^2}. \quad (4.9)$$

Proof. First we compute the Fourier coefficients $\alpha_m^0(f'')$. We have

$$\alpha_m^0(f'') = \int_0^T f''(t) \psi_m(t) dt.$$

We use partial integration two times and it follows

$$\alpha_m^0(f'') = [f'(t) \psi_m(t)]_0^T - [f(t) \psi_m'(t)]_0^T + \int_0^T f(t) \psi_m''(t) dt.$$

The derivatives of $\psi_m(t)$ are given by

$$\psi_m'(t) = \frac{\sqrt{2}}{\sqrt{T}} \cos\left(\left(\frac{\pi}{2} + m\pi\right) \frac{t}{T}\right) \left(\frac{\pi}{2} + m\pi\right) \frac{1}{T},$$

and

$$\begin{aligned} \psi_m''(t) &= -\frac{\sqrt{2}}{\sqrt{T}} \sin\left(\left(\frac{\pi}{2} + m\pi\right) \frac{t}{T}\right) \left(\frac{\pi}{2} + m\pi\right)^2 \frac{1}{T^2} \\ &= -\psi_m(t) \left(\frac{\pi}{2} + m\pi\right)^2 \frac{1}{T^2}. \end{aligned}$$

Since we have $\psi_m(0) = 0$, $f(0) = 0$ and $\psi'(T) = 0$ it follows

$$\mathfrak{a}_m^0(f'') = f'(T)\psi_m(T) - \left(\frac{\pi}{2} - m\pi\right)^2 \frac{1}{T^2} \int_0^T f(t)\psi_m(t) dt.$$

Thus we have

$$\mathfrak{a}_m^0(f'') = f'(T)\psi_m(T) - \left(\frac{\pi}{2} - m\pi\right)^2 \frac{1}{T^2} \mathfrak{a}_m^0(f),$$

which is equal to

$$\mathfrak{a}_m^0(f) = [f'(T)\psi_m(T) - \mathfrak{a}_m^0(f'')] \frac{T^2}{\left(\frac{\pi}{2} + m\pi\right)^2}.$$

With the triangle inequality it follows

$$|\mathfrak{a}_m^0(f)| \leq \left[|f'(T)\psi_m(T)| + \int_0^T |f''(t)\psi_m(t)| dt \right] \frac{T^2}{\left(\frac{\pi}{2} + m\pi\right)^2}.$$

Now we use the fact that $\psi_m(t) \leq 1$ and thus we have

$$|\mathfrak{a}_m^0(f)| \leq \left[|f'(T)| + \int_0^T |f''(t)| dt \right] \frac{T^2}{\left(\frac{\pi}{2} + m\pi\right)^2},$$

and this is equal to (4.9). Further since

$$\frac{1}{\left(\frac{\pi}{2} + m\pi\right)^2} \leq \frac{1}{m^2},$$

for all $m > 1$ and due to the absolutely convergence of $\sum_{k=1}^{\infty} \frac{1}{k^2}$, the Fourier series $\sum_{m=0}^{\infty} \mathfrak{a}_m^0(f)\psi_m(t)$ converges absolutely to $f(t)$. \square

Remark 4.3. *Lemma 4.9 provides another way to guarantee sufficient accuracy for the truncated Fourier series:*

For $\varepsilon > 0$, choose $N_F \in \mathbb{N}$, s.t.

$$|\mathfrak{a}_{N_F}^0(f)| \geq \varepsilon \quad \text{and} \quad |\mathfrak{a}_{N_F+1}^0(f)| < \varepsilon.$$

From now on for the readers' convenience we write \mathfrak{a}_m^0 instead of $\mathfrak{a}_m^0(f)$, i.e. we neglect the argument.

Consider a sequence of identically distributed random variables $(\xi_m)_{m \geq 0}$ on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. These random variables neither need to be independent nor they need to have the same probability distribution function. We later choose independent and identically Gaussian distributed random variables $\xi_m \sim \mathcal{N}(\mathbb{1}_n, \Sigma)$ with mean value $\mathbb{1}_n$ and diagonal covariance matrix $\Sigma \in \mathbb{R}_+^{n \times n}$. Mention that Σ contains the variances σ_i^2 on its diagonal. For $\omega \in \Omega$ consider the random function

$$f^\omega(t) = \sum_{m=0}^{\infty} \xi_m(\omega) \alpha_m^0 \psi_m(t).$$

Then according to an extension of the Paley-Zygmund Theorem (see [Hill, 2012] and [Cuzick and Lai, 1980]) it follows that if $f \in L^2([0, T])$, then f^ω converges almost surely in L^2 .

We shortly outline how to use this approach for functions that do not satisfy $f(0) = 0$. Let a function $f \in L^2([0, T])$ with $f(0) \neq 0$ be given. We define $\tilde{f}(t) := f(t) - f(0)$, thus $\tilde{f}(0) = 0$. For $m = 0, 1, 2, \dots$, we define the constants α_m^0 for \tilde{f} and get a representation of \tilde{f} as (4.7). From this we get the Fourier series representation

$$f(t) = \left(\sum_{m=0}^{\infty} \alpha_m^0 \psi_m(t) \right) + f(0),$$

and the random Fourier series representation

$$f^\omega(t) = \left(\sum_{m=0}^{\infty} \xi_m(\omega) \alpha_m^0 \psi_m(t) \right) + f(0).$$

In a last step in this section we show how to choose the diagonal covariance matrix Σ s.t. the truncated random Fourier series $f_{N_F}^\omega$ almost surely satisfies a lower bound, i.e.

$$f_{N_F}^\omega \geq \underline{y} \quad \text{almost surely.} \quad (4.10)$$

We assume w.l.o.g. that f is one-dimensional. For $m = 0, \dots, N_F$ consider $\xi_m \sim \mathcal{N}(1, \sigma^2)$ with mean value 1 and variance $\sigma^2 \geq 0$. Since we have a stochastic setting here, surely (4.10) cannot be guaranteed by 100% but we can give an upper bound for the variance σ^2 to almost guarantee that (4.10) is satisfied.

The idea for finding a bound for the variance s.t. (4.10) is satisfied almost surely comes from the classical robust optimization: We first choose an appropriate uncertainty set and then we derive bounds for σ^2 s.t even the worst case within the uncertainty set satisfies (4.10). Since the random variables ξ_n are Gaussian distributed, the confidence interval $[\mu - 3\sigma, \mu + 3\sigma]$ is a good choice for the uncertainty set due to the fact that

$$\int_{\mu-3\sigma}^{\mu+3\sigma} \rho_{\xi}(z) dz = 0.9973 \dots,$$

so 99.73% out of 10000 random Gaussian distributed numbers are inside $[\mu - 3\sigma, \mu + 3\sigma]$. Next we split the Fourier series in positive and negative terms. Let

$$\gamma^* := \operatorname{argmin}_{t \in [0, T]} f(t),$$

be the argument of the minimal value of f and let

$$\begin{aligned} \mathfrak{J}_+ &:= \{m = 0, 1, \dots \mid \mathbf{a}_m^0 \psi_m(\gamma^*) \geq 0\} \\ \text{and } \mathfrak{J}_- &:= \{m = 0, 1, \dots \mid \mathbf{a}_m^0 \psi_m(\gamma^*) \leq 0\} \end{aligned}$$

denote the index sets of positive and negative terms. Then we have

$$f(\gamma^*) = \sum_{m \in \mathfrak{J}_+} \mathbf{a}_m^0 \psi_m(\gamma^*) + \sum_{m \in \mathfrak{J}_-} \mathbf{a}_m^0 \psi_m(\gamma^*).$$

Even the worst case of random numbers $\xi_m(\omega) \in [\mu - 3\sigma, \mu + 3\sigma]$ should satisfy (4.10). The worst case is if all positive terms get smaller and all negative terms (in absolute values) get larger, i.e.

$$\sum_{m \in \mathfrak{J}_+} (\mu - 3\sigma) \mathbf{a}_m^0 \psi_m(\gamma^*) + \sum_{m \in \mathfrak{J}_-} (\mu + 3\sigma) \mathbf{a}_m^0 \psi_m(\gamma^*) \geq \underline{y}.$$

This is equivalent to

$$\mu f(\gamma^*) - 3\sigma \left[\sum_{m \in \mathfrak{J}_+} \mathbf{a}_m^0 \psi_m(\gamma^*) - \sum_{m \in \mathfrak{J}_-} \mathbf{a}_m^0 \psi_m(\gamma^*) \right] \geq \underline{y},$$

and it directly follows

$$\sigma \leq \frac{\mu f(\gamma^*) - \underline{y}}{3 \left[\sum_{m \in \mathfrak{J}_+} \mathbf{a}_m^0 \psi_m(\gamma^*) - \sum_{m \in \mathfrak{J}_-} \mathbf{a}_m^0 \psi_m(\gamma^*) \right]}. \quad (4.11)$$

Remark 4.4. *The computation of (4.11) is not expensive in a numerical sense since all terms*

$$\mathbf{a}_m^0 \psi_m(t),$$

have to be computed anyway for all $m = 0, \dots, N_F$.

Remark 4.5. If f is not one-dimensional, due to the diagonality of Σ the bound for the variance stated in (4.11) can be found for every dimension of f . Analogously bounds for Σ can be found s.t.

$$f_{N_F}^\omega \leq \bar{y},$$

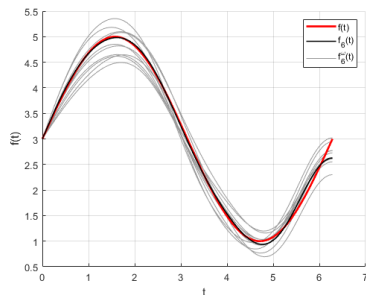
almost surely. We mention that the bounds for Σ can equivalently be computed for different uncertainty sets $[\mu - x\sigma, \mu + x\sigma]$ with $x \in \mathbb{R}_+$.

Example. For $t \in [0, 2\pi]$ consider the functions

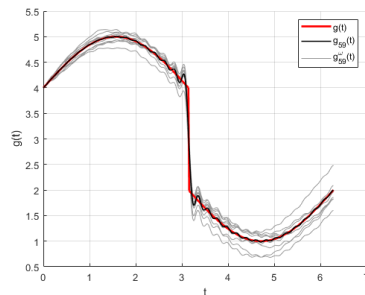
$$f(t) = 2\sin(t) + 3 \quad \text{and} \quad g(t) = \begin{cases} \sin(t) + 4 & t \in [0, \pi], \\ \sin(t) + 2 & t \in (\pi, 2\pi]. \end{cases} \quad (4.12)$$

For the Fourier representation of f and g we choose $\vartheta = 0.001$ (cf. Remark 4.1). This leads to a Fourier series representation of f that is truncated after 7 terms (i.e. $N_F = 6$) and a Fourier series representation of g that is truncated after 60 terms (i.e. $N_F = 59$). For the L^∞ -error we choose $\vartheta = 0.05$, which leads to a Fourier series representation of f that is truncated after 14 terms (i.e. $N_F = 13$). Due to the discontinuity in g the L^∞ -error is always larger than or equal to 1, so the L^∞ truncation error can never be satisfied for g with $\vartheta = 0.05$.

We define random variables $(\xi_f)_m \sim \mathcal{N}(1, \sigma_f^2)$ and $(\xi_g)_m \sim \mathcal{N}(1, \sigma_g^2)$. We want the random data $f_6^\omega(t)$ and $g_{59}^\omega(t)$ to be non negative almost surely. Then (4.11) leads to $\sigma_f \leq 0.1580$ and $\sigma_g \leq 0.0703$. In Figure 4.1 we plotted respectively 10 random scenarios for the functions $f_6(t)$ and $g_{59}(t)$.



(a) Random scenarios for the function $f_6(t)$



(b) Random scenarios for the function $g_{59}(t)$

Figure 4.1.: Random scenarios for the functions f and g defined in (4.12)

So we have shown how to create random scenarios resp. random time dependent functions with no loss in regularity. We use this procedure later to simulate time dependent random gas demand.

4.3. Mathematical Modelling and Problem Description

In this section we introduce a mathematical model for the dynamic gas transport in a pipeline network. The setting is similar to the stationary case (see *Section 3*). Further we write the model in terms of Riemann invariants and discuss the existence of a solution.

4.3.1. Model Derivation

Consider a connected, directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the vertex set $\mathcal{V} = \{v_0, \dots, v_n\}$ and a set of edges $\mathcal{E} = \{e_1, \dots, e_m\} \subseteq \mathcal{V} \times \mathcal{V}$ ($n, m \in \mathbb{N}$). Contrary to the stationary case we do not focus on tree-structured graphs in this section. Also the numbering of the nodes and edges is not important here. Every edge $e \in \mathcal{E}$ represents a pipe with (positive) length L^e . For $(t, x) \in [0, T] \times [0, L^e]$ we consider the isothermal Euler equations for horizontal pipes and ideal gases

$$\left\{ \begin{array}{l} \rho_t^e + q_x^e = 0, \\ q_t^e + \left((c^e)^2 \rho^e + \frac{(q^e)^2}{\rho^e} \right)_x = -\frac{\lambda_F^e}{2D^e} \frac{q^e |q^e|}{\rho^e} \end{array} \right. \quad (\text{qISO})$$

with density ρ^e and flow q^e . The constants $c^e, \lambda_F^e, D^e \in \mathbb{R}_{\geq 0}$ denote the speed of sound in the gas, the pipe friction coefficient and the pipe diameter of the edge $e \in \mathcal{E}$ respectively. For simplicity we assume that these parameters are constant on every edge. Further ρ^e and q^e represent the restriction of the pressure and the flow (defined over the network) to a single edge $e \in \mathcal{E}$.

We state that all nodes $v \in \mathcal{V}$ with $|\mathcal{E}_0(v)| > 1$ are inner nodes and that all nodes $v \in \mathcal{V}$ with $|\mathcal{E}_0(v)| = 1$ are terminal nodes. We assume w.l.o.g. that gas inflow and gas outflow only occurs at terminal nodes. If this is not the case, one can add a frictionless pipe and an additional node where the gas inflow or gas outflow occurs, as it is shown in *Figure 4.2*.

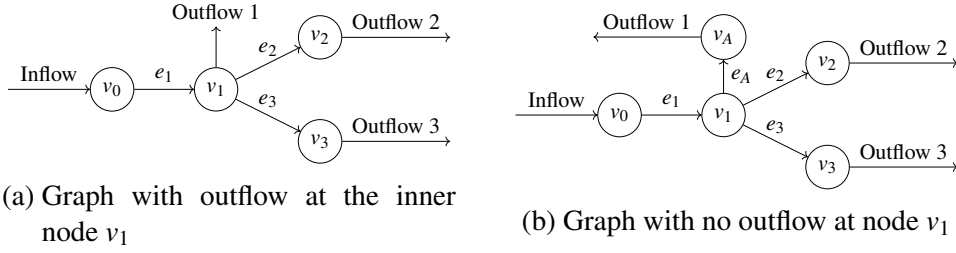


Figure 4.2.: Example on how to add an additional node v_A and an additional frictionless pipe e_A s.t. gas outflow does not occur at an inner node

Next we define coupling conditions for all inner nodes of the graph. Therefor for $v \in \mathcal{V}$ and $e \in \mathcal{E}_0(v)$ we define

$$x^e(v) := \begin{cases} 0 & \text{if } e \in \mathcal{E}_+(v), \\ L^e & \text{if } e \in \mathcal{E}_-(v), \end{cases}$$

as the end of edge e , that corresponds to node v . Then similar to the stationary case in *Chapter 3* we assume conservation of mass for all inner nodes, i.e. for all nodes $v \in \mathcal{V}$ with $|\mathcal{E}_0(v)| > 1$ we have

$$\sum_{e \in \mathcal{E}_0(v)} (D^e)^2 q^e(t, x^e(v)) = 0 \quad \forall t \in [0, T]. \quad (4.13)$$

Further we assume continuity in density for all inner nodes, i.e. for all nodes $v \in \mathcal{V}$ with $|\mathcal{E}_0(v)| > 1$ we have

$$\rho^e(t, x^e(v)) = \rho^f(t, x^f(v)) \quad \forall e, f \in \mathcal{E}_0(v). \quad (4.14)$$

Since we assume ideal gas, the continuity in density conforms with the continuity in pressure. Next we define boundary conditions for terminal nodes, i.e. for all nodes $v \in \mathcal{V}$ with $|\mathcal{E}_0(v)| = 1$ we define

$$\rho^e(t, 0) := \rho_0(t) \quad \text{if } e \in \mathcal{E}_+(v), \quad (4.15)$$

and

$$q^e(t, L^e) := b(t) \quad \text{if } e \in \mathcal{E}_-(v). \quad (4.16)$$

As in the stationary case the function $\rho(\cdot)$ states the (time dependent) inlet density and $b(\cdot)$ states the (time dependent) gas outflow.

Last we define initial conditions for the density and for the flow, i.e. for all $e \in \mathcal{E}$ we define

$$\rho^e(0, x) := \rho_{\text{ini}}^e(x), \quad (4.17)$$

and

$$q^e(0, x) = q_{\text{ini}}^e(x). \quad (4.18)$$

So the full model is given by the isothermal Euler equations (qISO), coupling conditions (4.13) and (4.14), boundary conditions (4.15) and (4.16) and initial conditions (4.17) and (4.18):

$$\left\{ \begin{array}{l} \text{For all } e \in \mathcal{E} \text{ we have} \\ \rho_t^e + q_x^e = 0, \\ q_t^e + \left((c^e)^2 \rho^e + \frac{(q^e)^2}{\rho^e} \right)_x = -\frac{\lambda_F^e}{2D^e} \frac{q^e |q^e|}{\rho^e}, \\ \rho^e(0, x) = \rho_{\text{ini}}^e(x), \\ q^e(0, x) = q_{\text{ini}}^e(x), \\ \text{for all } v \in \mathcal{V} \text{ with } |\mathcal{E}_0(v)| > 1 \text{ we have} \\ \sum_{e \in \mathcal{E}_0(v)} (D^e)^2 q^e(t, x^e(v)) = 0, \\ \rho^e(t, x^e(v)) = \rho^f(t, x^f(v)) \quad e, f \in \mathcal{E}_0(v), \\ \text{for all } v \in \mathcal{V} \text{ with } |\mathcal{E}_0(v)| = 1 \text{ we have} \\ \rho^e(t, 0) = \rho_0^e(t) \quad e \in \mathcal{E}_+(v), \\ q^e(t, L^e) = b^e(t) \quad e \in \mathcal{E}_-(v). \end{array} \right. \quad (\text{dynModel})$$

4.3.2. The Model in Terms of Riemann Invariants

In this section we write the model (dynModel) in terms of Riemann invariants and discuss the existence of a solution. We already discussed Riemann invariants in *Section 2.2*. The Riemann invariants for the quasilinear isothermal Euler equations (qISO) are stated in *Lemma 2.1*. We have

$$R_1^e(\rho^e, q^e) = \ln(\rho^e) + \frac{q^e}{c^e \rho^e}, \quad (4.19)$$

and

$$R_2^e(\rho^e, q^e) = \ln(\rho^e) - \frac{q^e}{c^e \rho^e}. \quad (4.20)$$

Thus the quasilinear isothermal Euler equations (qISO) in terms of Riemann invariants are given by

$$\begin{bmatrix} R_1^e \\ R_2^e \end{bmatrix}_t + \begin{bmatrix} \lambda_1^e & 0 \\ 0 & \lambda_2^e \end{bmatrix} \begin{bmatrix} R_1^e \\ R_2^e \end{bmatrix}_x = -\frac{c^e \lambda_F^e}{8D^e} (R_1^e - R_2^e) |R_1^e - R_2^e| \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad (\text{qISOR})$$

where the eigenvalues λ_1^e and λ_2^e are stated in (2.7). For the initial conditions we have

$$\begin{aligned} R_1^e(0, x) &= \ln(\rho^e(0, x)) + \frac{q^e(0, x)}{c^e \rho^e(0, x)} \\ &= \ln(\rho_{\text{ini}}^e(x)) + \frac{q_{\text{ini}}^e(x)}{c^e \rho_{\text{ini}}^e(x)}, \end{aligned} \quad (4.21)$$

and

$$\begin{aligned} R_2^e(0, x) &= \ln(\rho^e(0, x)) - \frac{q^e(0, x)}{c^e \rho^e(0, x)} \\ &= \ln(\rho_{\text{ini}}^e(x)) - \frac{q_{\text{ini}}^e(x)}{c^e \rho_{\text{ini}}^e(x)}. \end{aligned} \quad (4.22)$$

Next we state the coupling conditions in terms of Riemann invariants.

Lemma 4.6. *For all inner nodes $v \in \mathcal{V}$ (i.e. all nodes $v \in \mathcal{V}$ with $|\mathcal{E}_0(v)| > 1$) and $e \in \mathcal{E}_0(v)$ let the linear equation*

$$R_1^e(t, x^e(v)) = -R_2^e(t, x^e(v)) + \frac{2}{\sum_{f \in \mathcal{E}_0(v)} (D^f)^2} \sum_{g \in \mathcal{E}_0(v)} (D^g)^2 R_2^g(t, x^g(v)), \quad (4.23)$$

hold. Then the coupling conditions (4.13) and (4.14) hold.

Proof. For $v \in \mathcal{V}$ with $|\mathcal{E}_0(v)| > 1$ and for $e \in \mathcal{E}_0(v)$ the equation (4.23) is equivalent to

$$R_1^e(t, x^e(v)) + R_2^e(t, x^e(v)) = \frac{2}{\sum_{f \in \mathcal{E}_0(v)} (D^f)^2} \sum_{g \in \mathcal{E}_0(v)} (D^g)^2 R_2^g(t, x^g(v)),$$

which implies that

$$R_1^e(t, x^e(v)) + R_2^e(t, x^e(v)) = R_1^f(t, x^f(v)) + R_2^f(t, x^f(v)), \quad (4.24)$$

for all $e, f \in \mathcal{E}_0(v)$. Thus (4.14) holds. Further (4.23) is equivalent to

$$\sum_{f \in \mathcal{E}_0(v)} (D^f)^2 (R_1^e(t, x^e(v)) + R_2^e(t, x^e(v))) = 2 \sum_{g \in \mathcal{E}_0(v)} (D^g)^2 R_2^g(t, x^g(v)). \quad (4.25)$$

Due to (4.24), we can equivalently write (4.25) as

$$\sum_{f \in \mathcal{E}_0^f(v)} (D^f)^2 \left(R_1^f(t, x^f(v)) + R_2^f(t, x^f(v)) \right) = 2 \sum_{g \in \mathcal{E}_0^g(v)} (D^g)^2 R_2^g(t, x^g(v)),$$

which is equivalent to

$$\sum_{f \in \mathcal{E}_0^f(v)} (D^f)^2 R_1^f(t, x^f(v)) = \sum_{g \in \mathcal{E}_0^g(v)} (D^g)^2 R_2^g(t, x^g(v)).$$

This implies

$$\sum_{f \in \mathcal{E}_0^f(v)} (D^f)^2 \left(R_1^f(t, x^f(v)) - R_2^f(t, x^f(v)) \right) = 0,$$

and thus, the conservation of mass (4.13) holds. \square

Last we write the boundary conditions defined at the terminal nodes in terms of Riemann invariants, i.e. at the nodes $v \in \mathcal{V}$ with $|\mathcal{E}_0(v)| = 1$. If $e \in \mathcal{E}_+(v)$ then from (4.19), (4.20) and (4.15) it follows

$$\rho_0^e(t) = \exp \left(\frac{R_1^e(t, 0) + R_2^e(t, 0)}{2} \right),$$

which is equivalent to

$$R_1^e(t, 0) = 2 \ln(\rho_0^e(t)) - R_2^e(t, 0). \quad (4.26)$$

Further if $e \in \mathcal{E}_-(v)$ then from (4.19), (4.20) and (4.16) it follows

$$b^e(t) = c \frac{R_1^e(t, L^e) - R_2^e(t, L^e)}{2} \exp \left(\frac{R_1^e(t, L^e) + R_2^e(t, L^e)}{2} \right),$$

which is equivalent to

$$\frac{R_2^e(t, L^e) - R_1^e(t, L^e)}{2} \exp \left(\frac{R_2^e(t, L^e) - R_1^e(t, L^e)}{2} \right) = -\frac{1}{c} b^e(t) \exp(-R_1^e(t, L^e)).$$

With the *Lambert W function* it follows

$$R_2^e(t, L^e) = 2 W \left(-\frac{1}{c} b^e(t) \exp(-R_1^e(t, L^e)) \right) + R_1^e(t, L^e). \quad (4.27)$$

The Lambert W function is also known as omega function or product logarithm (see e.g. [Lambert, 1758], [Corless et al., 1996] and [Chapeau-Blondeau and Monir, 2002]) and it is defined as the inverse of the map $\omega \mapsto \omega \exp(\omega)$, i.e. we have $\omega = W(z)$ if and only if $\omega \exp(\omega) = z$.

Since we assume positive boundary flow $b^e(t)$ the argument of the Lambert W function is negative, which might cause problems because the Lambert W function is defined for arguments larger than or equal to $(-1/e)$. But for positive flow $q^e(t, x)$ the Riemann invariant $R_1^e(t, x)$ is quite large since the pressures of the gas usually take values around 50bar, which is $5 \cdot 10^6$ Pa. So the argument of the Lambert W function is negative but usually larger than $(-1/e)$, even close to zero.

The full model in terms of Riemann invariants is given by the hyperbolic balance law (qISOR), coupling conditions (4.23), initial conditions (4.21) and (4.22) and boundary conditions (4.26) and (4.27):

$$\left\{ \begin{array}{l}
 \text{For all } e \in \mathcal{E} \text{ we have} \\
 \\
 (R_1^e)_t + \lambda_1^e (R_1^e)_x = \frac{c^e \lambda_F^e}{8D^e} (R_1^e - R_2^e) |R_1^e - R_2^e| \\
 (R_2^e)_t + \lambda_2^e (R_2^e)_x = -\frac{c^e \lambda_F^e}{8D^e} (R_1^e - R_2^e) |R_1^e - R_2^e| \\
 \\
 R_1^e(0, x) = \ln(\rho_{\text{ini}}^e(x)) + \frac{q_{\text{ini}}^e(x)}{c^e \rho_{\text{ini}}^e(x)}, \\
 R_2^e(0, x) = \ln(\rho_{\text{ini}}^e(x)) - \frac{q_{\text{ini}}^e(x)}{c^e \rho_{\text{ini}}^e(x)}, \\
 \\
 \text{for all } v \in \mathcal{V} \text{ with } |\mathcal{E}_0(v)| > 1 \text{ we have} \\
 \\
 R_1^e(t, x^e(v)) = -R_2^e(t, x^e(v)) \\
 \quad + \frac{2}{\sum_{f \in \mathcal{E}_0(v)} (D^f)^2} \sum_{g \in \mathcal{E}_0(v)} (D^g)^2 R_2^g(t, x^g(v)) \quad e \in \mathcal{E}_0^-(v), \\
 \\
 \text{for all } v \in \mathcal{V} \text{ with } |\mathcal{E}_0(v)| = 1 \text{ we have} \\
 \\
 R_1^e(t, 0) = 2 \ln(\rho_0(t)) - R_2^e(t, 0) \quad e \in \mathcal{E}_+(v), \\
 \\
 R_2^e(t, L^e) = 2W\left(-\frac{1}{c} b^e(t) \exp(-R_1^e(t, L^e))\right) \\
 \quad + R_1^e(t, L^e) \quad e \in \mathcal{E}_-(v).
 \end{array} \right. \quad (\text{dynModelR})$$

Remark 4.7. *While we have classical Dirichlet boundary conditions in (dyn-Model), the boundary conditions in (dynModelR) are a mix of Dirichlet boundary conditions and a feedback law. This does not affect the existence of a solution.*

In [Gugat and Ulbrich, 2018]) the authors state a well-posedness result for general quasilinear hyperbolic 2×2 systems. Our system differs from their model in the boundary conditions and in the fact that we assume ideal gas while they assume real gas, but their proof works analogously for our system. Their statement is that for Lipschitz continuous data with sufficiently small Lipschitz constants, a solution in the sense of characteristic curves exist on a sufficiently small time interval.

To be more precise on that, we need to assume that the eigenvalues $\lambda_1^e(R_1^e, R_2^e)$ and $\lambda_2^e(R_1^e, R_2^e)$ are Lipschitz continuous, the source term of (qISOR) is Lipschitz continuous with respect to x on $[0, T] \times [0, L^e]$, the initial and the boundary data are Lipschitz continuous and the C^0 -compatibility conditions are satisfied between the initial and the boundary conditions. Then (dynModelR) has a unique solution on $[0, T]$ (where T depends on the length of the pipes and on the Lipschitz constants of the eigenvalues), that satisfies the corresponding integral equation along the characteristic curves. Further the unique solution is Lipschitz continuous with respect to x . More details and the full proof can be found in [Gugat and Ulbrich, 2018, Section 5].

For $D \subseteq \mathbb{R}$ let

$$\text{Lip}(D) := \left\{ f \in C(D) \mid \sup_{x,y \in D, x \neq y} \frac{|f(x) - f(y)|}{|x - y|} < \infty \right\}$$

be the space of Lipschitz continuous functions on $D \subseteq \mathbb{R}$. Then for appropriate eigenvalues, initial data and boundary data, we can find a solution (R_1^e, R_2^e) on every edge $e \in \mathcal{E}$ with

$$R_1^e, R_2^e \in C([0, T], \text{Lip}((0, L^e); \mathbb{R})). \quad (4.28)$$

With this we can define the feasible set for the dynamic case.

4.3.3. Problem Description

In this section we want the density to satisfy box constraints at all nodes where gas leaves the network. For ideal gases, density and pressure are directly proportional, so the bounds for the density in this chapter are equivalent to the bounds for the pressure in the stationary setting of *Chapter 3*. Let

$$\mathcal{V}_{\text{out}} := \{v \in \mathcal{V} \mid \mathcal{E}_0(v) = \mathcal{E}_-(v)\},$$

be the set of all nodes, where gas leaves the network, i.e. where the boundary conditions (4.16) hold. Let $q^v(t)$ be the flow of the gas at node $v \in \mathcal{V}$ and let $\rho^v(t)$ be the density of the gas at node $v \in \mathcal{V}$. Let bounds for the density $\rho^{\min}, \rho^{\max} \in \mathbb{R}^{|\mathcal{V}_{\text{out}}|}$ with $0 < \rho^{\min} \leq \rho^{\max}$ (component-by-component) be given. For $t^* \in [0, T]$, we demand

$$\rho^v(t^*) \in [\rho^{v,\min}, \rho^{v,\max}] \quad \forall v \in \mathcal{V}_{\text{out}},$$

where $\rho^{v,\min}$ and $\rho^{v,\max}$ are the bounds for the density at node $v \in \mathcal{V}_{\text{out}}$. We define the set of feasible loads for the dynamic setting as follows:

Definition 4.8. For density bounds $\rho^{\min}, \rho^{\max} \in \mathbb{R}^{|\mathcal{V}_{\text{out}}|}$ with $0 < \rho^{v,\min} \leq \rho^{v,\max}$ ($v \in \mathcal{V}_{\text{out}}$) and for a point in time $t^* \in [0, T]$, the set

$$M(t^*) := \left\{ \begin{array}{l} b : [0, T] \rightarrow \mathbb{R}^{|\mathcal{V}_{\text{out}}|} \\ b_i \in \text{Lip}([0, T]) \end{array} \left| \begin{array}{l} \text{The solution } \rho^v(t), q^v(t) \text{ of } (\text{dynModel}) \\ \text{satisfies the box constraints} \\ \rho^v(t^*) \in [\rho^{v,\min}, \rho^{v,\max}] \text{ at time } t^* \end{array} \right. \right\} \quad (4.29)$$

is called the set of feasible loads.

Surely the densities and the flows depend on the Lipschitz continuous boundary functions $b(t)$ but we neglect this dependence here for better readability.

As in the stationary setting of *Chapter 3*, the boundary functions $b(t)$ represent the loads of the network, i.e. the consumers' gas demand. Due to the liberalization of the gas market, from a gas network owners' point of view the gas demand cannot be known a priori, so it can be seen as random. We consider this randomness in the sense of Fourier series as it is described in *Section 4.2*. Our aim in this section is to answer the following question:

Q_{prob}: For a given inlet pressure, can we guarantee, that every consumer receives their demanded gas, s.t. the gas pressure in the network is never neither too high nor too low, in at least $\alpha\%$ of all scenarios?

Since we are in a dynamic setting in this chapter, the scenarios depend on the time. We want to compute the probability that a random boundary function is feasible, i.e.

$$\mathbb{P}(b \in M(t) \forall t \in [0, T]). \quad (4.30)$$

Usually the aim is to choose box constraints for the density such that the probability is large, e.g. close to one. This can be interpreted as robustness (in a probabilistic sense) against perturbations in the outflow.

In the next section we apply the KDE approach (introduced in *Section 1.1*) to the dynamic setting. The aim is to compute the probability (4.30). As in the stationary setting, the load $b(t)$ is considered to be random but we have no information about the structure of the set of feasible loads $M(t)$. The densities $\rho^v(t)$ at the nodes $v \in \mathcal{V}_{\text{out}}$ are also random but the set of feasible densities

$$\mathcal{P}_{\min}^{\max} := \otimes_{v \in \mathcal{V}_{\text{out}}} \left[\rho^{v, \min}, \rho^{v, \max} \right], \quad (4.31)$$

is a well-known cuboid. For $t \in [0, T]$ let $\rho_{\text{out}}(t)$ be the densities at the nodes $v \in \mathcal{V}_{\text{out}}$. Then we have

$$\mathbb{P}(b \in M(t) \forall t \in [0, T]) = \mathbb{P}(\rho_{\text{out}}(t) \in \mathcal{P}_{\min}^{\max} \forall t \in [0, T]). \quad (4.32)$$

So we use the KDE approach to estimate the probability density function of the gas densities at the outflow nodes and integrate this probability density over the set of feasible pressures to compute the desired probability. Since an explicit solution of (qISO) is not given here, we do not use the SRD approach in the dynamic setting (cf. *Section 3.7*).

4.4. Application of the KDE

In this section we use the KDE (introduced in *Section 1.1*) to compute the probability for a random load function to be feasible, where the feasible set is defined in (4.29).

Let $n_{\text{out}} := |\mathcal{V}_{\text{out}}|$ be the number of nodes where gas leaves the network and let a Lipschitz continuous boundary function $b : [0, T] \rightarrow \mathbb{R}^{n_{\text{out}}}$ be given. Consider a sequence of n_{out} -dimensional Gaussian distributed random variables $(\xi_m)_{m \geq 0}$ on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where

$$\xi_m \sim \mathcal{N}(\mu, \Sigma).$$

Here $\mu \in \mathbb{R}^{n_{\text{out}}}$ is the mean value and $\Sigma \in \mathbb{R}^{n_{\text{out}} \times n_{\text{out}}}$ is the positive definite covariance matrix. For $N_{\text{KDE}} \in \mathbb{N}$ and $m \geq 0$, let

$$\mathcal{A}_m := \{a_m^{\omega,1}, \dots, a_m^{\omega, N_{\text{KDE}}}\} \subseteq \mathbb{R}^{n_{\text{out}}},$$

be an independent and identically distributed sampling of the random variable ξ_m . Using the idea of random Fourier series (introduced in *Section 4.2*) let

$$\mathcal{B}_{\mathcal{A}} := \{b^{\omega,1}(t), \dots, b^{\omega, N_{\text{KDE}}}(t)\},$$

with

$$b^{\omega,i}(t) = \sum_{m=0}^{\infty} a_m^{\omega,i} \alpha_m^0 \psi_m(t) \quad (i = 1, \dots, N_{\text{KDE}}),$$

be the corresponding sampling of random boundary functions. Note that since b is Lipschitz continuous, the $b^{\omega,i}$ are also Lipschitz continuous almost surely. We assume that $b^{\omega,i}(t)$ is non negative almost surely, i.e.

$$\mathbb{P}(b^{\omega,i}(t) \geq 0 \forall t \in [0, T]) \approx 1 \quad (i = 1, \dots, n_{\text{out}}).$$

Otherwise problems can occur as it is described in *Remark 3.9*. The non negativity of $b^{\omega,i}(t)$ can be guaranteed almost surely by an appropriate choice of μ and Σ as it is described in *Section 4.2* (see (4.11)).

Let Lipschitz continuous initial data ρ_{ini}^e , q_{ini}^e be given on every edge $e \in \mathcal{E}$ and let inlet pressures ρ_0^e be given on every node, where gas enters the network. Then for every $b^{\omega,i} \in \mathcal{B}_{\mathcal{A}}$ there exists a solution of (dynModel) that is continuous in time and Lipschitz continuous in space (this has been discussed in *Section 4.3.2*). Let

$$\mathcal{P}_{\mathcal{B}} := \{\rho_{\text{out}}(t, b^{\omega,1}), \dots, \rho_{\text{out}}(t, b^{\omega, N_{\text{KDE}}})\},$$

be the corresponding sampling of gas densities at the nodes where gas leaves the network, i.e. the $\rho^v(t, b^{\omega,i})$ is contained in $\rho_{\text{out}}(t, b^{\omega,i})$ iff $v \in \mathcal{V}_{\text{out}}$. We explicitly state that the gas density depends on the gas outflow here since this is important for following analysis.

In contrast to the stationary case of *Chapter 3* we cannot directly approximate the probability density function of the gas density due to the time dependence. Roughly speaking we need to get rid of the term „ $\forall t \in [0, T]$ “. This term is the robust part of the probust constraint

$$\mathbb{P}(\rho_{\text{out}}(t) \in \mathcal{P}_{\text{min}}^{\text{max}} \forall t \in [0, T]) \quad (i = 1, \dots, N_{\text{KDE}}),$$

so we use an idea of the classical robust optimization: the worst case analysis. This means that the density bounds are satisfied for all times, iff the minimal and the maximal density in $[0, T]$ satisfy the density bounds.

So for $i = 1, \dots, N_{\text{KDE}}$ we define

$$\underline{\rho}_{\text{out}}(b^{\omega,i}) := \min_{t \in [0, T]} \rho_{\text{out}}(t, b^{\omega,i}),$$

and

$$\bar{\rho}_{\text{out}}(b^{\omega,i}) := \max_{t \in [0, T]} \rho_{\text{out}}(t, b^{\omega,i}),$$

as minimal and maximal gas densities of ρ_{out} on $[0, T]$. The minimal and the maximal gas densities have to be understood component-by-component and they exist due to the continuity of ρ^v and the compactness of $[0, T]$. For $i = 1, \dots, N_{\text{KDE}}$ it follows

$$\begin{aligned} \rho_{\text{out}}(t, b^{\omega,i}) &\in \mathcal{D}_{\min}^{\max} \quad \forall t \in [0, T] \\ &\Updownarrow \\ \underline{\rho}_{\text{out}}(b^{\omega,i}), \bar{\rho}_{\text{out}}(b^{\omega,i}) &\in \mathcal{D}_{\min}^{\max}. \end{aligned} \tag{4.33}$$

We can now approximate the probability density function of the minimal and maximal gas densities, which leads to a $(2n_{\text{out}})$ -dimensional (time independent) probability density function $\varrho_{\rho, \text{KDE}}$. Since there is no time dependency anymore we can use the theory from the stationary case in *Section 3.3*.

Remark 4.9. *In hyperbolic systems, the information from the boundary needs time to travel through the system. We need to guarantee that the solution takes its minimal resp. maximal values in the time period when the solution depends on the uncertainty. We illustrate this with an easy example:*

Consider a linear transport equation without source term, with positive parameter c and with initial and boundary condition r_0 and r_B on an interval $[0, T] \times [0, L]$

$$\begin{aligned} r_t(t, x) + dr_x(t, x) &= 0, \\ r(0, x) &= 0, \\ r(t, 0) &= r_B(t). \end{aligned}$$

Then its solution is given by

$$r(t, x) = \begin{cases} r_0(x - dt) & \text{if } t < \frac{x}{d}, \\ r_B\left(t - \frac{x}{d}\right) & \text{if } t \geq \frac{x}{d}. \end{cases}$$

If the boundary condition is positive for all $t > 0$, then the minimal value of the solution is in the interval $[0, L/d]$. That means that even if r_b is random, zero is an upper bound for the minimal value of every scenario, a scenario with minimal value larger than zero cannot exist. If this is the case, the probability distribution function for the minimal values contains a discontinuity and a probability density function in the classical sense might not exist so the convergence theory stated in Section 1.1.3 does not hold.

So either one has to guarantee that a minimal and a maximal value exist outside of $[0, L/d]$ (e.g. by choosing $r_B(t) = \sin(t)$ with T sufficiently large) or one consider bounds on the solution only for a time interval $[t_1, T]$ with $t_1 > L/d$.

We assume from now on that initial, boundary and coupling conditions satisfy the property that the minimal and maximal values of the solution of (dyn-Model) depend on the uncertainty.

For Gaussian kernel functions and for diagonal bandwidth matrices H^{\min} of the sampling of minimal densities $\underline{\rho}_{\text{out}}(b^{\omega,i})$ with $(h_j^{\min})^2 := H_{j,j}^{\min}$ and H^{\max} of the sampling of maximal densities $\bar{\rho}_{\text{out}}(b^{\omega,i})$ with $(h_j^{\max})^2 := H_{j,j}^{\max}$ we have

$$\begin{aligned} \varrho_{\rho, N_{\text{KDE}}}(z) &= \frac{1}{N_{\text{KDE}} \prod_{j=1}^{n_{\text{out}}} h_j^{\min} \prod_{j=1}^{n_{\text{out}}} h_j^{\max}} \\ &\quad \cdot \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^{n_{\text{out}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_{1,j} - \underline{\rho}_{\text{out},j}(b^{\omega,i})}{h_j^{\min}}\right)^2\right) \\ &\quad \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_{2,j} - \bar{\rho}_{\text{out},j}(b^{\omega,i})}{h_j^{\max}}\right)^2\right), \end{aligned} \tag{4.34}$$

where $z \in \mathbb{R}^{2n_{\text{out}}}$, $z = [z_1, z_2]^\top$ with $z_1, z_2 \in \mathbb{R}^{n_{\text{out}}}$ (cf. (3.15)). To be more precise we define a sampling for the minimal and maximal gas densities

$$\bar{\mathcal{P}}_{\mathcal{B}} := \left\{ \begin{bmatrix} \underline{\rho}_{\text{out}}(b^{\omega,1}) \\ \bar{\rho}_{\text{out}}(b^{\omega,1}) \end{bmatrix}, \dots, \begin{bmatrix} \underline{\rho}_{\text{out}}(b^{\omega, N_{\text{KDE}}}) \\ \bar{\rho}_{\text{out}}(b^{\omega, N_{\text{KDE}}}) \end{bmatrix} \right\} \subseteq \mathbb{R}^{2n_{\text{out}}}.$$

Let Σ_{\min}^{\max} be the covariance matrix of the sampling $\bar{\mathcal{P}}_{\mathcal{B}}$. We use the KDE introduced in (1.6) with Gaussian kernel and diagonal bandwidth matrix $H_{i,i} = h^2(\Sigma_{\min}^{\max})_{i,i}$ (h defined in (1.7)) to get an approximation of the probability density function of the minimal and maximal gas densities.

Then we have

$$\begin{aligned} \varrho_{\rho, N_{\text{KDE}}}(z) &= \\ &= \frac{1}{N_{\text{KDE}} \det(H)^{\frac{1}{2}}} \sum_{i=1}^{N_{\text{KDE}}} \frac{1}{(2\pi)^{\frac{n_{\text{out}}}{2}}} \exp\left(-\frac{1}{2} \left\| H^{-\frac{1}{2}} \left(z - \begin{bmatrix} \underline{\rho}_{\text{out}}(b^{\omega, i}) \\ \bar{\rho}_{\text{out}}(b^{\omega, i}) \end{bmatrix} \right) \right\|_2^2\right). \end{aligned} \quad (4.35)$$

Due to the diagonality of H and due to the definition of the sampling $\underline{P}_{\mathcal{B}}$ we have

$$H = \begin{bmatrix} H^{\min} & \mathbb{0} \\ \mathbb{0} & H^{\max} \end{bmatrix},$$

and thus (4.34) and (4.35) are equivalent.

Assume that the minimal and maximal gas densities have an absolute continuous probability distribution function with probability density function ϱ_{ρ} . Then from *Section 1.1* we know that

$$\varrho_{\rho, N_{\text{KDE}}} \xrightarrow{N_{\text{KDE}} \rightarrow \infty} \varrho_{\rho} \quad \mathbb{P} - \text{almost surely.}$$

With this approximation (4.34) of the probability density function of the minimal and maximal gas densities we can compute the probability (4.30) by using (4.32) and (4.33). We have

$$\mathbb{P}_{\text{KDE}}(\rho_{\text{out}}(t) \in \mathcal{D}_{\min}^{\max} \forall t \in [0, T]) = \int_{\mathcal{D}_{\min}^{\max} \times \mathcal{D}_{\min}^{\max}} \varrho_{\rho, N_{\text{KDE}}}(z) dz.$$

We insert the approximated probability density function (4.34). It follows

$$\begin{aligned} \mathbb{P}_{\text{KDE}}(\rho_{\text{out}}(t) \in \mathcal{D}_{\min}^{\max} \forall t \in [0, T]) &= \\ &= \int_{\mathcal{D}_{\min}^{\max} \times \mathcal{D}_{\min}^{\max}} \frac{1}{N_{\text{KDE}} \prod_{j=1}^{n_{\text{out}}} h_j^{\min} \prod_{j=1}^{n_{\text{out}}} h_j^{\max}} \\ &\quad \cdot \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^{n_{\text{out}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_{1,j} - \underline{\rho}_{j,\text{out}}(b^{\omega, i})}{h_j^{\min}}\right)^2\right) \\ &\quad \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_{2,j} - \bar{\rho}_{j,\text{out}}(b^{\omega, i})}{h_j^{\max}}\right)^2\right) dz. \end{aligned}$$

Note that $\underline{\rho}_{j,\text{out}}$ is the j -th component of $\underline{\rho}_{\text{out}}$. Due to the continuity of $\varrho_{\rho, N_{\text{KDE}}}$ and the compactness of $\mathcal{D}_{\min}^{\max} \times \mathcal{D}_{\min}^{\max}$ we can apply Fubini's theorem.

Further since every factor of the product depends only on one dimension, we can swap the integral and the product. Thus we have

$$\begin{aligned} \mathbb{P}_{\text{KDE}}(\rho_{\text{out}}(t) \in \mathcal{D}_{\min}^{\max} \forall t \in [0, T]) &= \\ &= \frac{1}{N_{\text{KDE}} \prod_{j=1}^{n_{\text{out}}} h_j^{\min} \prod_{j=1}^{n_{\text{out}}} h_j^{\max}} \\ &\quad \cdot \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^{n_{\text{out}}} \int_{\rho_j^{\min}}^{\rho_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_{1,j} - \underline{\rho}_{j,\text{out}}(b^{\omega,i})}{h_j^{\min}}\right)^2\right) dz_{1,j} \\ &\quad \cdot \int_{\rho_j^{\min}}^{\rho_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_{2,j} - \bar{\rho}_{j,\text{out}}(b^{\omega,i})}{h_j^{\max}}\right)^2\right) dz_{2,j}, \end{aligned}$$

where $\rho_j^{\min} = \rho^{v_j, \min}$ and $\rho_j^{\max} = \rho^{v_j, \max}$. For $i = 1, \dots, N_{\text{KDE}}$ and $j = 1, \dots, n_{\text{out}}$ we define the functions $\varphi_{i,j}^{\min}, \varphi_{i,j}^{\max} : \mathbb{R} \rightarrow \mathbb{R}$ with

$$\varphi_{i,j}^{\min}(x) := \frac{x - \underline{\rho}_{j,\text{out}}(b^{\omega,i})}{\sqrt{2} h_j^{\min}} \quad \text{and} \quad \varphi_{i,j}^{\max}(x) := \frac{x - \bar{\rho}_{j,\text{out}}(b^{\omega,i})}{\sqrt{2} h_j^{\max}},$$

and thus it follows

$$\begin{aligned} \mathbb{P}_{\text{KDE}}(\rho_{\text{out}}(t) \in \mathcal{D}_{\min}^{\max} \forall t \in [0, T]) &= \\ &= \frac{1}{N_{\text{KDE}} \prod_{j=1}^{n_{\text{out}}} h_j^{\min} \prod_{j=1}^{n_{\text{out}}} h_j^{\max}} \\ &\quad \cdot \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^{n_{\text{out}}} \int_{\rho_j^{\min}}^{\rho_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\left(\varphi_{i,j}^{\min}(z_{1,j})\right)^2\right) dz_{1,j} \\ &\quad \cdot \int_{\rho_j^{\min}}^{\rho_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\left(\varphi_{i,j}^{\max}(z_{2,j})\right)^2\right) dz_{2,j}. \end{aligned}$$

We set $\tau_{i,j}^{\min} := \varphi_{i,j}^{\min}(z_{1,j})$ and $\tau_{i,j}^{\max} := \varphi_{i,j}^{\max}(z_{2,j})$ and use integration by substitution. Then with

$$\frac{d}{dx} \varphi_{i,j}^{\min}(x) = \frac{1}{\sqrt{2} h_j^{\min}} \quad \text{and} \quad \frac{d}{dx} \varphi_{i,j}^{\max}(x) = \frac{1}{\sqrt{2} h_j^{\max}},$$

it follows

$$\begin{aligned} \mathbb{P}_{\text{KDE}}(\rho_{\text{out}}(t) \in \mathcal{D}_{\min}^{\max} \forall t \in [0, T]) &= \\ &= \frac{1}{N_{\text{KDE}}} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^{n_{\text{out}}} \int_{\varphi_{i,j}^{\min}(\rho_j^{\min})}^{\varphi_{i,j}^{\min}(\rho_j^{\max})} \frac{1}{\sqrt{\pi}} \exp\left(-\left(\tau_{i,j}^{\min}\right)^2\right) d\tau_{i,j}^{\min} \\ &\quad \cdot \int_{\varphi_{i,j}^{\max}(\rho_j^{\min})}^{\varphi_{i,j}^{\max}(\rho_j^{\max})} \frac{1}{\sqrt{\pi}} \exp\left(-\left(\tau_{i,j}^{\max}\right)^2\right) d\tau_{i,j}^{\max}. \end{aligned}$$

Last we insert the *Gauss error function* (defined in (3.18), which leads to

$$\begin{aligned} \mathbb{P}_{\text{KDE}}(\rho_{\text{out}}(t) \in \mathcal{D}_{\min}^{\max} \forall t \in [0, T]) &= \\ &= \frac{1}{N_{\text{KDE}} 2^{2n_{\text{out}}}} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^{n_{\text{out}}} \left[\text{erf}\left(\varphi_{i,j}^{\min}(\rho_j^{\max})\right) - \text{erf}\left(\varphi_{i,j}^{\min}(\rho_j^{\min})\right) \right] \\ &\quad \cdot \left[\text{erf}\left(\varphi_{i,j}^{\max}(\rho_j^{\max})\right) - \text{erf}\left(\varphi_{i,j}^{\max}(\rho_j^{\min})\right) \right]. \end{aligned} \quad (4.36)$$

Thus the probability (4.30) can efficiently be computed by (4.36). Due to *Scheffé's lemma* (stated in *Section 1.1.3*) the approximated probability converges \mathbb{P} -almost surely to the exact probability for $N_{\text{KDE}} \rightarrow \infty$. So for N_{KDE} sufficiently large, (4.36) almost surely provides a sufficiently accurate result for the probability of a random load function to be feasible.

4.5. Optimization Problems with Probabilistic Constraints

In this section we consider a probabilistic constrained optimization problem that fits to the setting of this chapter. We analyze it in terms of the existence of an optimal solution by applying the theory from *Section 3.6*. We use the KDE approach that was applied in *Section 4.4* to approximate the probabilistic constraint and we also analyze the existence of an optimal solution for the approximated problem. Last in this section we come to one of the central results of this thesis: necessary optimality conditions for the approximated probabilistic constrained optimization problem. A KDE approach has not been applied yet to approximate probabilistic constrained optimization problems in the context of gas transport in pipeline networks, so kernel density estimation provides a new access to this topic.

Let a convex function

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \rho^{\max} \mapsto f(\rho^{\max}),$$

be given. For a probability level $\alpha \in (0, 1)$ consider the optimization problem

$$\begin{aligned} \min_{\rho^{\max} \in \mathbb{R}^{n_{\text{out}}}} \quad & f(\rho^{\max}) \\ \text{s.t.} \quad & \mathbb{P}(b \in M(t, \rho^{\max}) \quad \forall t \in [0, T]) \geq \alpha, \\ & \rho^{\max} \geq \rho^{\min}. \end{aligned} \quad (4.37)$$

The feasible set $M(t)$ is defined in (4.29). We write here $M(t, \rho^{\max})$ to explicitly state that the feasible set depends on the upper bound ρ^{\max} for the gas density. Due to (4.32) we can write (4.37) equivalently as

$$\begin{aligned} \min_{\rho^{\max} \in \mathbb{R}^{n_{\text{out}}}} \quad & f(\rho^{\max}) \\ \text{s.t.} \quad & \mathbb{P}(\rho_{\text{out}}(t) \in \mathcal{D}_{\min}^{\max} \quad \forall t \in [0, T]) \geq \alpha, \\ & \rho^{\max} \geq \rho^{\min}, \end{aligned}$$

where the second constraint has to be understood component-by-component. Due to (4.33) the probability in the first constraint is equivalent to

$$\mathbb{P}(\underline{\rho}_{\text{out}}, \bar{\rho}_{\text{out}} \in \mathcal{D}_{\min}^{\max}),$$

where the minimum and the maximum have to be understood component-by-component. So we can write (4.37) equivalently as

$$\begin{aligned} \min_{\rho^{\max} \in \mathbb{R}^{n_{\text{out}}}} \quad & f(\rho^{\max}) \\ \text{s.t.} \quad & \mathbb{P}(\underline{\rho}_{\text{out}}, \bar{\rho}_{\text{out}} \in \mathcal{D}_{\min}^{\max}), \\ & \rho^{\max} \geq \rho^{\min}. \end{aligned} \tag{4.38}$$

We assume, that $[\underline{\rho}_{\text{out}}(t), \bar{\rho}_{\text{out}}(t)] \in \mathbb{R}^{2n_{\text{out}}}$ has an absolutely continuous probability distribution with probability density function $\varrho_{\text{out}, \rho}$ on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. This assumption is quite strong (see *Remark 4.9*) but it is necessary for the convergence of the probability.

Since the probabilistic constrained optimization problem (4.38) does not depend on the time anymore, we can apply the theory from *Section 3.6* here. Note that the constraint $\rho^{\min} \geq \rho^{\max}$ cannot be active.

Corollary 4.10. *Under the same assumptions as in Lemma 3.16 there exists a solution of (4.37). Let $\rho^{*, \max} \in \mathbb{R}^{n_{\text{out}}}$ be a solution of (4.37). Then we have*

$$\mathbb{P}(b \in M(t, \rho^{*, \max}) \quad \forall t \in [0, T]) = \alpha.$$

Next we consider the approximated probabilistic constrained optimization problem

$$\begin{aligned} \min_{\rho^{\max} \in \mathbb{R}^{n_{\text{out}}}} \quad & f(\rho^{\max}) \\ \text{s.t.} \quad & \mathbb{P}_{\text{KDE}}(b \in M(t, \rho^{\max}) \quad \forall t \in [0, T]) \geq \alpha, \\ & \rho^{\max} \geq \rho^{\min}. \end{aligned} \tag{4.39}$$

Equivalently to (4.38) we can write (4.39) as

$$\begin{aligned} \min_{\rho^{\max} \in \mathbb{R}^{n_{\text{out}}}} \quad & f(\rho^{\max}) \\ \text{s.t.} \quad & \mathbb{P}_{\text{KDE}}(\underline{\rho}_{\text{out}}, \bar{\rho}_{\text{out}} \in \mathcal{P}_{\text{min}}^{\max}), \\ & \rho^{\max} \geq \rho^{\min}, \end{aligned} \quad (4.40)$$

which allows us to also use the theory of *Section 3.6*.

Corollary 4.11. *Under the same assumptions as in Corollary 3.20 there exists a solution of (4.39). Let $\rho^{*,\max} \in \mathbb{R}^{n_{\text{out}}}$ be a solution of (4.39). Then we have*

$$\mathbb{P}_{\text{KDE}}(b \in M(t, \rho^{*,\max}) \quad \forall t \in [0, T]) = \alpha.$$

Further *Theorem 3.22* guarantees under weak assumptions that the solution of (4.39) is almost surely δ -close to the solution (4.37) for N_{KDE} sufficiently large.

Last in this section we state necessary optimality conditions for the approximated probabilistic constrained optimization problem (4.39). Hence we first compute the derivative of the approximated probabilistic constraint. Due to (4.36) we have

$$\begin{aligned} \mathbb{P}_{\text{KDE}}(b \in M(t, \rho^{\max}) \quad \forall t \in [0, T]) &= \\ &= \frac{1}{N_{\text{KDE}} 2^{2n_{\text{out}}}} \sum_{i=1}^{N_{\text{KDE}}} \prod_{j=1}^{n_{\text{out}}} \left[\text{erf}\left(\varphi_{i,j}^{\min}(\rho_j^{\max})\right) - \text{erf}\left(\varphi_{i,j}^{\min}(\rho_j^{\min})\right) \right] \\ &\quad \cdot \left[\text{erf}\left(\varphi_{i,j}^{\max}(\rho_j^{\max})\right) - \text{erf}\left(\varphi_{i,j}^{\max}(\rho_j^{\min})\right) \right]. \end{aligned}$$

We define

$$g^{\alpha} : \mathbb{R}^{n_{\text{out}}} \rightarrow \mathbb{R}, \quad g : \rho^{\max} \mapsto \alpha - \mathbb{P}_{\text{KDE}}(b \in M(t, \rho^{\max}) \quad \forall t \in [0, T]).$$

We compute the partial derivative of g^{α} with respect to ρ_k^{\max} ($k \in \{1, \dots, n_{\text{out}}\}$).

With the product rule it follows

$$\begin{aligned}
& \frac{\partial}{\partial \rho_k^{\max}} g^\alpha(\rho^{\max}) = \\
& = -\frac{1}{N_{\text{KDE}} 2^{2n_{\text{out}}}} \sum_{i=1}^{N_{\text{KDE}}} \left[\prod_{j=1, j \neq k}^{n_{\text{out}}} \left[\text{erf}\left(\varphi_{i,j}^{\min}(\rho_j^{\max})\right) - \text{erf}\left(\varphi_{i,j}^{\min}(\rho_j^{\min})\right) \right] \right. \\
& \quad \left. \cdot \left[\text{erf}\left(\varphi_{i,j}^{\max}(\rho_j^{\max})\right) - \text{erf}\left(\varphi_{i,j}^{\max}(\rho_j^{\min})\right) \right] \right] \\
& \quad \cdot \left[\frac{d}{d\rho_k^{\max}} \text{erf}\left(\varphi_{i,k}^{\min}(\rho_k^{\max})\right) \left[\text{erf}\left(\varphi_{i,k}^{\max}(\rho_k^{\max})\right) - \text{erf}\left(\varphi_{i,k}^{\max}(\rho_k^{\min})\right) \right] \right. \\
& \quad \left. + \frac{d}{d\rho_k^{\max}} \text{erf}\left(\varphi_{i,k}^{\max}(\rho_k^{\max})\right) \left[\text{erf}\left(\varphi_{i,k}^{\min}(\rho_k^{\max})\right) - \text{erf}\left(\varphi_{i,k}^{\min}(\rho_k^{\min})\right) \right] \right].
\end{aligned}$$

The derivative of the Gauss error function is given by

$$\frac{d}{dx} \text{erf}(x) = \frac{2}{\sqrt{\pi}} \exp(-x^2),$$

and the derivatives of $\varphi_{i,k}^{\min}$ and $\varphi_{i,k}^{\max}$ with respect to ρ_k^{\max} are given by

$$\frac{d}{d\rho_k^{\max}} \varphi_{i,k}^{\min}(\rho^{\max}) = \frac{1}{\sqrt{2} h_k^{\min}} \quad \text{and} \quad \frac{d}{d\rho_k^{\max}} \varphi_{i,k}^{\max}(\rho^{\max}) = \frac{1}{\sqrt{2} h_k^{\max}}.$$

Thus by applying the chain rule it follows

$$\begin{aligned}
& \frac{\partial}{\partial \rho_k^{\max}} g^\alpha(\rho^{\max}) = \\
& = -\frac{1}{N_{\text{KDE}} 2^{2n_{\text{out}}}} \sum_{i=1}^{N_{\text{KDE}}} \left[\prod_{j=1, j \neq k}^{n_{\text{out}}} \left[\text{erf}\left(\varphi_{i,j}^{\min}(\rho_j^{\max})\right) - \text{erf}\left(\varphi_{i,j}^{\min}(\rho_j^{\min})\right) \right] \right. \\
& \quad \left. \cdot \left[\text{erf}\left(\varphi_{i,j}^{\max}(\rho_j^{\max})\right) - \text{erf}\left(\varphi_{i,j}^{\max}(\rho_j^{\min})\right) \right] \right] \\
& \quad \cdot \frac{\sqrt{2}}{\sqrt{\pi}} \left[\frac{1}{h_k^{\min}} \exp\left(-(\varphi_{i,k}^{\min}(\rho_k^{\max}))^2\right) \left[\text{erf}\left(\varphi_{i,k}^{\max}(\rho_k^{\max})\right) - \text{erf}\left(\varphi_{i,k}^{\max}(\rho_k^{\min})\right) \right] \right. \\
& \quad \left. + \frac{1}{h_k^{\max}} \exp\left(-(\varphi_{i,k}^{\max}(\rho_k^{\max}))^2\right) \left[\text{erf}\left(\varphi_{i,k}^{\min}(\rho_k^{\max})\right) - \text{erf}\left(\varphi_{i,k}^{\min}(\rho_k^{\min})\right) \right] \right].
\end{aligned} \tag{4.41}$$

So the k -th component of the gradient $\nabla g^\alpha(\rho^{\max})$ is given by (4.41). As it is mentioned in *Remark 3.27*, the linear independence constraint qualification (LICQ) holds for (4.39) and we can state the necessary optimality conditions for the approximated probabilistic constrained optimization problem (4.39).

Corollary 4.12. *Let $\rho^{*,\max} \in \mathbb{R}^{n_{out}}$ be a (local) optimal solution of (4.39). Since the LICQ holds in $\rho^{*,\max}$, there exists a multiplier $\mu^* \geq 0$, s.t.*

$$\begin{aligned}\nabla f(\rho^{*,\max}) + \mu^* \nabla g^\alpha(\rho^{*,\max}) &= 0, \\ g^\alpha(\rho^{*,\max}) &\leq 0, \\ \mu^* g^\alpha(\rho^{*,\max}) &= 0,\end{aligned}$$

where the components of the gradient of g^α are stated in (4.41). Thus $(\rho^{*,\max}, \mu^*) \in \mathbb{R}^{n_{out}} \times \mathbb{R}$ is a Karush-Kuhn-Tucker point.

Due to the fact that if the minimal and maximal gas densities in $[0, T]$ satisfy the density bounds then the gas density satisfies the density bounds for all times in $[0, T]$, we were able to use the theory on the probabilistic constrained optimization problems that we derived in *Section 3.6*. The necessary optimality conditions of *Corollary 4.12* can be used to characterize the solution of the approximated problem (4.39) as it is shown in *Figure 3.12*.

4.6. A Numerical Example

Last in this chapter we present an example for (qISO) on a single edge. We first solve an deterministic optimization problem to get the optimal deterministic upper density bound. Then we consider uncertain gas outflow and we compute the probability that the uncertain gas outflow stays in the deterministic bounds. Last we solve the corresponding probabilistic constrained optimization problem to get the optimal probabilistic upper density bound.

Consider (qISO) on a single edge with

$$\rho^{v_1}(t) \in [\rho^{\min}, \rho^{\max}],$$

as it is shown in *Figure 4.3*.

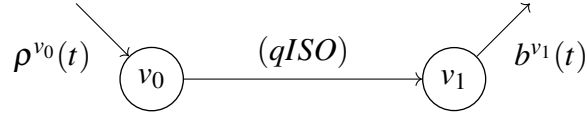


Figure 4.3.: Dynamic setting on a single edge

Motivated by the application we demand that the pressure at node v_1 has to stay larger than 40bar. This is a common lower bound for the pressure in realistic scenarios¹. The corresponding lower bound for the density directly follow by the ideal gas equation. All values and constants are shown in *Table 4.1.*

Variable	Letter	Value	Unit
lower density bound	ρ^{\min}	34	kg/m ³
speed of sound in the gas	c	343	m/s
pipe friction coefficient	λ^F	0.1	
pipe diameter	D	0.5	m
pipe length	L	30	km
final time	T	24	h

Table 4.1.: Values for the example with one edge.

We assume that the inlet pressure at node v_0 is 55bar, thus we set

$$\rho_0(t) = 46.75 \frac{\text{kg}}{\text{m}^3}.$$

Note that the time period is a full day and the gas demand at night is lower than the gas demand during the day. We set

$$b(t) = -15 \sin\left(\frac{\pi}{12 \cdot 60^2} (t + 60^2 \pi)\right) + \frac{15}{2} \cos\left(\frac{\pi}{6 \cdot 60^2} (t + 60^2 \pi)\right) + 150,$$

as gas outflow (see *Figure 4.4*). For the initial conditions $\rho_{\text{ini}}(x)$ and $q_{\text{ini}}(x)$ we simulate (qISO) with the constant boundary conditions $\rho_0(0)$ and $b(0)$ for 48 hours, which is sufficient that the state reaches an equilibrium. Then the resulting state after 48 hours is stationary and satisfies the C^0 -compatibility between initial and boundary conditions. We use this stationary state as initial conditions $\rho_{\text{ini}}(x)$ and $q_{\text{ini}}(x)$.

¹Realistic scenarios for different networks can be found on <https://gaslib.zib.de/>

For the initial and boundary conditions stated above we solve (qISO) for $(t, x) \in [0, 24\text{h}] \times [0, 30\text{km}]$. The solver was provided by CRC² Transregio 154: *Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks*, subproject C04. It is based on a conforming mixed finite elements method for space discretization and an implicit Euler method for time discretization. The resulting nonlinear system is solved with a *Newton* method. The solution is plotted for every hour in *Figure A.3* and *Figure A.4* in the *Appendix A.2*.

Now we compute the lowest upper density bound $\rho_{\text{det}}^{*,\text{max}}$ (for deterministic boundary data $b(t)$), s.t. the density at node v_1 stays between ρ^{min} and $\rho_{\text{det}}^{*,\text{max}}$ for the full time period $[0, 24\text{h}]$. Let $\rho^{v_1}(t)$ be the density at node v_1 . For $w > 0$ consider the optimization problem

$$\left\{ \begin{array}{l} \min_{\rho_{\text{det}}^{\text{max}} \in \mathbb{R}} w \rho_{\text{det}}^{\text{max}}, \\ \text{s.t. } \rho^{v_1}(t) \in [\rho^{\text{min}}, \rho_{\text{det}}^{\text{max}}] \quad \forall t \in [0, T], \\ \rho_{\text{det}}^{\text{max}} \geq \rho^{\text{min}}. \end{array} \right. \quad (4.42)$$

This optimization problem can be equivalently written as

$$\left\{ \begin{array}{l} \min_{\rho_{\text{det}}^{\text{max}} \in \mathbb{R}} w \rho_{\text{det}}^{\text{max}}, \\ \text{s.t. } \min_{t \in [0, T]} \rho^{v_1}(t) \in [\rho^{\text{min}}, \rho_{\text{det}}^{\text{max}}], \\ \max_{t \in [0, T]} \rho^{v_1}(t) \in [\rho^{\text{min}}, \rho_{\text{det}}^{\text{max}}], \\ \rho_{\text{det}}^{\text{max}} \geq \rho^{\text{min}}. \end{array} \right. \quad (4.43)$$

Since ρ^{v_1} is continuous, the minimum and the maximum exist on $[0, T]$ and the existence of an optimal solution of (4.42) directly follows. We solve this problem numerically in *MATLAB*[®] 2019a using the *fmincon.m* routine with an interior point algorithm, which returns

$$\rho_{\text{det}}^{*,\text{max}} = 42.15 \frac{\text{kg}}{\text{m}^3},$$

as lowest upper density bound in the deterministic case. This corresponds to an upper pressure bound of 49.6bar.

²Collaborative Research Center

In the next step, we show that $\rho_{\text{det}}^{*,\text{max}}$ is not robust in the sense of probabilities with respect to uncertain boundary data. We write $b(t)$ as Fourier series (as it was introduced in *Section 4.2*) using the basis functions $\psi_m(t)$ (defined in (4.5)) and the Fourier coefficients α_m^0 (defined in (4.6)). For $m = 0, 1, \dots$ consider Gaussian distributed random variables

$$\xi_m \sim \mathcal{N}\left(1, \sqrt{0.1}\right),$$

with mean value 1 and standard deviation $\sqrt{0.1}$ on an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then the random boundary functions are given by

$$b^\omega(t) = \left(\sum_{m=0}^{\infty} \xi_m(\omega) \alpha_m^0 \psi_m(t) \right) + b(0).$$

Note that since $b(0) \neq 0$ the Fourier coefficients are given for the shifted function $b(t) - b(0)$. For the implementation, we cut the Fourier series after 11 terms, i.e. $N_F = 10$. For visualization the deterministic boundary data $b(t)$ and 15 corresponding random scenarios are shown in *Figure 4.4*.

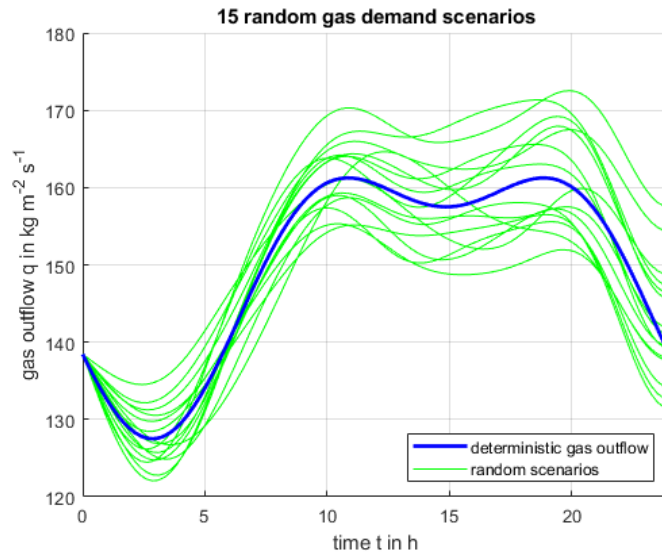


Figure 4.4.: Deterministic gas demand $b(t)$ (plotted in blue) and 15 random scenarios $b^\omega(t)$ (plotted in green)

The probability density function of the gas density is given by (4.34) and it is shown in *Figure 4.5*. We use (4.36) with $5 \cdot 10^3$ samples to compute the probability, that the gas density is larger than ρ^{\min} and smaller than $\rho_{\det}^{*,\max}$. We have

$$\mathbb{P}_{N_{\text{KDE}}} \left(\rho^{v_1}(t) \in \left[\rho^{\min}, \rho_{\det}^{*,\max} \right] \forall t \in [0, T] \right) = 46.65\%.$$

So for $\alpha > 0.4665$ the deterministic optimal gas density bound $\rho_{\det}^{*,\max}$ is not robust in a probabilistic sense.

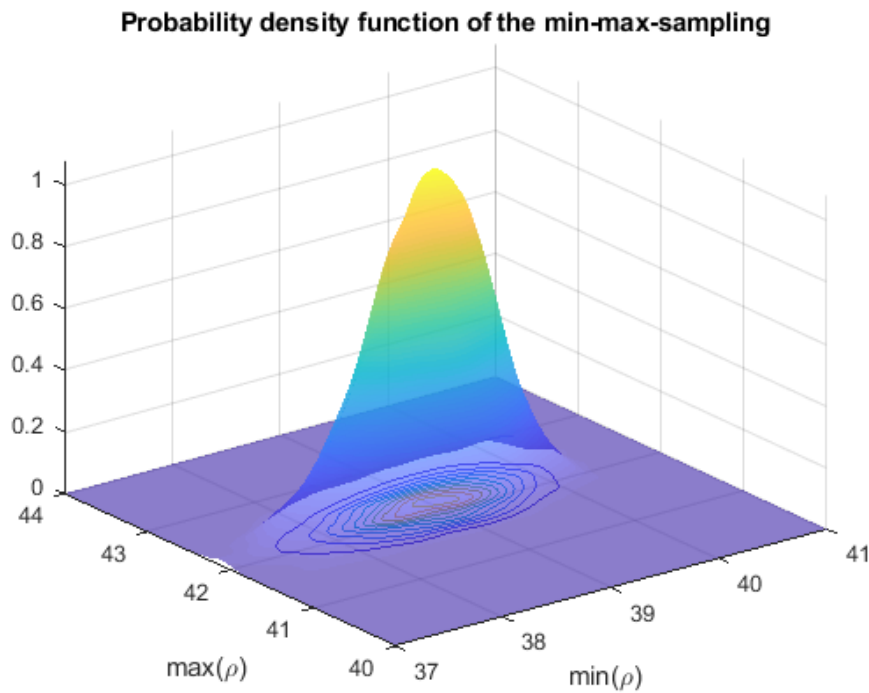


Figure 4.5.: Probability density function of the minimal and maximal gas density

Remark 4.13. *At first glance it was to be expected that the probability is about 50% since we perturbed $b(t)$ with Gaussian distributed numbers. But in general this is not true as it is shown in [Schuster et al., 2021]. There the authors have a similar setting on a network with 11 pipes and the resulting probability is about 35%.*

Last in this example we compute an upper gas density bound that is more robust in a probabilistic sense against uncertain gas outflow. For

$$\alpha := 0.90,$$

consider the following probabilistic constrained optimization problem:

$$\left\{ \begin{array}{l} \min_{\rho_{\text{prob}}^{\text{max}} \geq \rho^{\text{min}}} w \rho_{\text{prob}}^{\text{max}}, \\ \text{s.t. } \mathbb{P}_{\text{KDE}} \left(\rho^{v_1}(t) \in [\rho^{\text{min}}, \rho_{\text{prob}}^{\text{max}}] \quad \forall t \in [0, T] \right) \geq 0.90. \end{array} \right.$$

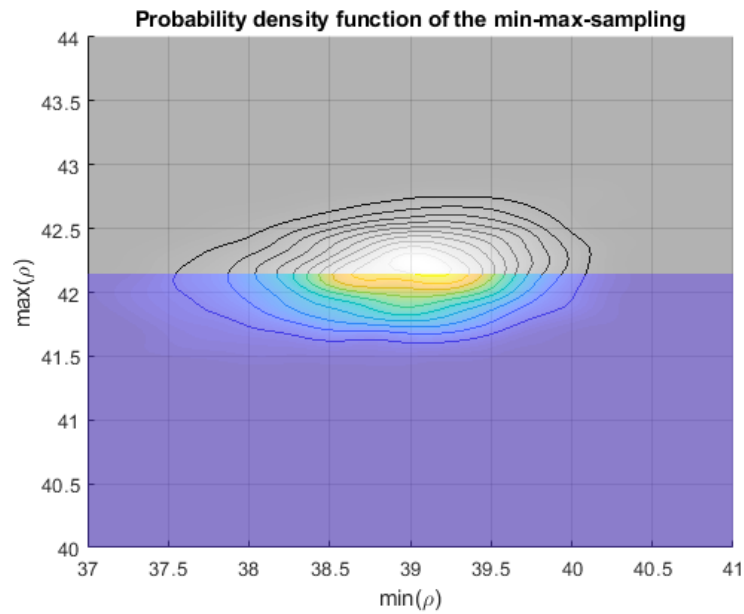
This probabilistic constrained optimization problem can be equivalently written as

$$\left\{ \begin{array}{l} \min_{\rho_{\text{prob}}^{\text{max}} \geq \rho^{\text{min}}} w \rho_{\text{prob}}^{\text{max}}, \\ \text{s.t. } \mathbb{P} \left(\begin{array}{l} \min_{t \in [0, T]} \rho^{v_1}(t) \in [\rho^{\text{min}}, \rho_{\text{prob}}^{\text{max}}] \\ \max_{t \in [0, T]} \rho^{v_1}(t) \in [\rho^{\text{min}}, \rho_{\text{prob}}^{\text{max}}] \end{array} \right) \geq 0.90. \end{array} \right.$$

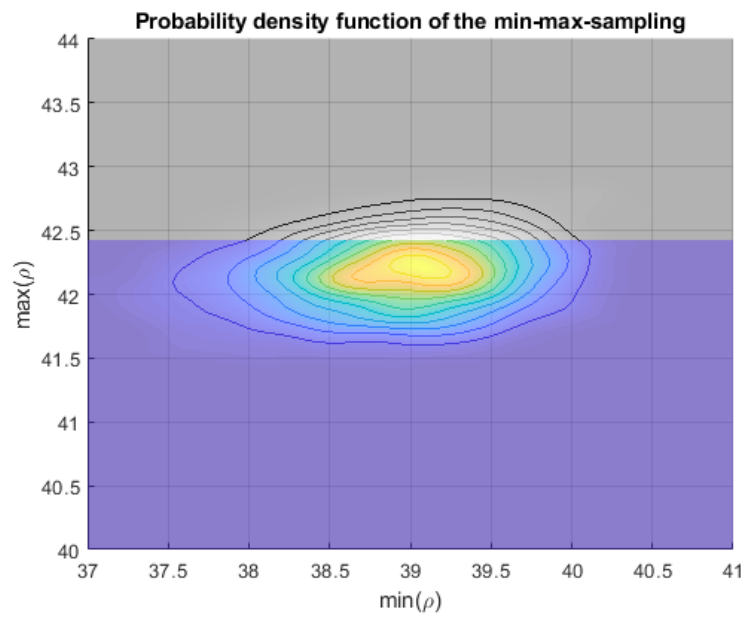
Analogously to the deterministic problem (4.42) the minimum and the maximum exist on $[0, T]$ and the existence of a solution follows from *Lemma 3.16*. We also solve this problem numerically in *MATLAB*[®] 2019a using the *fmincon.m* routine with an interior point algorithm, which returns

$$\rho_{\text{prob}}^{*, \text{max}} = 42.49 \frac{\text{kg}}{\text{m}^3},$$

as lowest upper density bound. This corresponds to an upper pressure bound of 50bar. The solution $\rho_{\text{prob}}^{*, \text{max}}$ is slightly larger than $\rho_{\text{det}}^{*, \text{max}}$ but this slight increase guarantees 90% robustness (in a probabilistic sense) against uncertain gas outflow. The effect of this slight increase in the upper density bound is visualized in *Figure 4.6*. The part of the probability density function of the minimal and maximal gas density that satisfies the optimal deterministic resp. the optimal probabilistic gas density bounds is shown in *Figure 4.6 (a)* resp. *Figure 4.6 (b)*.



(a) Probability density function of the minimal and maximal gas density in $[\rho^{\min}, \rho_{\det}^{*,\max}]$



(b) Probability density function of the minimal and maximal gas density in $[\rho^{\min}, \rho_{\text{prob}}^{*,\max}]$

Figure 4.6.: Probability density functions of the minimal and maximal gas densities for the lowest upper density bound in the deterministic and the stochastic setting

Summary, Conclusion and Perspectives

In the last chapter of this thesis we shortly summarize the content and we highlight the key ideas and main results once more. The aim of this thesis is to give a contribution to the analysis of optimization problems with probabilistic constraints. Among other things, this requires the computation of probabilities. We introduce a kernel density estimation (KDE) approach to estimate unknown probability density functions and we introduce the spheric radial decomposition (SRD) to reduce the dimension of integration.

Both tools are applied to a stationary gas transport model on a network with uncertain gas outflow and both tools provide an efficient and well working possibility to compute the probability that a random load vector (the gas outflow) is feasible. Feasibility is defined by box constraints for the pressure at the network junctions. A significant advantage of the SRD is, that it can be applied to a model extension with unknown inlet pressure and compressor control, whereas we cannot apply the KDE to this model extension. On the contrary we can apply the KDE approach to a dynamic gas transport model on a network with uncertain gas outflow, even if an explicit solution is not known for this model. *Section 4.4* and *Section 4.5* are a novelty in the mathematics of gas transport resp. probabilistic constrained optimization in the context of gas transport.

In *Section 3.6* and *Section 4.5* we use the KDE approach to approximate the probabilistic constraints and we prove that weak additional assumptions imply that the optimal solutions of the approximated probabilistic constrained optimization problems are close to the solutions of the exact probabilistic constrained optimization problem. This is another key result of this thesis.

In a nutshell we show that the KDE approach provides new and interesting possibilities to analyze and solve probabilistic constrained optimization problems resp. to compute the probabilities in the probabilistic constraints. These probabilities can be seen as a measure for the robustness of certain

variables against uncertainty e.g. in the boundary conditions. Since classical robust optimization often leads to overconservative results, probabilistic robustness takes into account a weighting of uncertain scenarios by probability distribution functions resp. probability density functions. Many applications like e.g. gas transport in pipeline networks benefit from a probabilistic robustness instead of a classical robustness analysis.

Finally we mention two promising extensions of this work. Since the optimization part in this thesis was finite dimensional, a consequential next step would be to consider infinite dimensional probabilistic constrained optimization problems like e.g. time dependent optimal control problems with probabilistic constraints. In [Farshbaf-Shaker et al., 2020] the authors analyze a similar problem. They consider an optimal Neumann control problem with a probabilistic constraint for the terminal energy for a system governed by the wave equation with uncertain initial data. As shown in this thesis, our KDE approach works fine for time dependent uncertainty. So it could also be used to approximate probabilities in infinite dimensional probabilistic constrained optimization problems like the optimal Neumann control problem in [Farshbaf-Shaker et al., 2020]. Another promising extension of this work is to combine the KDE approach with a stochastic collocation method. Since we need to solve the model for every scenario to get a sampling of the pressures resp. densities at the outflow nodes, it would be wise to use a stochastic collocation method to approximate the stochastic solution of the model. Then the sampling of pressures resp. densities at the outflow nodes would result in an evaluation of the stochastic solution, which would increase the speed of computation enormously.

Even if the KDE probability does not guarantee strong convergence, it almost surely converges pointwise to the exact probability. Albert Einstein once said¹

„Das, wobei unsere Berechnungen versagen, nennen wir Zufall.“

¹<https://www.zitat-service.de/quotations/419>

A. Appendix

A.1. Comparison of radial symmetric and product kernel functions

The following figures show the difference between two dimensional radial symmetric kernel functions and the corresponding product kernels.

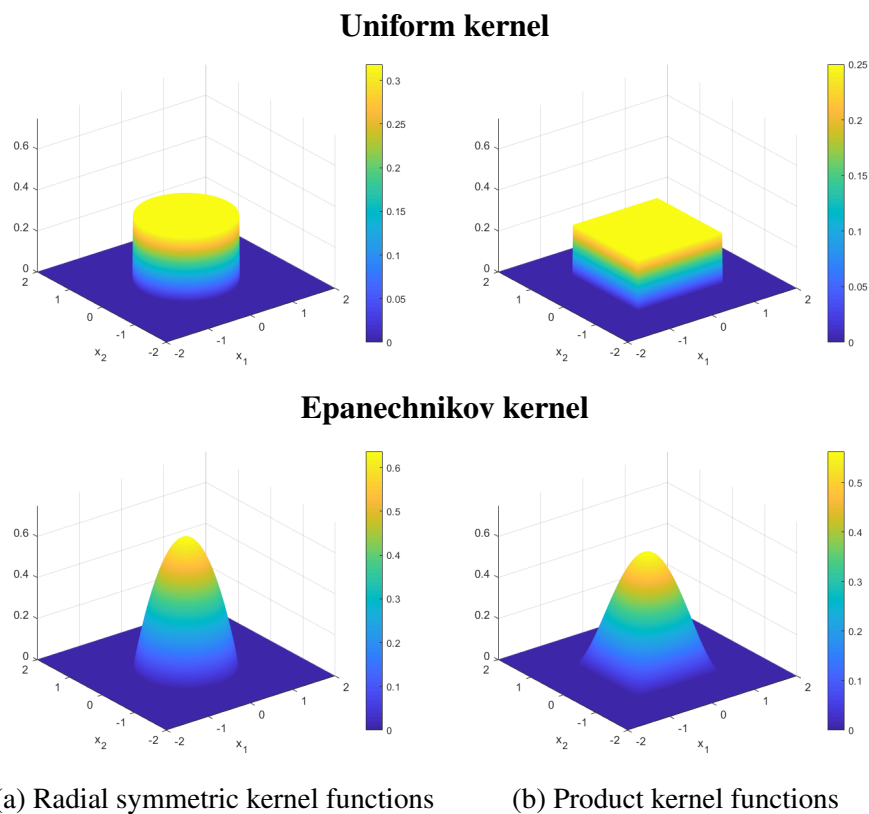
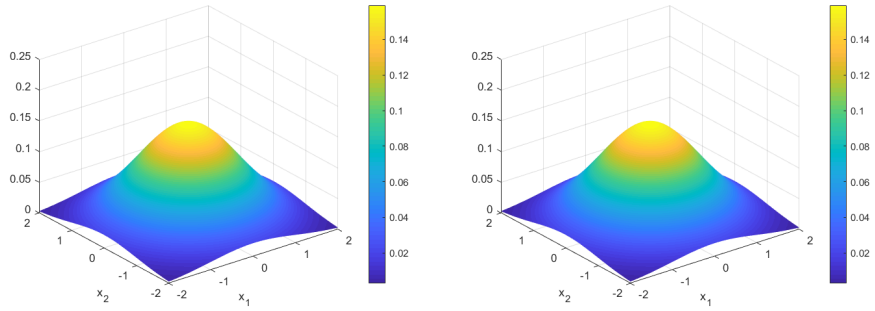
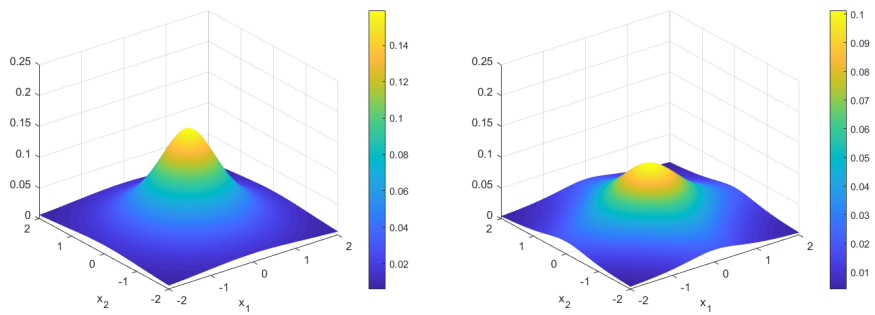


Figure A.1.: Comparison of the radial symmetric kernel functions and the product kernel functions with compact support given in *Section 1.1.2* in 2D

Gaussian kernel



Cauchy kernel



(a) Radial symmetric kernel functions

(b) Product kernel functions

Figure A.2.: Comparison of the radial symmetric kernel functions and the product kernel functions with global support given in *Section 1.1.2* in 2D

A.2. Solution of ISO1 on one edge

The following figure shows the solution of (qISO) on one edge for every hour in $t \in [0, 24\text{h}]$ (cf. *Section 4.6*). The blue lines show the density in $\text{kg} \cdot \text{m}^{-3}$ and the red lines show the flow in $\text{kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$.

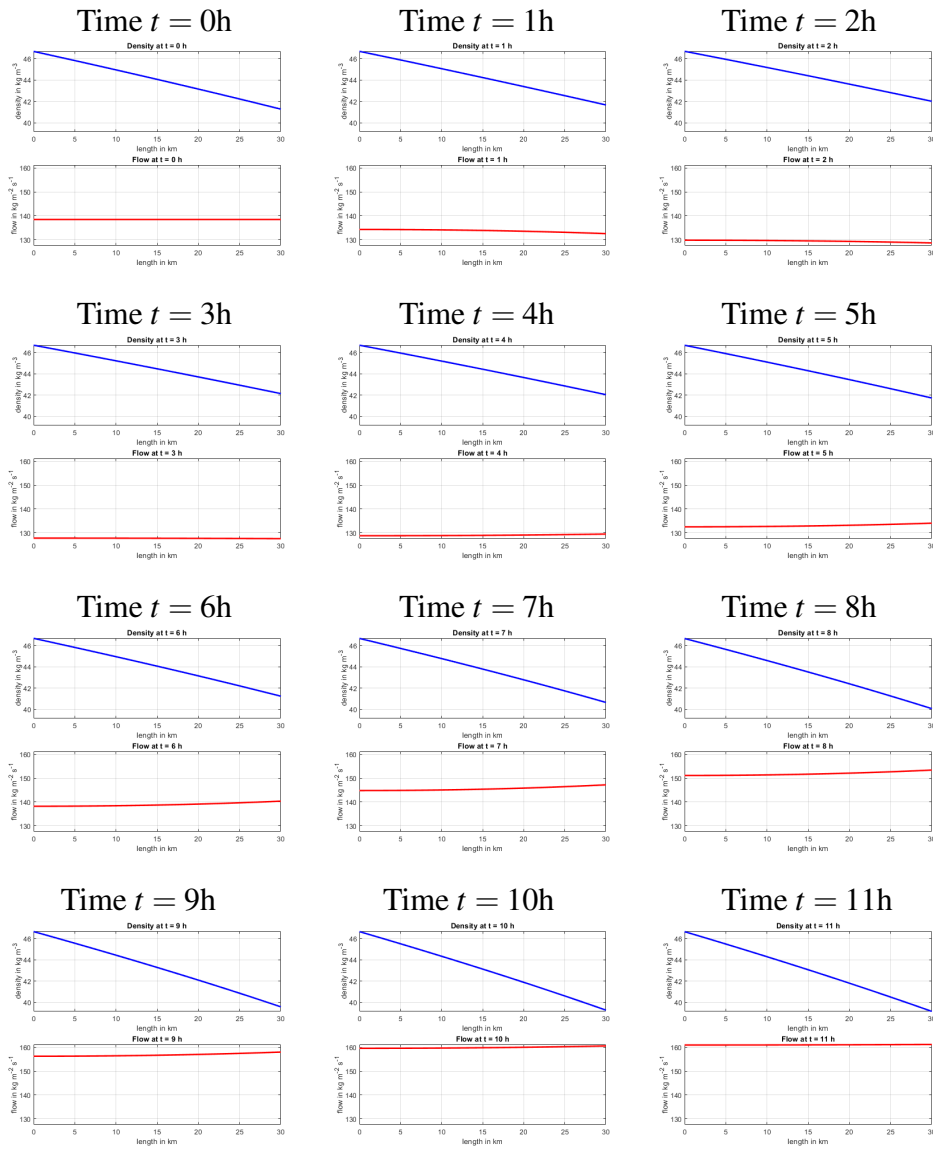


Figure A.3.: Solution of (qISO) on one edge from $t = 0$ to $t = 11\text{h}$

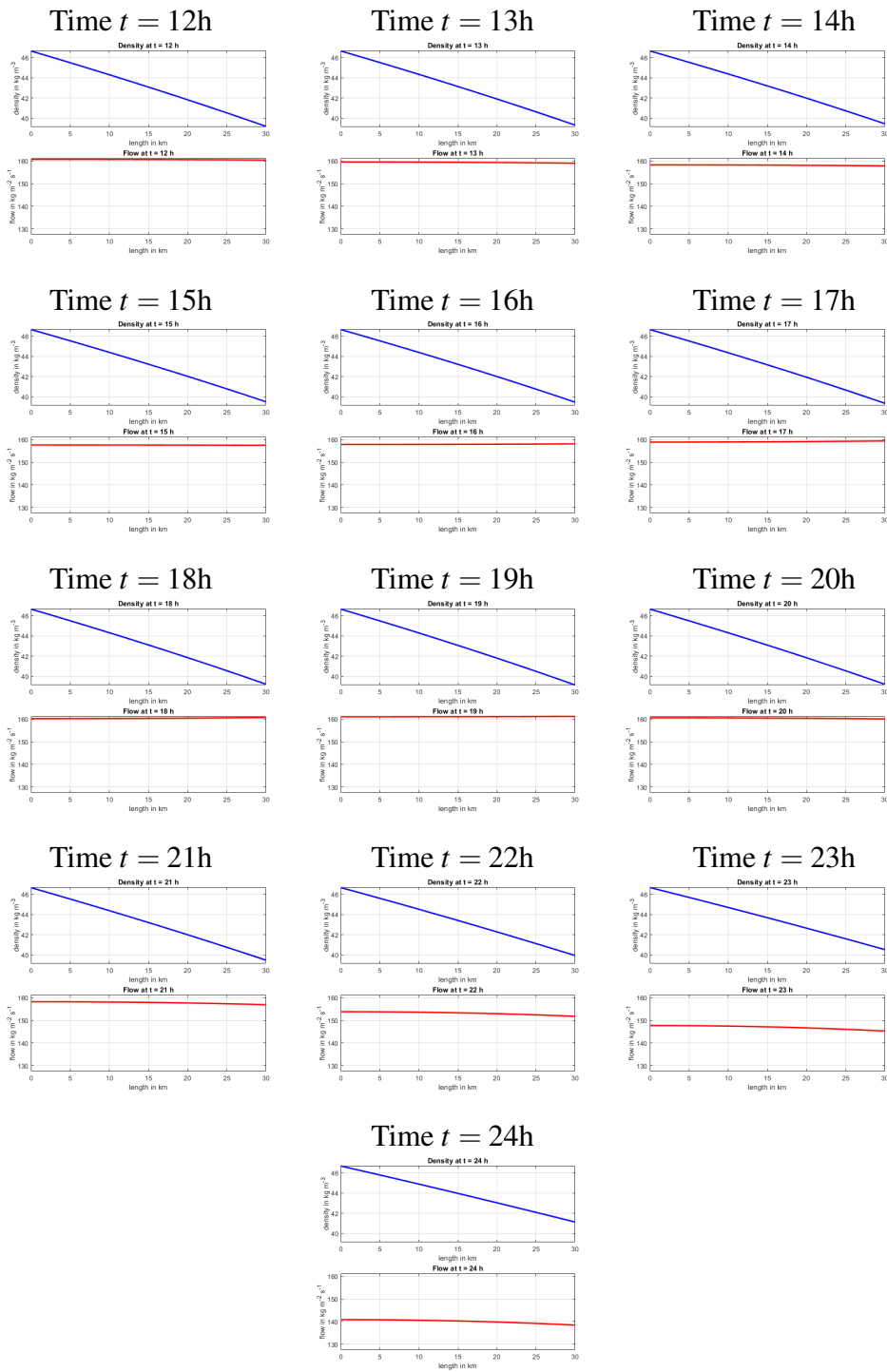


Figure A.4.: Solution of (q)ISO on one edge from $t = 12$ to $t = 24$ h

Bibliography

- [Abell et al., 1998] Abell, M. L., Braselton, J. P., and Rafter, J. A. (1998). *Statistics with Mathematica*. Academic Press.
- [Adelhütte et al., 2020] Adelhütte, D., Aßmann, D., Gradón, T. G., Gugat, M., Heitsch, H., Henrion, R., Liers, F., Nitsche, S., Schultz, R., Stingl, M., and Wintergerst, D. (2020). Joint model of probabilistic-robust (probust) constraints applied to gas network optimization. *Vietnam J. Math.*
- [Ahuja et al., 1993] Ahuja, R., Magnanti, T., and Orlin, J. (1993). *Network Flows: Theory, Algorithms and Applications*. Prentice Hall.
- [Almeida et al., 2014] Almeida, J., Velásquez, J., and Barbieri, R. (2014). A methodology for calculating the natural gas compressibility factor for a distribution network. *Petrol. Sci. Technol.*, 32(21).
- [Andrews, 1998] Andrews, L. C. (1998). *Special Functions of Mathematics for Engineers*. SPIE Press, 2 edition.
- [Banda and Herty, 2011] Banda, M. and Herty, M. (2011). Towards a space mapping approach to dynamic compressor optimization of gas networks. *Optim. Control Appl. Meth.*, 32:253–269.
- [Banda et al., 2006a] Banda, M., Herty, M., and Klar, A. (2006a). Coupling conditions for gas networks governed by the isothermal Euler equations. *Netw. Heretog. Media*, 1(2):295–314.
- [Banda et al., 2006b] Banda, M., Herty, M., and Klar, A. (2006b). Gas flow in pipeline networks. *Netw. Heretog. Media*, 1(1):41–56.
- [Bressan et al., 2014] Bressan, A., Canic, S., Herty, M., and Piccoli, B. (2014). Flows on networks: Recent results and perspectives. *EMS Surv. Math. Sci*, 1.
- [Caillaud et al., 2018] Caillaud, J.-B., Cerf, M., Sassi, A., Trélat, E., and Zidani, H. (2018). Solving chance constrained optimal control problems in aerospace via kernel density estimation. *Optimal Control Appl. Methods*, 39:1833–1858.
- [Chapeau-Blondeau and Monir, 2002] Chapeau-Blondeau, F. and Monir, A. (2002). Numerical evaluation of the Lambert W function and application to generation of generalized Gaussian noise with exponent 1/2. *IEEE Trans. Signal Process*, 50:2160–2165.

- [Chen, 2012] Chen, X. (2012). Lipschitz continuous solutions to the Cauchy problem for quasi-linear hyperbolic systems. *Chin. Ann. Math. Ser. B*, 33:521–536.
- [Colombo et al., 2008] Colombo, R., Guerra, G., Herty, M., and Sachers, V. (2008). Modeling and optimal control of networks of pipes and canals. *arXiv: Analysis of PDEs*.
- [Colombo et al., 2009] Colombo, R., Guerra, G., Herty, M., and Schleper, V. (2009). Optimal control in networks of pipes and canals. *SIAM J. Control Optim.*, 48:2032–2050.
- [Colombo and Marcellini, 2009] Colombo, R. and Marcellini, F. (2009). Smooth and discontinuous junctions in the p-system. *J. Math. Anal. Appl.*, 361.
- [Corless et al., 1996] Corless, R. M., Gonnet, G. H., Jeffrey, D. J., and Knuth, D. E. (1996). On the LambertW function. *Adv. Comput. Math.*, 5:329–359.
- [Cormen et al., 2009] Cormen, T., Leiserson, C., Rivest, R., and Stein, C. (2009). *Introduction to Algorithms*. The MIT Press, 3 edition.
- [Cuzick and Lai, 1980] Cuzick, J. and Lai, T. (1980). On random Fourier series. *Trans. Amer. Math. Soc.*, 261(1):53–80.
- [Devroye and Györfi, 1985] Devroye, L. and Györfi, L. (1985). *Nonparametric density estimation: the LI view*. Wiley series in probability and mathematical statistics. Wiley.
- [Domschke et al., 2017] Domschke, P., Hiller, B., Lang, J., and Tischendorf, C. (2017). Modellierung von Gasnetzwerken: Eine Übersicht. Technical Report 2717, Technische Universität Darmstadt.
- [Duller, 2018] Duller, C. (2018). *Einführung in die nichtparametrische Statistik mit SAS, R und SPSS*. Springer, 2 edition.
- [Elstrodt, 2011] Elstrodt, J. (2011). *Maß- und Integrationstheorie*. Springer.
- [Farshbaf-Shaker et al., 2018] Farshbaf-Shaker, M., Henrion, R., and Hömberg, D. (2018). Properties of chance constraints in infinite dimensions with an application to pde constrained optimization. *Set-Valued Var. Anal.*, 26:821–841.
- [Farshbaf-Shaker et al., 2020] Farshbaf-Shaker, M. H., Gugat, M., Heitsch, H., and Henrion, R. (2020). Optimal Neumann boundary control of a vibrating string with uncertain initial data and probabilistic terminal constraints. *SIAM J. Control. Optim.*, 58:2288–2311.
- [Garavello and Piccoli, 2009] Garavello, M. and Piccoli, B. (2009). Conservation laws on complex networks. *Ann. l H. Poincaré-An.*, 26:1925–1951.

- [Genz and Bretz, 2009] Genz, A. and Bretz, F. (2009). *Computation of Multivariate Normal and t Probabilities*. Lecture Notes in Statistics. Springer.
- [Gerster and Herty, 2019] Gerster, S. and Herty, M. (2019). Discretized feedback control for systems of linearized hyperbolic balance laws. *Math. Control Relat. F.*, 9(3):517–539.
- [González-Gradón et al., 2017] González-Gradón, T., Heitsch, H., and Henrion, R. (2017). A joint model of probabilistic/robust constraints for gas transport management in stationary networks. *Comput. Manag. Sci.*, 14:443–460.
- [Gotzes et al., 2016] Gotzes, C., Heitsch, H., Henrion, R., and Schultz, R. (2016). On the quantification of nomination feasibility in stationary gas networks with random load. *Mathematical Methods of Operations Research*, 84:427–457.
- [Gramacki, 2018] Gramacki, A. (2018). *Nonparametric Kernel Density Estimation and Its Computational Aspects*. Springer International Publishing.
- [Gugat et al., 2015] Gugat, M., Hante, F. M., Hirsch-Dick, M., and Leugering, G. (2015). Stationary states in gas networks. *Networks and Heterogeneous Media*, 10(2):295–320.
- [Gugat and Herty, 2011] Gugat, M. and Herty, M. (2011). Existence of classical solutions and feedback stabilization for the flow in gas networks. *ESAIM Control Optim. Calc. Var.*, 17(1):28–51.
- [Gugat et al., 2012] Gugat, M., Herty, M., Klar, A., Leugering, G., and Schleper, V. (2012). Well-posedness of networked hyperbolic systems of balance laws. *Internat. Ser. Numer. Math.*, 160:123–146.
- [Gugat et al., 2011] Gugat, M., Hirsch-Dick, M., and Leugering, G. (2011). Gas flow in fan-shaped networks: Classical solutions and feedback stabilization. *SIAM J. Control Optim.*, 49(5):2101–2117.
- [Gugat et al., 2020] Gugat, M., Schultz, R., and Schuster, M. (2020). Convexity and starshapedness of feasible sets in flow networks. *Netw. Heterog. Media*, 15:171–195.
- [Gugat et al., 2018] Gugat, M., Schultz, R., and Wintergerst, D. (2018). Networks of pipelines for gas with nonconstant compressibility factor: Stationary states. *Comput. Appl. Math.*, 37:1066–1097.
- [Gugat and Schuster, 2018] Gugat, M. and Schuster, M. (2018). Stationary gas networks with compressor control and random loads: Optimization with probabilistic constraints. *Mathematical Problems in Engineering (Article ID 7984079)*.

- [Gugat and Ulbrich, 2017] Gugat, M. and Ulbrich, S. (2017). The isothermal Euler equations for ideal gas with source term: Product solutions, flow reversal and no blow up. *J. Math. Anal. Appl.*, 454:439–452.
- [Gugat and Ulbrich, 2018] Gugat, M. and Ulbrich, S. (2018). Lipschitz solutions of initial boundary value problems for balance laws. *Math. Models Methods Appl. Sci.*, 28(5):921–951.
- [Gugat and Wintergerst, 2018] Gugat, M. and Wintergerst, D. (2018). Transient flow in gas networks: Traveling waves. *Int. J. Appl. Math. Comput. Sci.*, 28(2):341–348.
- [Härdle and Müller, 1997] Härdle, W. and Müller, M. (1997). Multivariate and semi-parametric kernel regression. SFB 373 Discussion Papers 1997,26, Humboldt University of Berlin.
- [Härdle et al., 2004] Härdle, W., Werwatz, A., Müller, M., and Sperlich, S. (2004). *Nonparametric and Semiparametric Models*. Series in Statistics. Springer.
- [Hill, 2012] Hill, M. (2012). Convergence of random Fourier series. In *REU participant papers*.
- [Holle et al., 2020] Holle, Y., Herty, M., and Westdickenberg, M. (2020). New coupling conditions for isentropic flow on networks. Preprint on webpage <https://arxiv.org/abs/2004.09184>.
- [Imekraz et al., 2016] Imekraz, R., Robert, D., and Thomann, L. (2016). On random Hermite series. *Trans. Amer. Math. Soc.*, 368(4):2763–2792.
- [Kessy et al., 2018] Kessy, A., Lewin, A., and Strimmer, K. (2018). Optimal whitening and decorrelation. *The American Statistician*, 72(4):309–314.
- [Koch et al., 2015] Koch, T., Hiller, B., Pfetsch, M. E., and Schewe, L. (2015). *Evaluating Gas Network Capacities*. MOS-SIAM.
- [Lambert, 1758] Lambert, J. H. (1758). Observationes variae in mathesis puram. *Acta helv.*, 3:128–168.
- [Langrené and Warin, 2019] Langrené, N. and Warin, X. (2019). Fast and stable multivariate kernel density estimation by fast sum updating. *Journal of Computational and Graphical Statistics*, 28(3):596–608.
- [Lowe and Clarke, 2002] Lowe, C. and Clarke, J. (2002). A class of exact solutions for the Euler equations with sources: Part i. *Math. Comput. Modelling*, 36:275–291.
- [Lowe and Clarke, 2003] Lowe, C. and Clarke, J. (2003). A class of analytical solutions to the Euler equations with source terms: Part ii. *Math. Comput. Modelling*, 38:1101–1117.

- [Luskin, 1981] Luskin, M. (1981). On the existence of global smooth solutions for a model equation for fluid flow in a pipe. *J. Math. Anal. Appl.*, 84:614–630.
- [Marcus and Pisier, 1981] Marcus, M. B. and Pisier, G. (1981). *Random Fourier Series with Applications to Harmonic Analysis*. Princeton University Press.
- [Meleshko, 2015] Meleshko, S. V. (2015). *Methods for Constructing Exact Solutions of Partial Differential Equations*. Springer US, 1 edition.
- [Míos-Mercado and Borraz-Sánchez, 2015] Míos-Mercado, R. and Borraz-Sánchez, C. (2015). Optimization problems in natural gas transportation systems: A state-of-the-art review. *Applied Energy*, 147:536–555.
- [Osiadacz, 1987] Osiadacz, A. (1987). *Simulation and Analysis of Gas Networks*. E. & F.N. Spon.
- [Paley and Zygmund, 1930] Paley, R. E. A. C. and Zygmund, A. (1930). On some series of functions, (1). *Math. Proc. Cambridge Philos. Soc.*, 26:337–357.
- [Paley and Zygmund, 1932] Paley, R. E. A. C. and Zygmund, A. (1932). On some series of functions, (2). *Math. Proc. Cambridge Philos. Soc.*, 28:190–205.
- [Parzen, 1962] Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.*, 33(3):1065–1076.
- [Prékopa, 1995] Prékopa, A. (1995). *Stochastic Programming*. Springer, 1 edition.
- [Roald et al., 2020] Roald, L., Sundar, K., Zlotnik, A., Misra, S., and Andersson, G. (2020). An uncertainty management framework for integrated gas-electric energy systems. *Proceedings of the IEEE*, 108:1–23.
- [Römisch and Schultz, 1991a] Römisch, W. and Schultz, R. (1991a). Distribution sensitivity in stochastic programming. *Math. Program.*, 50:197–226.
- [Römisch and Schultz, 1991b] Römisch, W. and Schultz, R. (1991b). Stability analysis for stochastic programs. *Ann. Oper. Res.*, 30:241–266.
- [Römisch and Schultz, 1993] Römisch, W. and Schultz, R. (1993). Stability of solutions for stochastic programs with complete recourse. *Math. Oper. Res.*, 18:590–609.
- [Rose et al., 2016] Rose, D., Schmidt, M., Steinbach, M. C., and Willert, B. M. (2016). Computational optimization of gas compressor stations: MINLP models versus continuous reformulations. *Math. Methods Oper. Res.*, 83:409–444.
- [Schmidt et al., 2014] Schmidt, M., Steinbach, M., and Willert, B. (2014). High detail stationary optimization models for gas networks: Validation and results. *Optim. Eng.*, 16(1).

- [Schuster et al., 2021] Schuster, M., Strauch, E., Gugat, M., and Lang, J. (2021). Probabilistic constrained optimization on flow networks. *Optim. Eng. (accepted)*. Preprint on webpage <https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/searchtype/all/docId/325/start/0/rows/10>.
- [Shea, 1988] Shea, B. L. (1988). Algorithm as 239: Chi-squared and incomplete gamma integral. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 37(3):466–473.
- [Silverman, 1986] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Springer.
- [van Ackooij et al., 2018] van Ackooij, W., Aleksovska, I., and Munoz-Zuniga, M. (2018). (Sub-)differentiability of probabilistic functions with elliptical distributions. *Set-Valued Var. Anal.*, 26:887–910.
- [van Ackooij et al., 2016] van Ackooij, W., Frangioni, A., and Oliveira, W. (2016). Inexact stabilized Benders’ decomposition approaches: With application to chance-constrained problems with finite support. *Comput. Optim. Appl.*, 65:637–669.
- [van Ackooij and Henrion, 2014a] van Ackooij, W. and Henrion, R. (2014a). Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distribution. *SIAM J. Optim.*, 24:1864–1889.
- [van Ackooij and Henrion, 2014b] van Ackooij, W. and Henrion, R. (2014b). Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. *SIAM J. Optim.*, 24(4):1864–1889.
- [van Ackooij et al., 2020] van Ackooij, W., Henrion, R., and Pérez-Aros, P. (2020). Generalized gradients for probabilistic/robust (proburst) constraints. *Optimization*, 69:1451–1479.