Global Sensitivity Analysis of Ordinary Differential Equations

Adaptive Density Propagation Using Approximate Approximations

Dissertation zur Erlangung des Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich Mathematik & Informatik
der Freien Universität Berlin

vorgelegt von

Andrea Yeong Weiße

April 2009

1. Gutachter: Prof. Dr. Christof Schütte
2. Gutachter: Dr. Wilhelm Huisinga

Tag der Disputation: 1. Juli 2009

"Le doute n'est pas une condition agréable,
mais la certitude est absurde."

(Doubt is not a pleasant condition, but certainty is absurd.)

– Voltaire

Abstract

Ordinary differential equations play an important role in the modeling of many real-world processes. To guarantee reliable results, model design and analysis must account for uncertainty and/or variability in the model input. The propagation of uncertainty & variability through the model dynamics and their effect on the output is studied by sensitivity analysis. Global sensitivity analysis is concerned with variations in the model input that possibly span a large domain. Two major problems that complicate the analysis are high-dimensionality and quality control, i.e. controlling the approximation error of the estimated output uncertainty. Current numerical approaches to global sensitivity analysis mainly focus on scalability to high-dimensional models. However, to what extent the estimated output uncertainty approximates the true output uncertainty generally remains unclear.

In this thesis we suggest an error-controlled approach to global sensitivity analysis of ordinary differential equations. The approach exploits an equivalent formulation of the problem as a partial differential equation, which describes the evolution of the state uncertainty in terms of a probability density function. We combine recent advances from numerical analysis and approximation theory to solve this partial differential equation. The method automatically controls the approximation error by adapting both temporal and spatial discretization of the numerical solution. Error control is realized using a Rothe method that provides a framework for estimating temporal and spatial errors such that the discretization can be adapted accordingly. We use a novel technique called approximate approximations for the spatial discretization; it is the first time that these are used in the context of an adaptive Rothe scheme.

We analyze the convergence of the method and investigate the performance of approximate approximations in the adaptive scheme. The method is shown to converge, and the theoretical results directly indicate how to design an efficient implementation. Numerical examples illustrate the theoretical results and show that the method yields highly accurate estimates of the true output uncertainty. Furthermore, approximate approximations have favorable properties in terms of readily available error estimates and high approximation order at feasible computational costs. Recent advances in the theory of approximate approximations, based on a meshfree discretization of the state space, promise that the applicability of the adaptive density propagation framework developed herein can be extended to higher-dimensional problems.

Acknowledgments

I would like to thank my advisors Wilhelm Huisinga and Christof Schütte for giving me the chance of taking part in this research project. This thesis owes a lot to Wilhelm's patience to sit down together with me and calmly work out the maths that often seemed so obscure. Whenever drowned in the details, Christof's "three keypoint policy" was helpful in digging the big picture out of the confusion. My gratitude also goes to Ralf Kornhuber for drawing my attention to the work by Bornemann and, not least, for always having encouraging words, as well as to Michael Wulkow for helping me during the early stages of this project.

In the course of this research I had the great opportunity to spend time in both the International Max Planck Research School (Berlin) and the Hamilton Institute (Ireland). I thank Hannes Luz and the people from the IMPRS for generating such a nice working and social environment, and the people from the Hamilton Institute who warmly welcomed me as one of them. Evelyn Dittmer deserves special recognition for her lovely company during my time in Berlin; I truly appreciate that in those two years we shared an office she never seemed to mind my many distractive habits. I am also very grateful to Stephan Menz, Philipp Metzner, Oliver Mason and Utz Pape for reading parts of this thesis and providing valuable comments and suggestions. My friend Nathalie Véron, too, was brave enough to be my guinea pig, reading parts of the thesis and figuring out whether they made any sense to the experimentalist eye; she deserves my sincere gratitude for that.

I am deeply thankful to my family, my parents, my sister and my brother, as well as my dear friend Sharif. They always thought that they could not support me, but they did. In particular, I would like to mention my mother. This little woman has put so many efforts into making the three of us appreciate education as the privilege it is. She told us the story about her own mother, my grandmother, who had suffered all her life from the fact that she, as a woman, did not have the chance to attend school, but had to sneak into her children's school books to merely ease her curiosity. This work pays tribute to such admirable spirit.

And finally, Diego, how can I thank you for all you did? It is thanks to you that the last months, despite everything, have not been a disaster. With joy I am awaiting the upcoming months, when I can try my best to support you in the same way.

Andrea Weiße Berlin, April 2009

Contents

Introduction

Mathematical modeling is a key tool for the analysis of a wide range of real-world phenomena ranging from physics and engineering to chemistry, biology and economics [50, 43, 82]. The recently growing influence of modeling in the analysis of biological processes [83] poses challenging mathematical problems. Among the different modeling approaches, *ordinary differential equations* (ODE) are particularly important and have led to significant advances [6, 16]. Ordinary differential equations model the temporal evolution of the relevant variables by describing their *deterministic* dynamics. The study of dynamical systems with ODEs is a mature field and therefore, there is a rich literature devoted to their analysis [2, 20] and solution [30, 31, 25].

ODEs are used to model biological processes on various levels ranging from gene expression [23, 24] or signaling processes on the cellular level [35] to the kinetics of drugs on the whole-body level [87]. All these processes have in common that their modeling with ODEs bears a considerable degree of uncertainty and/or variability in both initial conditions and parameters [4, 17, 52]. This is particularly the case when models are considered in a population-wide context. Then, uncertainty commonly corresponds to noisy measurements or the lack of knowledge about individual systems, whereas variability refers to variations over time in individual systems or within the population [5, 7]. The propagation of uncertainty and variability through the system dynamics can lead to considerable variations in the model outputs, see Figure 1, and neglecting this may lead to unreliable conclusions.

The systematic study of how uncertainty and variability affect the model outputs is called *sensitivity analysis* and is a crucial step of any practical modeling approach [13, 17]. Sensitivity analysis of ODEs can be addressed from different mathematical perspectives, which give access to different numerical methods. The advantages and disadvantages of those motivated the development of a novel approach, which is presented in this thesis.

Sensitivity analysis of ODEs Depending on the problem under study, the uncertainty and variability of an ODE model may affect initial values, the parameters, or both. These will be referred to as the *model input*.

In many cases, uncertainty can be regarded as small variations, or perturbations, around reference input values, while variability generally refers to larger variations. Effects of small variations are often studied using a local approach. *Local sensitivity analysis* is based on linearized solutions of the ODE around a reference input values. Linearization facilitates the analysis of the problem considerably. It involves the computation of partial derivatives of the ODE with respect to the uncertain input variables, so called sensitivity indices, which describe the variance of the output uncertainty [74, 86]. The two terms, local and linear sensitivity analysis, are often used interchangeably [25, 86].

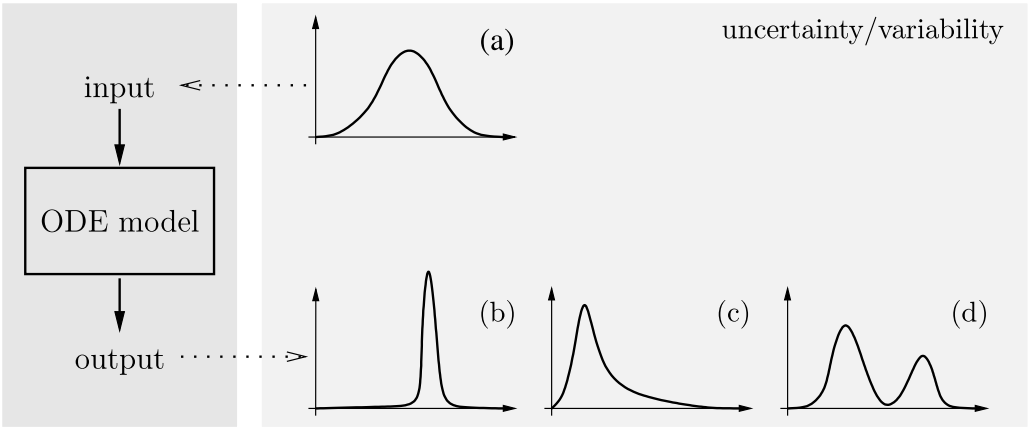

Figure 1: Sensitivity analysis: (a) uncertainty in the model input propagates through the model dynamics and yields an uncertain model output. Three scenarios of output uncertainty are shown: (b) small output uncertainty, the model is not very sensitive to the input; (c) & (d) input uncertainty seriously affects the output uncertainty, the model is sensitive, and simple descriptors such as mean or variance fail to capture the structure of the output uncertainty.

The linear approach provides a good estimate of the true sensitivity only when variations are small, or when the model dynamics are linear. In the case of larger variations, the sensitivity of a nonlinear model should therefore be studied globally. *Global sensitivity analysis* commonly considers the input values as random variables with a given probability distribution. The problem can then be transformed to a system of *ODEs with random initial values*. By extending the state space to include the model parameters, this approach can account for variations in initial values and parameters within a single framework. A straightforward approach is to solve this system for a set of sampled input values. An estimate of the sensitivity of the model can then be obtained from the outputs produced with each of the sampled values. Sampling-based approaches are called Monte Carlo (MC) methods [32, 61, 77, 78] and are widely used for sensitivity analysis of ODEs [36, 62].

Based on the probability density function of the random initial values, the problem can be recast as a *density propagation* problem. The evolution of the density function is described by a first-order linear *partial differential equation* (PDE). Costanza & Seinfeld [19] first proposed to perform sensitivity analysis by numerically solving this PDE, and since then the approach is often referred to as stochastic sensitivity analysis [74, 86]. The numerical analysis of PDEs is a broad field of ongoing research and an extensive literature is dedicated to it [29, 73, 84]. Therefore, the density propagation approach gives access to a rich theory and methodology that facilitate a highly accurate estimation of the output uncertainty.

Limitations of existing approaches In most applications, ODE models describe nonlinear dynamics, so that analytical solutions are generally not available and one must resort to numerical methods. In addition, many applications require sensitivity analysis with respect to numerous input variables, and thus the numerical methods have to deal with *high-dimensions*.

Local sensitivity methods can cope with high-dimensionality comparatively well, but they are limited to problems with small input variations. In the case of global sensitivity analysis,

the input space (and thus the complexity of the problem) grows exponentially with the number of dimensions. This limitation is common to all numerical methods. One way to circumvent this problem is to study the sensitivity of each input variable separately. This approach requires moderate effort but does not reflect the true nature of the problem, since correlations between variables are lost.

Currently, most global sensitivity methods that can be applied to high-dimensional problems are based on an MC method [36, and references therein]. As for all numerical methods, the discretization (for MC methods the representation of the uncertainty by means of a discrete sample) inherently carries on approximation errors of the estimated output uncertainty. In the random setting, estimates of these errors are generally hard to obtain [21]. It thus remains a problem to judge if the sensitivity has been analyzed with sufficient accuracy.

The density propagation approach facilitates accuracy control, since the numerical analysis of PDEs provides methodology that is specially tailored for an error-controlled, or adaptive solution, see e.g. [29]. In our view, *adaptive density propagation* promises the most accurate estimates of the sensitivity of ODE models. However, this strategy is generally limited to low-dimensional problems [74, 86], since most PDE methods become inefficient for dimensions higher than two or three [12, 28].

Objective of this work MC methods continue to be fundamental for the global sensitivity analysis of high-dimensional ODE models. Many practical studies could however benefit from a global method that, at high accuracy, can be applied to medium-dimensional models. This thesis intends to provide a theoretical framework to address such global sensitivity analysis problems. In this work an adaptive density propagation method is developed. The method allows us to control both temporal and spatial errors via an adaptation of the discretization. This is implemented by combining recent results from the fields of numerical analysis and approximation theory.

Numerical analysis offers different approaches to the solution of PDEs. Among these, the *Rothe approach* [75, 76] is particularly important in terms of adaptivity, see e.g. [29]. Rothe methods are based on a temporal semi-discretization of the PDE. This results in stationary spatial problems, which can be solved using approximation methods.

The proposed method is based on a *Rothe scheme with multiplicative error correction* that was introduced by Bornemann [8, 9] for the solution of parabolic PDEs. Multiplicative error correction aims at improved temporal adaptivity by avoiding numerical cancellation in the temporal error estimates. This approach also allows for a separate estimation of temporal and spatial errors, which in turn provides the basis for the decision when to refine the temporal or the spatial discretization.

A novel approximation method called *approximate approximations* is used for the solution of the stationary spatial problems. This method was developed by Maz'ya & Schmidt [64, 63, 66] and has successfully been employed for the solution of elliptic and time-dependent PDEs mostly of order two or higher [48, 68, 81]. Favorable analytical properties with respect to the approximation of differential operators [65], together with a sound convergence theory [66] make approximate approximations attractive for their use in an adaptive Rothe context. To our knowledge, they have not been used for this purpose so far.

The analysis shown in this work reveals that, to guarantee convergence of the overall numerical scheme, dependencies between the spatial and temporal discretization have to be taken into account. These impose high accuracy constraints on the spatial discretization. Approximate approximations prove more favorable in this respect as compared to classical spatial discretization methods, because they allow these dependencies to be efficiently resolved.

To fully understand the numerical aspects of the suggested approach, we confine the analysis to approximate approximations on uniform grids. In practice, this implies that the method, in its current shape, is restricted to low-dimensional problems. We intend to establish a theoretical foundation for adaptive Rothe methods with approximate approximations in the context of global sensitivity analysis of ODEs. At a later stage, we plan to combine the methodology with approximate approximations on non-uniform or scattered grids in order to extend its applicability to problems with higher dimensions. There is ongoing research on approximate approximations with scattered grids [27, 44, 54]. A combination of the framework presented herein with approximate approximations on scattered grids may provide a powerful tool for the global sensitivity analysis of ODEs with moderate input dimensions.

Thesis overview In Part I, the mathematical setting is presented: ODEs with random initial values are discussed in Chapter 1, and the two equivalent approaches to solving this problem—the ODE-based and the PDE-based approach—are presented. Chapter 2 then gives an overview on existing numerical methods to address the problem. We focus on the conceptual frameworks of those methods rather than discussing their differences in detail.

Part II constitutes the main part of the thesis. The Rothe approach with multiplicative error correction is described in Chapter 3, and Chapter 4 gives an introduction to approximate approximations. Then, in Chapter 5, we propose an algorithm for adaptive density propagation that combines the Rothe method with approximate approximations. The convergence of the suggested method is analyzed in Chapter 6, and numerical examples are presented in Chapter 7.

Finally, in Part III, we conclude by summarizing the results obtained in this work and discussing possible extensions of the method and applications to other problems.

Some technical material has been allocated to the Appendix. This includes basic concepts for semi-discretization of PDEs in time (Part A), derivations of spatial error estimates within the multiplicative error correction (Part B) as well as formulas of derivatives of the basis functions of approximate approximations (Part C).

Part I

Preliminaries

Chapter 1

Ordinary differential equations with random initial values

The objective of sensitivity analysis is to study the effect of uncertainty and variability on the model output. In this thesis we are interested in global sensitivity analysis. We assume that the uncertainty or variability is specified in terms of a probability distribution of the input variables under study. Considering parameter variables as part of the state space allows us to formulate global sensitivity analysis as the solution of ODEs with random initial values.

In this chapter we present the mathematical setting for ODEs with random initial values together with their solution. In Section 1.1, the problem is formally stated, and Section 1.2 describes its solution using the theory of Frobenius-Perron operators. This perspective allows us to derive an equivalent characterization of the solution by means of a first-order linear PDE.

1.1 Problem statement

We are interested in problems where the state $z \in \mathbb{R}^n$ of the system can be described by an ordinary differential equation of the form

$$\dot{z} = f(z \,|\, p)\,, \quad \text{with} \quad z(0) = z_0\,. \tag{1.1}$$

The right hand side $f(\cdot|p) : \mathbb{R}^n \to \mathbb{R}^n$ may depend on parameters $p \in \mathbb{R}^m$. Since we are interested in a sensitivity analysis with respect to a model input consisting of both initial conditions z_0 and parameters p, we consider the extended state variable $x := \begin{pmatrix} z & p \end{pmatrix}^T \in \mathbb{R}^d$, with $d = n + m$. This allows us to study the effects of variations in z_0 and p simultaneously by setting

$$\dot{x} = F(x) := \begin{pmatrix} f(z \,|\, p) \\ 0 \end{pmatrix}, \quad \text{with} \quad x(0) = x_0 = \begin{pmatrix} z_0 \\ p \end{pmatrix}. \tag{1.2}$$

Let $|\cdot|$ denote a vector norm on $\mathbb{R}^d$ (e.g. the Euclidean norm). Then, the following theorem gives conditions for the existence and uniqueness of a solution $x(t)$, $t \geq 0$.

Theorem 1.1.1 (**Existence Theorem of Picard-Lindelöf**, [25, Theorem 2.7])**.** *Let F be locally Lipschitz continuous, i.e., there exists $L \geq 0$ such that*

$$|F(x) - F(y)| \leq L \cdot |x - y|\,, \qquad \forall\, x \in \mathbb{R}^d,\ y \in B_\kappa(x),$$

where $B_\kappa(x) := \{y \in \mathbb{R}^d, |y - x|_2 \leq \kappa\}$ *denotes an open neighborhood around* x*. Then, the initial value problem* (1.2) *has a* unique *solution* $x(t)$*,* $t \geq 0$*.*

A sufficient condition for local Lipschitz continuity is continuous differentiability of F with respect to the state variable x, which will be assumed henceforth. Let us denote the evolution operator $\Phi_t : \mathbb{R}^d \to \mathbb{R}^d$ with

$$\Phi_t x_0 := x(t)\,, \tag{1.3}$$

which maps an initial state x_0 to its state at time t. The evolution operator has the following properties:

(*i*) $\Phi_0 x = x$ for all $x \in \mathbb{R}^d$,

(*ii*) $\Phi_t(\Phi_{t'} x) = \Phi_{t+t'} x$ for all $x \in \mathbb{R}^d$ and $t, t' \in \mathbb{R}$,

(*iii*) $\Phi_t x$ is differentiable with respect to x for all $t \in \mathbb{R}$.

Note that by the first two properties, $\{\Phi_t\}_{t\in\mathbb{R}}$ forms a group, and therefore Φ_t is invertible with $\Phi_t^{-1} = \Phi_{-t}$.

To mathematically characterize the uncertainty or variability in initial values, we assume that $x_0 = X_0$ is a *random variable*. Consequently, $\Phi_t x_0 = X_t$ is also a random variable and $\{X_t\}_{t\geq 0}$ a stochastic process. For any $t \geq 0$, let us denote with $u_t = u(t, \cdot)$, $u : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, the probability density function of the probability distribution of X_t, i.e.

$$\mathbb{P}[X_t \leq x] = \int_{-\infty}^{x} u_t(s)\ \mathrm{d}s\,. \tag{1.4}$$

The objective is to solve the following problem:

Problem 1.1.2 (**Random Initial Value Problem**)**.** *Let the system be described by an ODE of the form*

$$\dot{x} = F(x)\,.$$

Assume the initial value $x_0 = X_0$ *is a random variable and has a known probability distribution with density* u_0*. The problem is to compute the probability density function* u_t *associated with the random state* $x(t) = X_t$ *on a finite interval* $t \in [0, T]$*.*

1.2 Analytical solution

In the following we consider the solution of the Random Initial Value Problem 1.1.2. In Section 1.2.1, the temporal evolution of the probability density function is studied using the theory of Frobenius-Perron operators. The interpretation of Frobenius-Perron operators as a semigroup allows us then, in Section 1.2.2, to derive an equivalent formulation of the Random Initial Value Problem in terms of a first-order linear PDE. Finally, in Section 1.2.3, we show how pointwise solutions to this PDE can be obtained using the method of characteristics.

1.2.1 Evolution of density functions: the Frobenius-Perron operator

Below we give a brief introduction to the theory of Frobenius-Perron operators along with some required background on measure theory. A comprehensive treatment can be found in [55].

Measures, measure spaces and $\mathcal{L}_p$-spaces

As a matter of convention, let us denote the state space $\mathbb{R}^d$ by Ω. A set $\mathcal{B}$ containing subsets of Ω is called a *σ-algebra* if

(i) $\Omega \in \mathcal{B}$,

(ii) $B \in \mathcal{B} \quad \Rightarrow \quad \Omega \backslash B \ \in \mathcal{B}$,

(iii) for any countable collection $\{B_k\}_{k=1,2,\dots}$ of subsets of Ω: $B_k \in \mathcal{B} \Rightarrow \bigcup_k B_k \in \mathcal{B}$.

The σ-algebra generated by all closed intervals $[a,b] \subset \mathbb{R}^d$, $a,b \in \mathbb{R}^d$, is called the *Borel σ-algebra*. A real-valued function $\mu : \mathcal{B} \to \mathbb{R}$ is called a *measure* on Ω if

(i) $\mu(\emptyset) = 0$,

(ii) $\mu(B) \geq 0$ for all $B \in \mathcal{B}$,

(iii) for all countable sets $\{B_k\}_{k=1,2,\dots}$ of pairwise disjoint $B_k \in \mathcal{B}$: $\mu(\bigcup_k B_k) = \sum_k \mu(B_k)$,

and all sets $B \in \mathcal{B}$ are called *measurable sets*. We are particularly interested in probability measures $\mu = \mathbb{P}$, i.e. $\mu(\Omega) = 1$. The triple $(\Omega, \mathcal{B}, \mu)$ is called a *measure space*, and a *probability space* in case $\mu = \mathbb{P}$. Any function $u : \Omega \to \hat{\Omega}$ is called *measurable* in $(\Omega, \mathcal{B}, \mu)$, if for all $\hat{B} \subset \hat{\Omega}$ the pre-image of u

$$u^{-1}(\hat{B}) := \left\{ \omega \in \Omega,\ u(\omega) \in \hat{B} \right\}$$

is a measurable set, i.e. $u^{-1}(\hat{B}) \in \mathcal{B}$. We are specifically interested in real-valued measurable functions $u : \Omega \to \mathbb{R}$. For these,

$$||u||_{\mathcal{L}_p} := \left(\int_{x \in \Omega} |u(x)|^p \ \mu(\mathrm{d}x) \right)^{1/p} \quad \text{for } 1 \leq p < \infty \quad \text{and} \quad ||u||_{\mathcal{L}_\infty} := \sup_{x \in \Omega} |u(x)| \tag{1.5}$$

defines a norm on $(\Omega, \mathcal{B}, \mu)$, which is called the *$\mathcal{L}_p$-norm*. Moreover, the set of all functions u for which $||u||_{\mathcal{L}_p}$ is finite is called the *$\mathcal{L}_p(\Omega, \mathcal{B}, \mu)$-space*.

Remark 1.2.1. *Throughout this work, $||\cdot||$ will denote the $\mathcal{L}_p$-norm unless stated otherwise. We will also write $\mathcal{L}_p$ instead of $\mathcal{L}_p(\Omega, \mathcal{B}, \mu)$ whenever the measure space is clear from the context, or sometimes $\mathcal{L}_p(\Omega)$ when $\Omega \subset \mathbb{R}^d$ denotes a sub-domain of $\mathbb{R}^d$.*

This setting now allows us to define densities and probability density functions: Any positive function $u \in \mathcal{L}_1$ with $||u||_{\mathcal{L}_1} = 1$ is called a *density*, and furthermore u is called a density of the measure μ_u, if

$$\mu_u(B) = \int_B u(x) \ \mu(\mathrm{d}x)\,. \tag{1.6}$$

If additionally $\mu_u = \mathbb{P}$ is a probability measure, u is called a *probability density function*. For a Borel σ-algebra, the *Borel measure* that assigns to each interval its length (or hyper-area, if $d > 1$), uniquely defines a measure. We then write $\mathrm{d}x$ instead of $\mu(\mathrm{d}x)$.

A measurable transformation $\phi : \Omega \to \Omega$ is called *nonsingular*, if for all $B \in \mathcal{B}$

$$\mu(\phi^{-1}(B)) = 0 \quad \Rightarrow \quad \mu(B) = 0\,, \tag{1.7}$$

which means that only nullsets (with measure zero) can be mapped to nullsets.

Transformations of measures & densities

To study the solution to the Random Initial Value Problem 1.1.2, we consider the measure space $(\Omega, \mathcal{B}, \mu)$, where the state space is $\Omega = \mathbb{R}^d$, $\mathcal{B}$ denotes the Borel σ-algebra and μ the Borel measure. For a fixed time $t \geq 0$, the evolution operator Φ_t of the ODE denotes a transformation on the state space. The transformation of Ω through Φ_t causes a change in the probability distribution on Ω. The probability of a set B at time t must equal the probability of its pre-image $\Phi_t^{-1}(B)$, i.e.

$$\mathbb{P}[X_t \in B] = \int_B u_t(x)\ \mathrm{d}x = \int_{\Phi_t^{-1}(B)} u_0(x)\ \mathrm{d}x = \mathbb{P}[X_0 \in \Phi_t^{-1}(B)]\,. \tag{1.8}$$

Thus, the transformation of the probability distribution can be quantified by means of the density functions u_0 and u_t. Frobenius-Perron operators establish a functional relation between the initial and the transformed density by

$$\mathcal{P}_t u_0 = u_t\,, \tag{1.9}$$

where $\mathcal{P}_t$ is called the *Frobenius-Perron operator corresponding to* Φ_t. Similar to the evolution operator Φ_t, the Frobenius-Perron operator maps any initial density u_0 to its transformed version u_t at time t. Therefore, it describes the evolution of the probability density function associated with the random state X_t, as illustrated in Figure 1.1.

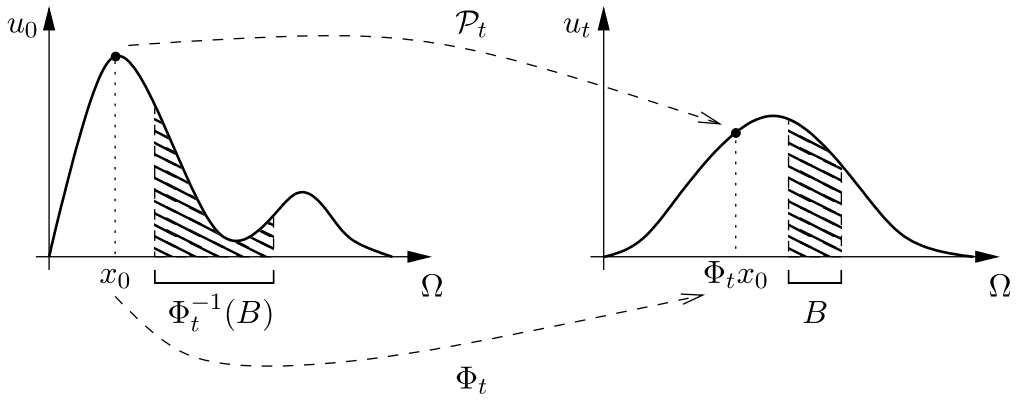

Figure 1.1: Shown are the two probability density functions u_0 and u_t. The conservation of probability mass on any set B and its pre-image $\Phi_t^{-1}(B)$ defines the Frobenius-Perron operator $\mathcal{P}_t$ corresponding to Φ_t, which relates the two density functions to each other.

The general definition of Frobenius-Perron operators follows from relation (1.8):

Definition 1.2.2 (**Frobenius-Perron operator**). *Let* $(\Omega, \mathcal{B}, \mu)$ *be a measure space and* $\Phi_t : \Omega \to \Omega$ *a nonsingular transformation. The Frobenius-Perron operator* $\mathcal{P}_t : \mathcal{L}_1 \to \mathcal{L}_1$ *corresponding to* Φ_t *is defined by*

$$\int_B \mathcal{P}_t u(x)\ \mu(\mathrm{d}x) = \int_{\Phi_t^{-1}(B)} u(x)\ \mu(\mathrm{d}x)\ , \quad \forall\ B \in \mathcal{B} \text{ and } u \in \mathcal{L}_1\,. \tag{1.10}$$

Nonsingularity of Φ_t ensures that the Frobenius-Perron operator is uniquely defined by (1.10) and follows from invertibility of Φ_t. Definition 1.2.2 further implies that $\mathcal{P}_t$ has the following properties:

(*i*) $\mathcal{P}_t$ is linear.

(*ii*) $u(x) \geq 0\,,\ \forall x \in \Omega \quad \Rightarrow \quad \mathcal{P}_t u(x) \geq 0\,,\ \forall x \in \Omega.$

(*iii*) $||\mathcal{P}_t u||_{\mathcal{L}_1} = ||u||_{\mathcal{L}_1}\ ,\ \forall u \in \mathcal{L}_1.$

(*iv*) For the concatenation $\Phi_t^n = \Phi_t \circ \overset{n}{\ldots} \circ \Phi_t$, the corresponding Frobenius-Perron operator is $\mathcal{P}_n = \mathcal{P}_t^n$.

Let us mention that by properties (*ii*) and (*iii*), $\mathcal{P}_t$ is a *Markov operator*, and thus if u_0 is a probability density function, then $u_t = \mathcal{P}_t u_0$ is as well. Since the evolution operator Φ_t is differentiable and invertible, an explicit form of $\mathcal{P}_t$ can be obtained. To illustrate this, let us consider $\Omega = \mathbb{R}$. Then for an interval $B = [a, x]$, relation (1.10) becomes

$$\int_a^x \mathcal{P}_t u_0(s)\ \mathrm{d}s = \int_{\Phi_t^{-1}([a,x])} u_0(s)\ \mathrm{d}s\,,$$

and by differentiation

$$\mathcal{P}_t u_0(x) = \frac{\mathrm{d}}{\mathrm{d}x} \int_{\Phi_t^{-1}([a,x])} u_0(s)\ \mathrm{d}s\,.$$

By the differentiability and invertibility of Φ_t, it follows that the evolution operator is monotone. Let us assume Φ_t is monotonically increasing, hence $\Phi_t^{-1}([a, x]) = [\Phi_{-t}a, \Phi_{-t}x]$, and we get

$$\begin{aligned} \mathcal{P}_t u_0(x) &= \frac{\mathrm{d}}{\mathrm{d}x} \int_{\Phi_{-t}a}^{\Phi_{-t}x} u_0(s)\ \mathrm{d}s \\ &= u_0(\Phi_{-t}x) \cdot \frac{\mathrm{d}}{\mathrm{d}x}\left(\Phi_{-t}x\right)\ . \end{aligned} \tag{1.11}$$

In [55, Chapter 3] it is shown that for $\Omega = \mathbb{R}^d$, (1.11) generalizes to

$$u_t(x) = \mathcal{P}_t u_0(x) = u_0\left(\Phi_{-t}x\right) \cdot \left|\frac{\mathrm{d}}{\mathrm{d}x}\left(\Phi_{-t}x\right)\right|\,, \tag{1.12}$$

where $\left|\frac{\mathrm{d}}{\mathrm{d}x}\left(\Phi_{-t}x\right)\right| := \det\left(\frac{\mathrm{d}}{\mathrm{d}x}\left(\Phi_{-t}x\right)\right)$, and $\frac{\mathrm{d}}{\mathrm{d}x}\left(\Phi_{-t}x\right)$ denotes the Jacobian of Φ_{-t}. Since Φ_t is invertible, we can rewrite (1.12) as

$$u_t(\Phi_t x) = \mathcal{P}_t u_0\left(\Phi_t x\right) = u_0(x) \cdot \left|\frac{\mathrm{d}}{\mathrm{d}x}\left(\Phi_t x\right)\right|^{-1}\ . \tag{1.13}$$

This means that the density u_t evaluated at the propagated point $x(t) = \Phi_t x_0$ differs from the initial density at the original point $x(0) = x_0$ by the factor $\left|\frac{\mathrm{d}}{\mathrm{d}x}(\Phi_t x)\right|^{-1}$. This factor accounts for local contractions or expansions of the evolution Φ_t. (Hamiltonian systems, for example, imply that $\left|\frac{\mathrm{d}}{\mathrm{d}x}(\Phi_t x)\right| = 1$ for all $x \in \Omega$ and thus, for all $t \geq 0$ the density remains constant along a trajectory $x(t) = \Phi_t x$.)

We next exemplify how in special cases the pointwise information from (1.12) can be used to obtain an explicit global solution.

Example 1.2.3 (**Evolution of a Normal distribution under linear dynamics**). *Assume that the ODE is linear with right hand side $F(x) = Ax$, $A \in \mathbb{R}^{d\times d}$. Then the evolution of the ODE with random initial value $x_0 = X_0$ is given by*

$$\Phi_t x_0 = e^{tA}\, x_0\,,$$

see e.g. [2], and its inverse by

$$\Phi_{-t} x_0 = e^{-tA}\, x_0\ ,$$

where e^{tA} denotes the matrix exponential.

Let the initial probability density function u_0 associated with X_0 be the density of a Normal distribution with mean μ_0 and covariance matrix Σ_0, i.e.

$$u_0(x) = \frac{1}{\sqrt{(2\pi)^d|\Sigma_0|}} \cdot \exp\left(-\frac{1}{2}(x-\mu_0)^T\Sigma_0^{-1}(x-\mu_0)\right)\,, \quad |\Sigma_0| := \det(\Sigma_0).$$

Applying the explicit formula for the Frobenius-Perron operator (1.12) *yields*

$$\begin{aligned} u_t(x) &= \frac{1}{\sqrt{(2\pi)^d|\Sigma_0|}} \cdot \exp\left(-\frac{1}{2}(e^{-tA}x-\mu_0)^T\Sigma_0^{-1}(e^{-tA}x-\mu_0)\right) \cdot \left|\frac{\mathrm{d}}{\mathrm{d}x}e^{-tA}x\right| \\ &= \frac{1}{\sqrt{(2\pi)^d|\Sigma_0||e^{tA}|^2}} \cdot \exp\left(-\frac{1}{2}(x-e^{tA}\mu_0)^T(e^{-tA})^T\Sigma_0^{-1}e^{-tA}(x-e^{tA}\mu_0)\right)\,, \end{aligned} \tag{1.14}$$

which is identical to the probability density function of the Normal distribution $\mathcal{N}(\mu_t, \Sigma_t)$ with parameters

$$\mu_t = e^{tA}\cdot\mu_0\,, \quad \textit{and} \quad \Sigma_t = (e^{tA})^T\cdot\Sigma_0\cdot e^{tA}\ . \tag{1.15}$$

□

In the above example it is seen that Gaussian densities are invariant to linear transformations. We will see in Chapter 2 that this property is used by some numerical methods to obtain a linearized estimate of the sensitivity of a model. In the following we will use the Frobenius-Perron operator to derive an equation that describes the temporal evolution of the probability density function u_t.

1.2.2 Equivalent formulation in terms of a PDE

So far we have considered the transformation of the initial density for a fixed time t. Using the properties of the evolution operator Φ_t, it can be shown that the family of Frobenius-Perron operators $\{P_t\}_{t\geq 0}$ forms a *semigroup*, i.e.

(*i*) $\mathcal{P}_0 u = u\,, \quad \forall\, u \in \mathcal{L}_1$ and

(*ii*) $\mathcal{P}_t(\mathcal{P}_{t'} u) = \mathcal{P}_{t+t'} u\,, \quad \forall\, u \in \mathcal{L}_1$ and $t, t' \geq 0$.

To derive a differential equation that describes the evolution of $u_t = \mathcal{P}_t u_0$ for $t \geq 0$, we will now consider a closely related operator, the *Koopman operator*.

Definition 1.2.4 (**Koopman operator**). *Let $(\Omega, \mathcal{B}, \mu)$ be a measure space and $\Phi_t : \Omega \to \Omega$ a nonsingular transformation. The operator $\mathcal{K}_t : \mathcal{L}_\infty \to \mathcal{L}_\infty$ defined by*

$$\mathcal{K}_t v(x) = v(\Phi_t x)\,, \quad \forall\, v \in \mathcal{L}_\infty\,, \tag{1.16}$$

is called the Koopman operator *with respect to Φ_t.*

It can be shown, see [55, Chapter 3], that the Koopman operator has the properties:

(*i*) $\mathcal{K}_t$ is linear.

(*ii*) $\mathcal{K}_t$ is a contraction on $\mathcal{L}_\infty$, i.e. $||\mathcal{K}_t v||_{\mathcal{L}_\infty} \leq ||v||_{\mathcal{L}_\infty}\,,\ \forall v \in \mathcal{L}_\infty$.

(*iii*) $\mathcal{K}_t$ is the adjoint operator of the Frobenius-Perron operator corresponding to Φ_t, i.e.

$$\langle \mathcal{P}_t u, v\rangle = \int_\Omega \mathcal{P}_t u(x) \cdot v(x)\,\mathrm{d}x = \int_\Omega u(x) \cdot \mathcal{K}_t v(x)\,\mathrm{d}x = \langle u, \mathcal{K}_t v\rangle\,, \qquad \forall\, u \in \mathcal{L}_1\,,\ v \in \mathcal{L}_\infty\,,$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product.

Furthermore, the family of Koopman operators $\{\mathcal{K}_t\}_{t\geq 0}$ with respect to $\{\Phi_t\}_{t\geq 0}$ forms a semigroup. We next derive a differential representation of $\mathcal{K}_t$ with respect to t. Assume $v \in \mathcal{L}_\infty$ is continuously differentiable and has compact support. By definition, the Koopman operator satisfies

$$\frac{\mathcal{K}_t v(x_0) - v(x_0)}{t} = \frac{v(\Phi_t x_0) - v(x_0)}{t} = \frac{v(x(t)) - v(x_0)}{t}\,.$$

Since v is continuously differentiable and has compact support, the mean value theorem yields

$$\frac{\mathcal{K}_t v(x_0) - v(x_0)}{t} = \sum_{i=1}^{d} \dot{x}_i(\theta t) \cdot v_{x_i}(x(\theta t)) = \sum_{i=1}^{d} F_i(\theta t) \cdot v_{x_i}(x(\theta t))\,, \qquad 0 < \theta < 1\,,$$

where v_{x_i} denotes the partial derivative of v with respect to x_i. Since v, and thus v_{x_i}, has compact support, the limit for $t \to 0$ exists and is given by

$$\lim_{t\to 0} \frac{\mathcal{K}_t v(x_0) - v(x_0)}{t} = \lim_{t\to 0} \left(\sum_{i=1}^{d} F_i\left(\theta t\right) \cdot v_{x_i}\left(x\left(\theta t\right)\right) \right) = \sum_{i=1}^{d} F_i(x_0) \cdot v_{x_i}(x_0)\,.$$

The differential operator defined by

$$\mathcal{A}_{\mathcal{K}} v(x) := \sum_{i=1}^{d} F_i(x) \cdot \frac{\partial}{\partial x_i} v(x)\,,$$

is called the *infinitesimal generator* of the semigroup of Koopman operators $\{\mathcal{K}_t\}_{t\geq 0}$ with respect to $\{\Phi_t\}_{t\geq 0}$. An explicit form of the infinitesimal generator of the semigroup of Frobenius-Perron operators, defined by

$$\mathcal{A} u(x) := \lim_{t\to 0} \frac{\mathcal{P}_t u(x) - u_0(x)}{t}\,,$$

can be derived using that the Koopman operator is the adjoint operator, i.e.

$$\langle \mathcal{P}_t u, v\rangle = \langle u, \mathcal{K}_t v\rangle\,, \qquad u \in \mathcal{L}_1\,,\ v \in \mathcal{L}_\infty\,.$$

Subtracting $\langle u, v\rangle$ from both sides and dividing by t yields

$$\left\langle \frac{\mathcal{P}_t u - u}{t}, v \right\rangle = \left\langle u, \frac{\mathcal{K}_t v - v}{t} \right\rangle\,.$$

For functions $u \in D(\mathcal{A})$ and $v \in D(\mathcal{A}_{\mathcal{K}})$ in the domains of $\mathcal{A}$ and $\mathcal{A}_{\mathcal{K}}$, taking the limit as $t \to 0$ further yields the relation

$$\langle \mathcal{A} u, v\rangle = \langle u, \mathcal{A}_{\mathcal{K}} v\rangle\,, \tag{1.17}$$

which, using the explicit form of $\mathcal{A}_{\mathcal{K}}$, can be written as

$$\langle \mathcal{A} u, v\rangle = \left\langle u, \sum_{i=1}^{d} \frac{\partial}{\partial x_i} v \cdot F_i \right\rangle = \sum_{i=1}^{d} \int_{\mathbb{R}^d} \left(\frac{\partial (u F_i v)}{\partial x_i} - v \cdot \frac{\partial (u F_i)}{\partial x_i} \right)\, \mathrm{d}x\,.$$

If v has compact support, then by the divergence theorem it follows that

$$\sum_{i=1}^{d} \int_{\mathbb{R}^d} \frac{\partial (u F_i v)}{\partial x_i}\, \mathrm{d}x = 0\,,$$

and therefore

$$\langle \mathcal{A} u, v\rangle = -\sum_{i=1}^{d} \int_{\mathbb{R}^d} v \cdot \frac{\partial (u F_i)}{\partial x_i}\, \mathrm{d}x = \left\langle -\sum_{i=1}^{d} \frac{\partial (u F_i)}{\partial x_i}, v \right\rangle = \langle -\mathrm{div}(F \cdot u), v\rangle$$

for all continuously differentiable functions $u \in D(\mathcal{A})$ and continuously differentiable functions $v \in D(\mathcal{A}_{\mathcal{K}_t})$ with compact support. Since $D(\mathcal{A})$ forms a dense subset of $\mathcal{L}_1$, compare [55, Remark 7.6.2 & Theorem 7.5.1], the semigroup of Frobenius-Perron operators is strongly continuous, i.e.

$$\lim_{t\to t_0} ||\mathcal{P}_t u - \mathcal{P}_{t_0} u|| = 0\,, \quad \forall\, u \in \mathcal{L}_1\,,\ t, t_0 \geq 0\,,$$

which allows us to state the following relation between the infinitesimal generator $\mathcal{A}$ and the differential equation describing the evolution of a density function under deterministic dynamics.

Proposition 1.2.5 (ODEs with random initial values & the infinitesimal generator of $\mathcal{P}_t$, [55, Chapter 7]). *Assume $F : \mathbb{R}^d \to \mathbb{R}^d$ is continuously differentiable. Then the evolution of the probability density function $u_t = u(t, \cdot)$ associated with the Random Initial Value Problem 1.1.2 with right hand side F is described by the first-order linear partial differential equation*

$$\frac{\partial}{\partial t} u = \mathcal{A}u = -\mathrm{div}(F \cdot u)\,, \qquad u(0, \cdot) = u_0\,. \tag{1.18}$$

Remark 1.2.6. *If $\mathrm{div}(F(x)) = 0$ for all $x \in \mathbb{R}^d$, as e.g., in Hamiltonian dynamics, then the above PDE is called the* Liouville equation.

1.2.3 Solution of the PDE along characteristics

First-order partial differential equations can be solved along characteristic curves or *characteristics*. Characteristics are curves $(t(s), x(s))_{s\in\mathbb{R}}$ in $\mathbb{R}^{d+1}$ along which the value $u(t(s), x(s))$ of a solution u is described by an ordinary differential equation. To understand this, let us recall the previously derived PDE (1.18), which can be rewritten as

$$\frac{\partial u}{\partial t} + \sum_{i=1}^{d} F_i \cdot \frac{\partial u}{\partial x_i} = -\mathrm{div}(F) \cdot u\ . \tag{1.19}$$

A solution $u(t(s), x(s)) =: z(s)$ parameterized by s has to satisfy

$$\frac{\mathrm{d}z}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s} u(t, x) = \frac{\partial u}{\partial t} \cdot \frac{\mathrm{d}t}{\mathrm{d}s} + \sum_{i=1}^{d} \frac{\partial u}{\partial x_i} \cdot \frac{\mathrm{d}x_i}{\mathrm{d}s}\,. \tag{1.20}$$

Comparison with (1.19) suggests to set $\frac{\mathrm{d}t}{\mathrm{d}s} = 1$ and $\frac{\mathrm{d}x_i}{\mathrm{d}s} = F_i$, so that

$$\frac{\mathrm{d}z}{\mathrm{d}s} = \frac{\partial u}{\partial t} + \sum_{i=1}^{d} F_i \cdot \frac{\partial u}{\partial x_i} = -\mathrm{div}(F) \cdot z\,. \tag{1.21}$$

Therefore, the PDE has been transformed to the system of ODEs

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s} t(s) &= 1 \\ \frac{\mathrm{d}}{\mathrm{d}s} x(s) &= F(x(s)) \\ \frac{\mathrm{d}}{\mathrm{d}s} z(s) &= -\underbrace{\mathrm{div}\left[F\left(x(s)\right)\right]}_{=:\lambda(x(s))} \cdot z(s) \end{aligned} \tag{1.22}$$

which can be solved analytically for initial values $t(0) = 0$, $x(0) = x_0$ and $z(0) = u(0, x_0)$. The solutions are:

$$\begin{aligned} t(s) &= s \\ x(t) &= \Phi_t x_0 \\ z(t) &= u(\Phi_t x_0, t) = e^{-\hat{\lambda}(t)} \cdot u(x_0, 0)\,, \end{aligned} \tag{1.23}$$

with $\hat{\lambda}(t) := \int_0^t \lambda(x(s))\,\mathrm{d}s$. This way of solving PDEs is called the *method of characteristics*. For a more detailed description, see e.g. [20, 26]. Comparison of the obtained solution with the explicit formula (1.13) for the Frobenius-Perron operator further implies that $e^{-\hat{\lambda}(t)} = \left|\frac{\mathrm{d}}{\mathrm{d}x}\Phi_t x\right|^{-1}$.

Concluding remarks In this chapter we have seen that the impact of uncertainty and variability in the input of ODE models can be studied using two equivalent characterizations of the problem. The first one is based on the solution of the ODE for a random initial value X_0 and yields the random state X_t with probability density function u_t. The second approach is based on the description of the the probability density function by means of a first-order linear PDE, and solution yields the density u_t. Frobenius-Perron operators as well as the method of characteristics provide a link between the two approaches. Next, in Chapter 2, we discuss numerical methods for both approaches.

Chapter 2

Numerical solutions for deterministic systems with random initial values

In this chapter we give an overview on numerical methods to solve the Random Initial Value Problem 1.1.2, i.e.,

$$\dot{x} = F(x)\,, \quad \text{with} \quad x(0) = x_0\,, \tag{2.1}$$

where $x_0 = X_0$ is a random variable with probability density function u_0. We distinguish between methods that solve the ODE directly and methods that solve the equivalent PDE formulation

$$\frac{\partial}{\partial t}u = -\mathrm{div}(F \cdot u) = \mathcal{A}u\,, \qquad u(0,\cdot) = u_0\,. \tag{2.2}$$

First, in Section 2.1, we focus on ODE-based methods, and later, in Section 2.2, on general strategies for the solution of PDEs. The applicability of methods from both approaches to the global sensitivity analysis of ODEs is then discussed in Section 2.3.

2.1 ODE-based approaches

Sensitivity analysis methods can be divided into local and global approaches. Local, or *linear*, sensitivity analysis considers small changes in the model input x_0 and studies their propagation along the solution $x(t) = \Phi_t x_0$ locally, based on a linearization of the dynamics. Here, Φ_t denotes the evolution operator of (2.1) as defined in the previous chapter. This strategy is briefly described in Section 2.1.1. Global sensitivity methods commonly rely on a sampling-based exploration of the input space. Sampling-based or Monte Carlo (MC) methods are described in Section 2.1.2.

2.1.1 Local sensitivity analysis of ODEs

Here, we follow the presentation of linear sensitivity in [25, Chapter 3] and consider a small change or perturbation δ_{x_0} around the initial value x_0. The ODE with a perturbed initial value $x_0 + \delta_{x_0}$ has the solution

$$x(t) = \Phi_t(x_0 + \delta_{x_0})\,. \tag{2.3}$$

Using the Taylor expansion of Φ_t around x_0, the linearized perturbation at time t,

$$\delta_x(t) = \Phi_t(x_0 + \delta_{x_0}) - \Phi_t x_0\,, \tag{2.4}$$

can be written as

$$\delta_x(t) = W_t \cdot \delta_{x_0} \,, \tag{2.5}$$

where $W_t \in \mathbb{R}^{d\times d}$ denotes the Jacobian of Φ_t evaluated at $\Phi_t x_0$, i.e.

$$W_t = \frac{\partial}{\partial x}\Phi_t x\bigg|_{x=\Phi_t x_0} \,. \tag{2.6}$$

Under the linearized dynamics, W_t propagates an initial perturbation along the trajectory $(\Phi_t x_0, t)$ and is therefore called propagation matrix (also Wronski or sensitivity matrix [86]). Since F and Φ_t are assumed to be differentiable with respect to x, we can establish an ODE for the evolution of the propagation matrix by

$$\frac{\mathrm{d}}{\mathrm{d}t}W_t = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial x}\Phi_t x\right) = \frac{\partial}{\partial x}\left(\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t x\right) = \frac{\partial}{\partial x}F(\Phi_t x)\cdot W_t \,. \tag{2.7}$$

Multiplication with δ_{x_0} and considering relation (2.5) yields the initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta_x(t) = \frac{\partial}{\partial x}F(\Phi_t x)\,\delta_x(t) \,, \quad \delta_x(0) = \delta_{x_0} \,, \tag{2.8}$$

which is called the *variational equation*. For linear ODEs, it can be solved analytically. For nonlinear ODEs, it yields a good approximation only for small times t and small initial perturbations δ_{x_0}. Then, the variational equation may be solved numerically, and the solution $\delta_x(t)$, $t \geq 0$, denotes a linearized perturbation along the trajectory $(\Phi_t x_0, t)$. By its deformation, i.e. expansion or contraction, the sensitivity of Φ_t with respect to the model input x_0 is studied. The linearized propagation of an initial perturbation is illustrated in Figure 2.1.

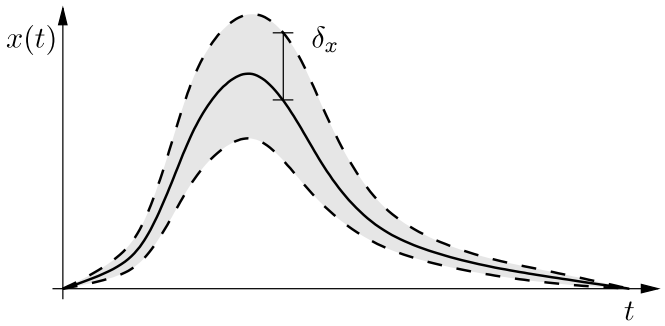

Figure 2.1: Linear sensitivity analysis of ODEs. The solid line denotes the evolution $x(t) = \Phi_t x_0$ of the unperturbed initial value x_0. Dashed lines indicate the evolution of the linearized perturbation $\delta_x(t)$ around $x(t)$. Solutions to the perturbed initial value problem are expected to remain within the grey shaded area.

In a probabilistic interpretation, δ_{x_0} can denote the standard deviation of a normally distributed initial value $X_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$ with

$$\mu_0 = x_0 \quad \text{and} \quad \Sigma_0 = \begin{pmatrix} (\delta_{x_0})_1^2 & & 0 \\ & \ddots & \\ 0 & & (\delta_{x_0})_d^2 \end{pmatrix} \,,$$

where $(\delta_{x_0})_i$ denotes the i-th component of δ_{x_0}. We have seen earlier in Example 1.2.3 that linear dynamics result in a Gaussian distribution of X_t. Then, the variational equation describes the evolution of a linearized estimate of the standard deviation $\delta_x(t)$ of the Gaussian distribution, and the shaded region in Figure 2.1 denotes error bounds or confidence intervals along the trajectory.

Often, linear or local sensitivity analysis refers to a description of the sensitivity by means of *sensitivity indices* or *sensitivity coefficients* [74, 86]. These denote the partial derivatives of the evolution Φ_t with respect to the uncertain variable $x(t)$ and can be extracted from the propagation matrix W_t.

Local or linear sensitivity analysis yields estimates of the uncertainty in the model output for small perturbations of the model input. Further, the estimate reflects the true sensitivity of the model only if perturbations remain small during propagation, or if the dynamics are linear.

2.1.2 Global sensitivity analysis using Monte Carlo methods

Since we assume that the uncertainty and/or variability in the model input $x_0 = X_0$ is captured by the probability distribution of X_0, a straightforward approach is to sample from this distribution, which yields a set of sample points $\{\xi_1, \dots, \xi_M\}$. If the sample size M is sufficiently large, then by the law of large numbers, the initial probability distribution can be approximated by

$$\mathbb{P}[X_0 \in B] = \int_B u_0(x)\ \mathrm{d}x \ \approx\ \frac{1}{M}\sum_{m=1}^{M} \mathbb{1}_{\{B\}}(\xi_m)\,, \quad \mathbb{1}_{\{B\}}(\xi) := \begin{cases} 1, & \xi \in B \\ 0, & \xi \notin B \end{cases}\,. \tag{2.9}$$

With a subsequent solution of the ODE for each of the sample points, the probability distribution at time t can analogously be estimated by

$$\mathbb{P}[X_t \in B] = \int_B u_t(x)\ \mathrm{d}x \ \approx\ \frac{1}{M}\sum_{m=1}^{M} \mathbb{1}_{\{B\}}(\Phi_t\xi_m)\,. \tag{2.10}$$

Similarly, the propagated sample points can be used to estimate other observables, including the mean or the variance of X_t.

The law of large numbers guarantees that the *expected* approximation error decays with[1] $\mathcal{O}(M^{-1/2})$ as $M \to \infty$. If an observable $\varphi : \mathbb{R}^d \to \mathbb{R}$ is differentiable and a function with bounded variation $V(\varphi) < \infty$, where

$$V(\varphi) := \int_\Omega |\nabla\varphi(x)|\ \mathrm{d}x\,,$$

then by the Koksma-Hlawka inequality, see e.g. [38, 67, 70, 71], the approximation error of the realization $\{\Phi_t\xi_1, \dots, \Phi_t\xi_M\}$ on a compact domain $\bar{\Omega} \subset \Omega = \mathbb{R}^d$ is proportional to $V(\varphi)$ with

$$\left|\frac{1}{M}\sum_{m=1}^{M}\varphi(\Phi_t\xi_m) - \int_{\bar{\Omega}}\varphi(x)\cdot u_t(x)\ \mathrm{d}x\right| \le V(\varphi)\cdot D_{t,M}(\bar{\Omega}, \xi_1, \dots, \xi_M)\,. \tag{2.11}$$

[1] See Definition A.1.1 in Appendix A for $\mathcal{O}$-notation.

The term $D_{t,M}(\bar{\Omega}, \xi_1, \ldots, \xi_M)$ is called the *discrepancy* of the sample on $\bar{\Omega}$. It is defined by

$$D_{t,M}(\bar{\Omega}, \xi_1, \ldots, \xi_M) := \sup_{B \subseteq \bar{\Omega}} \left| \left(\frac{1}{M} \cdot \sum_{m=1}^{M} \mathbb{1}_{\{B\}}(\Phi_t \xi_m) \right) - \mathbb{P}[X_t \in B] \right| \tag{2.12}$$

and measures how well the points $\{\Phi_t\xi_1, \ldots, \Phi_t\xi_M\}$ represent the distribution of X_t on $\bar{\Omega}$ (we consider the discrepancy with respect to nonuniform distributions, compare with [33]). In practice, we are interested in approximating the distribution of X_t accurately on a domain that contains most of the probability mass, i.e.

$$\mathbb{P}[X_t \notin \bar{\Omega}_t] \leq \varepsilon \ll 1\,, \quad \forall\, t \in [0, T]\,.$$

The region of interest thus depends on the distribution of X_t, and therefore on the evolution $\Phi_t(X_0) = X_t$. The following examples illustrate how expansions and contractions of the evolution can affect the approximation error in (2.11).

Example 2.1.1. *Consider the one-dimensional case $d = 1$ and a random variable X uniformly distributed on an interval $\Omega = [0, L]$, $0 < L < \infty$. Assume a set of equidistant points $x_m := m \cdot \frac{L}{(M+1)}$, $m = 1, \ldots, M$, is given. Since the state space is a compact interval, the discrepancy can be computed on the whole domain Ω:*

$$\begin{aligned} D_M(\Omega, x_1, \ldots, x_m) &= \sup_{0<\ell\leq L} \left| \left(\frac{1}{M} \sum_{m=1}^{M} \mathbb{1}_{[0,\ell)}(x_m) \right) - \frac{\ell}{L} \right| \\ &= \sup_{0<\ell\leq L} \left| \left\lfloor \frac{\ell}{M} \right\rfloor - \frac{\ell}{L} \right| = \left| \frac{L}{M} - 1 \right|, \end{aligned}$$

which implies that for constant M and increasing L—i.e. for an expanding state space Ω—the discrepancy increases (and so does the approximation error).

□

Example 2.1.2. *Now consider a normally distributed initial value X_0 with mean μ_0 and variance σ_0^2, and linear dynamics*

$$\dot{x} = F(x) = \alpha \cdot x\,, \quad x(0) = X_0\,.$$

For $\alpha > 0$, the evolution operator $\Phi_t x = e^{\alpha t} x$ denotes an expanding transformation of the state space, and we know from Example 1.2.3 that X_t, $t \geq 0$, is normally distributed with mean $\mu_t = e^{\alpha t} \cdot \mu_0$ and variance $\sigma_t^2 = e^{2\alpha t} \cdot \sigma_0^2$. Figure 2.2 depicts the approximation error of

(i) the mean estimated by

$$\hat{\mu_t} := \frac{1}{M} \cdot \sum_{m=1}^{M} \Phi_t \xi_m\,,$$

(ii) the probability of an interval $[a, b]$ estimated by

$$\hat{\mathbb{P}}[X_t \in [a, b]] := \frac{1}{M} \cdot \sum_{m=1}^{M} \mathbb{1}_{[a,b]}(\Phi_t \xi_m)\,,$$

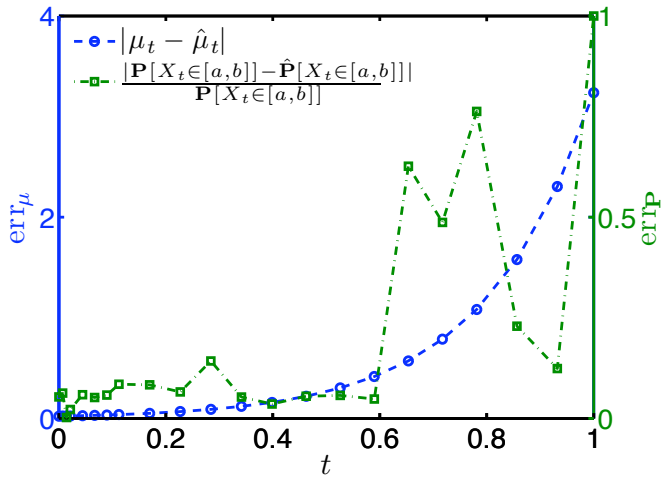

Figure 2.2: Error of the mean (dashed blue, left y-axis) and the probability of an interval $[a, b]$ (dotted-dashed green, right y-axis) estimated with $M = 1000$ sample points and linearly expansive dynamics.

for $M = 1000$, $\mu_0 = 0$, $\sigma_0^2 = 0.5$, $\alpha = 5$ *and* $[a, b] = [1, 1.5]$. *The example illustrates how the approximation errors increase exponentially with the exponentially expanding distribution of* X_t. *While for linear dynamics the distribution is globally expanded or contracted, for nonlinear dynamics, local contractions and expansions of the evolution operator* Φ_t *can moreover cause spatial inhomogeneities of the approximation error, which complicate error estimation.*

□

Sampling-based strategies as the one described above are referred to as MC-based sensitivity analysis [74, 86]. Numerous adaptations have been suggested that most aim at reducing the number of sample points while maintaining or improving the approximation quality [36], however not error-controlled, i.e., based on error estimates. (Among those, probably the most important extensions are the Fourier amplitude sensitivity test (FAST) [22, 79] and the surface response method [10, 37, 51].) Due to their simplicity (the possibility to use standard sampling techniques and ODE solvers) but most importantly due to their applicability in high dimensions, MC-methods and adaptations of those constitute a widely used tool for global sensitivity analysis [36].

2.2 Numerical solution of PDEs

Global sensitivity analysis of ODEs can equivalently be studied by solving the PDE (2.2) that describes the evolution of the probability density function associated with the random state variable $x(t) = X_t$. This approach is often referred to as *stochastic sensitivity analysis* [19, 74, 86]. In this section we give a brief overview of approaches to solving time-dependent PDEs. The numerical solution of PDEs is a broad field of research, and there is a rich literature devoted to it [29, 73, 84]. We confine ourselves to treating only those concepts with more detail that are relevant to discuss the approach proposed in this work.

Time-dependent PDEs are commonly solved by treating the spatial and temporal domains separately. A discretization in one of the domains is called *semi-discretization*. In the following section we classify methods by the order in which the semi-discretization is applied. A brief overview of the most important spatial discretization techniques is then given in Section 2.2.2, and Section 2.2.3 introduces concepts of temporal discretization. In Section 2.2.4 we focus on TRAIL, a method by Horenko et al. [39, 41, 42], that was developed for a particular case of our problem. We point out that this research was motivated by the attempt to transfer and apply TRAIL to the global sensitivity analysis of ODEs.

2.2.1 Method of lines & Rothe method

A solution $u : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ to a time-dependent PDE is a function in time and space. Semi-discretization in space corresponds to a computation of u at discrete space points or a representation in a finite-dimensional function space. Semi-discretization in time is the computation of u at discrete time points. Depending on the order of semi-discretization we distinguish between methods of lines, which first conduct a semi-discretization in space, and Rothe methods, which first apply semi-discretization in time. Semi-discretization results in a reduced problem, a temporal problem in case of the method of lines, and a spatial (stationary) problem in case of the Rothe method, as illustrated in Figure 2.3.

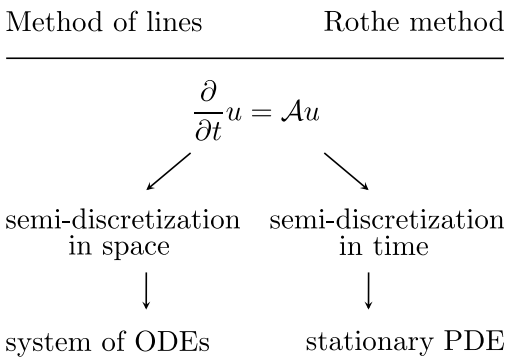

Figure 2.3: Spatio-temporal discretization by the method of lines and the Rothe method.

In the following we demonstrate by means of two examples how the reduced problems can be derived. For a comprehensive introduction we refer to [80, 84] for a numerical treatment by the method of lines and [75, 76, 46] with an emphasis on the Rothe method.

The method of lines

For first order PDEs, an initial semi-discretization in space results in a system of ODEs. Solving this system of ODEs yields a discrete solution along trajectories—or *lines*—in time, as illustrated in Figure 2.4, which is why the approach is called method of lines. We exemplify the derivation of such ODEs with a finite difference approximation of the spatial derivatives of the PDE.

Example 2.2.1 (Semi-discretization in space by first-order finite differences). *Consider the one-dimensional case $d = 1$. The PDE describing the evolution of an initial probability density function u_0 under deterministic dynamics F is given by*

$$\frac{\partial}{\partial t}u = \mathcal{A}u = -\frac{\partial}{\partial x}(F \cdot u) \quad \text{with} \quad u(0,\cdot) = u_0\,. \tag{2.13}$$

A first-order finite difference approximation of the spatial derivative of $u(t,\cdot) = u_t$ yields

$$\frac{\partial}{\partial x}u_t(y) \approx \frac{u_t(y) - u_t(z)}{y - z}\,, \quad y, z \in \mathbb{R}\,. \tag{2.14}$$

Given a finite set of points $x_m \in \mathbb{R}$, $m = 1, \ldots, M$, a substitution of $\frac{\partial}{\partial x}u$ for the above finite difference approximation transforms (2.13) to

$$\frac{\partial}{\partial t}u_t(x_m) = -F'(x_m) \cdot u_t(x_m) - F(x_m) \cdot \frac{u_t(x_m) - u_t(x_{m-1})}{x_m - x_{m-1}}\,. \tag{2.15}$$

With the remaining temporal derivative of u, the problem has been transformed into a system of ODEs. It can be solved using the initial values

$$u(0, x_m) = u_0(x_m)\,, \quad m = 1, \ldots, M\,,$$

where $u(t, x_1)$, $t \geq 0$, needs to be specified by a boundary condition. Solution yields a fully discrete solution $u(t_j, x_m)$ at discrete time points $t_j \in \mathbb{R}^+$, $j = 0, 1, \ldots$, and space points x_m, $m = 1, \ldots, M$.

□

The above example shows how ODEs for the function values $u_t(x_m)$ can be derived by an approximation of the spatial derivatives. Alternatively, u can be represented in a finite-dimensional function space, i.e. as a linear combination of a finite number of basis functions, see e.g. [29]. Then, a system of ODEs for the coefficients of the linear combination can be derived in a similar way. Higher-order PDEs, which require the inclusion of boundary conditions, yield differential algebraic equations (DAE) instead of ODEs. Due to the possibility of using standard numerical ODE (or DAE) solvers, the method of lines is a popular tool for the solution of time-dependent PDEs.

The Rothe method

The Rothe method first conducts a semi-discretization in time, which is why it is also referred to as the method of discretization in time [46, 75, 76]. The basic idea is to consider the PDE as an ODE in a function space. Semi-discretization in time then corresponds to the application of ODE discretization strategies and yields time-independent or stationary PDEs, see e.g. [29]. Solution of those yields an approximation of the function u at discrete time points as illustrated in Figure 2.4. We exemplify the derivation of the stationary problems using the implicit Euler method to approximate the temporal derivative.

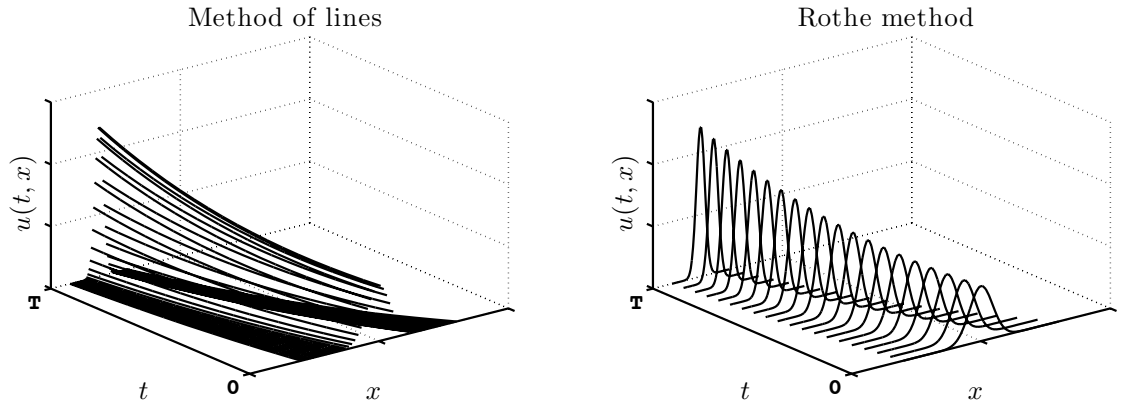

Figure 2.4: Spatio-temporal discretization by the method of lines (left) and the Rothe method (right). The method of lines computes temporal trajectories for discrete initial values $u_0(x_m)$, and the Rothe method approximates u_{t_j} at discrete time points. The method of lines and the Rothe method are sometimes referred to as method of vertical and horizontal lines.

Example 2.2.2 (**Semi-discretization in time by the implicit Euler**). *Consider the same time-dependent PDE as in Example 2.2.1. Now we approximate the temporal derivative for a fixed time step $\tau > 0$. Applying an implicit Euler approximation of the temporal derivative yields the sequence of stationary or elliptic PDEs*

$$\frac{u_{t_j} - u_{t_{j-1}}}{\tau} = \mathcal{A}\, u_{t_j}\,, \quad t_j = j \cdot \tau, \quad j = 1, 2, \ldots\,. \tag{2.16}$$

The stationary PDEs can be solved using spatial discretization techniques, which will be discussed later on. Their solution yields a sequence of approximations to u at discrete time points t_j.

□

In the above example we used a fixed time step τ. However, time steps can also be chosen differently in each integration step $t_{j+1} = t_j + \tau_j$. If reliable error estimates are available, time steps can be adjusted such that the estimated error in each integration step remains below a specified tolerance.

The main advantage of the Rothe method over the method of lines relies on the repeated solution of the stationary spatial problems. Instead of choosing a semi-discretization in space once, at $t = 0$, the resolution of the spatial approximation can be adapted according to the structure that a solution develops in the course of its temporal evolution. Therefore, the Rothe method allows for a fully adaptive integration of time-dependent PDEs.

2.2.2 Spatial discretization

Spatial discretization of PDEs can be achieved using various approximation methods. The methods summarized in this section can in general be used for the joint spatial and temporal discretization of PDEs. However, in view of the separate treatment of the temporal and

spatial domains discussed previously, we consider those methods for a spatial discretization only. We focus on the conceptual frameworks and mention only the most important methods; for a more comprehensive treatment, see e.g. [29, 73, 84].

Difference methods In difference methods, the spatial domain is discretized by a set of (preferably) uniform grid points. The pointwise information is then used to replace the derivatives in the PDE by finite difference quotients, which is why the methods are called *finite difference methods*. Discretizing the PDE in this way results in difference equations, which in the case of the method of lines are ODEs, and algebraic equations when used to discretize the stationary spatial problems within a Rothe method. An example of finite difference methods was shown in Example 2.2.1, where the implicit Euler scheme was used to obtain the finite difference quotients. Other difference quotients can be derived via Taylor expansions of the solution, see e.g. [29, 84]. An attractive feature of finite difference methods is their easy implementation. In addition, the Taylor expansion as a basis for the derivation of the difference quotients provides a straightforward convergence theory. However, difference methods require strong assumptions about the smoothness of solutions, and they are generally limited to domains with simple geometries.

Finite volume methods constitute a generalization of finite difference methods. They require less assumptions on the geometry of the spatial domain as well as on the structure of the grid points. Discretization is based on small control volumes surrounding each grid point. Balance equations between the volumes are then derived by considering conservation laws such as mass conservation, or in our case, the conservation of probability mass. Integrating the balance equations by parts (applying the divergence or Gauss's theorem), the PDE is transformed from an integral on the volume to an integral on the surfaces of the volume, see e.g. [56]. Since finite volume methods, by construction, inherently consider conservation laws, they are particularly attractive for conservative systems such as fluid dynamical problems or transport equations as the one derived in the previous chapter, see [53].

Ansatz methods While difference methods approximate the derivatives of a solution, ansatz methods approximate the solution of the PDE itself. This is commonly done by representing the solution in an *ansatz space*, as a linear combination of basis functions. A solution is then obtained by determining the coefficients of the basis functions, either by deriving ODEs that describe the temporal evolution (in case of the method of lines), or by solving algebraic equations (in case of Rothe methods). Methods that require the approximate function to satisfy the PDE at a set of grid points, are called collocation methods. It is often advantageous to satisfy the PDE in a weak sense, which leads to the formulation of variational problems, see e.g. [29]. The weak formulation requires less assumptions about the smoothness of a solution.

Methods that solve the variational problem in a finite-dimensional ansatz space are referred to as *Galerkin (ansatz) methods*. The most important class of Galerkin methods are *finite element methods*, which are based on a geometrical decomposition of the spatial domain into simple sub-domains, usually triangles, which are then called elements. The basis functions are constructed by defining simple functions on those elements, typically piecewise polynomial functions. A major advantage of finite element methods is the construction of

basis functions with small support. Discretizing the variational problem using these basis functions results in sparse, or localized problems, which can be solved more efficiently, see e.g. [3, 29, 73]. Finite element methods are often favorable for problems with a complex spatial domain, especially, if the domain changes in time, or when the solution lacks smoothness.

Spectral methods are closely related to finite element methods. Based on the same ideas, they rely on an approximation of the solution using orthogonal ansatz functions of generally global support, e.g., trigonometric functions or orthogonal polynomials. Although the sparseness is lost by this global approach, spectral methods are often favorable to solve problems with smooth solutions, in particular, if the boundary conditions are periodic, because they yield exponential, or spectral convergence, see e.g. [14, 15, 85].

Sparse grids & meshfree methods The methods mentioned above rely on the construction of grids. For a growing number of dimensions, the computational costs of these grids increase exponentially, which in practice limits the methods to problems with up to two or three dimensions.

Sparse grids are an alternative class of methods, which can be applied to problems in higher dimensions. They rely on the construction of grids that have a low discrepancy with respect to the uniform distribution, see Figure 2.5 (left) and compare (2.12) (note that for $d > 1$ uniform grids do not have minimal discrepancy). It can be shown that there is a one-to-one correspondence between sparse grids and so called *hyperbolic crosses*, see Figure 2.5 (right). With this relation, it is straightforward to show that the number

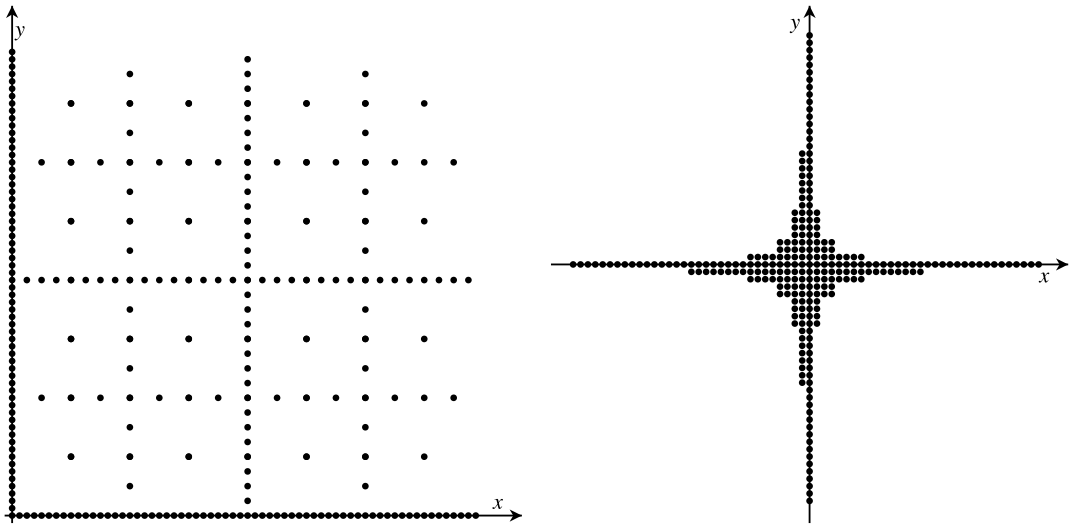

Figure 2.5: Two-dimensional sparse grid (left) and corresponding hyperbolic cross (right).

of grid points does not scale exponentially with the number of dimensions, see e.g. [58]. Therefore, sparse grids can be applied to solve considerably higher-dimensional problems. The solution of the PDE is then approximated in an ansatz space. Each basis function is associated with one grid point and defined by the tensor product of univariate basis functions that are associated with the coordinate in each dimension. Sparse grids yield

solutions of comparable approximation quality to the conventional methods, i.e. solutions of the same approximation order, however, at costs of additional smoothness assumptions, which depend on the number of dimensions, see e.g. [12, 92].

Another class of methods are *meshfree methods*, which constitute a comparably new field of research that has drawn considerable attention in recent years. The concept is rather general; various methods can be classified as meshfree methods (e.g. diffuse element methods, element-free Galerkin methods, generalized finite element methods, moving least squares, or smooth particle hydrodynamics methods, see e.g. [28]). Their common feature is that no assumption about the structure of the grid points is made. The solution of the PDE is approximated in an ansatz space with basis functions, often radial basis functions, which are centered at a set of *scattered* points. In contrast to finite element methods, the basis functions can have variable shapes. Furthermore, the multivariate basis functions only depend on univariate information, which, in the case of radial basis functions, is the distance from the center. Thus, virtually, the basis denotes a univariate basis, which makes these methods particularly attractive for a use in high dimensions, see e.g. [28, 57]. Moreover, since no assumptions are made about the structure of the grid points, they are favorable for problems with complex spatial geometries, especially when the domain changes in time.

2.2.3 Temporal discretization

Semi-discretization of PDEs in space results in systems of ODEs; temporal discretization then denotes the discretization of these ODEs. We omit a discussion of the various techniques for the discretization of ODEs; for a comprehensive treatment, see e.g. [25, 30, 31]. Here, we focus on the temporal semi-discretization of PDEs within the Rothe method.

The basic idea is to consider a time-dependent PDE as an ODE in a function space and treat errors introduced by the spatial discretization as perturbations. Analogously to the solution of ODEs, the semigroup describing the evolution of the solution—in our case the Frobenius-Perron operator $\mathcal{P}_t$—is approximated using rational functions. This is described in more detail in the Appendix, Section A.1. If the rational function satisfies certain properties, namely, *consistency* and *A-stability*, the discrete solution is guaranteed to converge to the analytical solution. Consistency of order $k \in \mathbb{N}$ further implies convergence of order k, see Section A.1 for definitions and Theorem A.1.6 or [11] for the convergence result. In the following example, the approximation by rational functions is illustrated by means of the implicit Euler method.

Example 2.2.3 (**Rational function of the implicit Euler method**). *In Example 2.2.2, the implicit Euler method was used for semi-discretization in time. The corresponding rational function is*

$$r(z) = \frac{1}{1-z}\,.$$

The discrete evolution operator is then defined by

$$R_\tau = r(\tau\mathcal{A}) = (\mathrm{Id} - \tau\mathcal{A})^{-1}\,, \quad \tau > 0\,,$$

where Id *denotes the identity operator. It can be shown that, if r is A-stable and consistent of order one, the discrete solution defined by*

$$u^{(1)}_{t_n} = R^n_\tau u_0 = (\mathrm{Id} - \tau\mathcal{A})^{-1}\, u^{(1)}_{t_{n-1}} \qquad \Leftrightarrow \qquad (\mathrm{Id} - \tau\mathcal{A})\, u^{(1)}_{t_n} = u^{(1)}_{t_{n-1}}\,, \tag{2.17}$$

with $u_0^{(1)} := u_0$ *and* $t_n = n \cdot \tau$, $n \in \mathbb{N}$, *converges to the solution* u_{t_n} *with order* $k = 1$.

□

Selecting the time step τ adaptively in each integration step permits to control the discretization error. Analogously to ODEs, the local discretization error, i.e. the error of one integration step, is typically estimated comparing two solutions of different (consistency and convergence) orders. The time step is then adjusted such that the estimated error remains below a specified accuracy, or tolerance condition. Error estimation and time step selection are described in Section A.2, in the Appendix.

Semi-discretization in time results in a sequence of stationary spatial problems (2.17), which are solved using spatial discretization techniques as introduced in the previous section. Spatial discretization yields solutions $\hat{u}_{t_j}^{(1)} \approx u_{t_j}^{(1)}$, which introduce additional local errors, or spatial perturbations,

$$\delta_{t_j} := \hat{u}_{t_j}^{(1)} - u_{t_j}^{(1)} .$$

Since the stationary problems are solved independently, the spatial resolution can as well be adjusted in each integration step to meet spatial accuracy conditions. This constitutes a major advantage of the Rothe method over the method of lines, where the spatial resolution is determined initially.

The following example illustrates the impact of the spatial perturbations to the global approximation error

$$\varepsilon_{\text{glob}} = \left|\left| u_T - \hat{u}_T^{(1)} \right|\right| .$$

By means of an ODE, with spatial perturbations in each integration step, we draw attention to dependencies between spatial accuracy and temporal discretization.

Example 2.2.4 (**ODE with spatial perturbations**). *Consider* $d = 1$ *and a linear ODE*

$$\dot{x} = \alpha \cdot x\,, \quad \alpha \in \mathbb{R}\,,\ x(0) = x_0\,.$$

We solve the ODE for $t \in [0, T]$ *using the implicit Euler method with a fixed time step* $\tau > 0$. *In each integration step, a random perturbation is added, i.e.*

$$\hat{x}(t_j) = \frac{1}{1 - \tau\alpha} \cdot \hat{x}(t_{j-1}) + \xi(t_j)\,, \quad t_j = j \cdot \tau,\ j = 1, \ldots, n\,,$$

where $\xi(t_j) \sim \mathcal{N}(0, \sigma^2)$ *are normally distributed random variables with standard deviation* $\sigma = \delta_x$. *To obtain representative results, the system is solved* $N = 1000$ *times, and the mean of the solutions* $\hat{x}_1(T), \ldots, \hat{x}_N(T)$ *is compared to the analytical solution* $x(T) = e^{T \cdot \alpha} \cdot x_0$.

The double-logarithmic plot in Figure 2.6 shows the global approximation error for different choices of the time step. The dashed line denotes the error for $\delta_x = 0.1$ *constant and indicates a growing error for* $\tau \to 0$. *This can be explained by the accumulation of spatial errors in the* $n = {}^{T}/_{\tau}$ *integration steps until* $t_n = T$, *where* n *grows for decreasing* τ. *Since* δ_x *was chosen independent of* τ, *the spatial perturbations can be expected to build up with* $\mathcal{O}(\tau^{-1})$ *as* $\tau \to 0$.

Consequently, to ensure convergence of the spatially perturbed solutions, $\delta_x = \delta_x(\tau)$ *cannot be chosen independently of* τ. *To further observe convergence of order* k, $\delta_x(\tau)$ *should vanish with* $\mathcal{O}(\tau^{k+1})$ *as* $\tau \to 0$. *The solid line in Figure 2.6 shows the global approximation error for* $\delta_x(\tau) = 0.1 \cdot \tau^{k+1}$. *Comparison with the dotted line indicates an error decay of order* $k = 1$, *as expected for the implicit Euler method.*

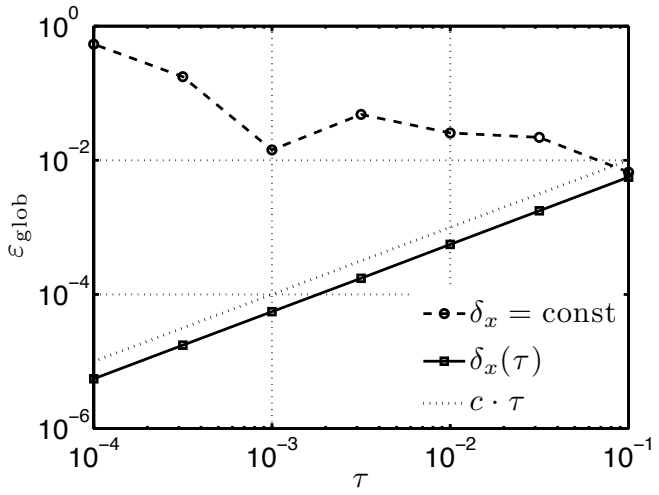

Figure 2.6: Global discretization error of the implicit Euler method for time steps τ.

□

The above example makes clear that for the discretization of time-dependent PDEs, the spatial accuracy cannot be chosen independently of the temporal discretization. This dependency will be analyzed more extensively in Chapter 6.

2.2.4 TRAIL

TRAIL stands for "Trapezoidal Rule for Adaptive Integration of Liouville dynamics". The method was developed by Horenko et al. [39, 41, 42] for the solution of Liouville-type equations in the context of quantum molecular dynamics. Its main feature is adaptivity in time and space based on a Rothe approach and a meshfree spatial discretization. Promising results in quantum molecular dynamics motivated its application to the general problem of ODEs with random initial values, which was the starting point of this research.

Adaptivity in time is based on the comparison of the second-order solution $u_t^{(2)}$ obtained via the trapezoidal rule,

$$r(z) = \frac{1+z}{1-z} \quad \Rightarrow \quad \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right) u_{t_j+\tau_j}^{(2)} = \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right) u_{t_j}^{(2)}, \quad j = 0, \ldots, n, \tag{2.18}$$

to the first-order implicit Euler solution $u_{t_j}^{(1)}$, compare Example 2.2.3. Spatial discretization relies on the representation of the solutions in an ansatz space with Gaussian basis functions, i.e.

$$\hat{u}_{t_j}^{(k)}(x) = \sum_{i=1}^{N} \omega_i \cdot \eta_i(t_j, x), \quad k = 1, 2,$$

with

$$\eta_i(t_j, x) = \frac{1}{\sqrt{2\pi}|\Sigma_i(t_j)|} \cdot \exp\left(-\frac{1}{2} \cdot [x - \mu_i(t_j)]^T \cdot (\Sigma_i(t_j))^{-1} \cdot [x - \mu_i(t_j)]\right).$$

The basis functions in each integration step are predicted by a linearized propagation of the previous basis functions along the characteristics of the PDE. Since the linearly predicted basis can be computed analytically (compare Example 1.2.3) same as the action of the generator $\mathcal{A}$ on the Gaussian basis functions (see Appendix, Part C), the stationary problems can be restated as linear least-squares problems. Solution of these yields the coefficients of the basis functions. Spatial adaptivity is achieved by an adaptive insertion or pruning of basis functions, based on residual error estimates of the optimization problems.

Note that the Gaussian basis functions are not positioned uniformly on the spatial domain, but scattered according to local error estimates. This meshfree construction allows for an application of the method to higher-dimensional problems, see e.g. [40], where the method was applied to a six-dimensional problem.

2.3 Discussion of the different approaches

Each of the different approaches shown in this chapter has its advantages and drawbacks, which we briefly discuss in the following.

ODE-based approaches Linear sensitivity analysis provides a powerful tool for the study of ODEs with perturbed initial values. It yields a *local and linearized* measure of the sensitivity and is thus only appropriate if perturbations are small. Consequently, it is not adequate for the global sensitivity analysis of ODEs, in particular when studying the impact of variability, which generally involves variations over a substantial domain.

Most approaches to global sensitivity analysis of ODEs are based on MC-methods, see e.g. [36] and references therein. Due to their flexibility and applicability to high dimensions, MC-based methods are the only choice for the analysis of many complex systems in practical applications. However, error estimation and control denotes a major challenge that has not been solved adequately.

PDE-based approaches The PDE-based approach to global sensitivity analysis gives access to a profound theory and broad methology. Methods of lines are generally simple to implement due to the possibility of using standard ODE solvers. Concerning error control, adaptive ODE solvers straightforwardly allow for temporal adaptivity. However, spatially adaptive methods of lines commonly rely on *a-posteriori* error estimates, which require a complete solution of the system, before the spatial discretization can be adapted, see e.g. [1]. In that respect, Rothe methods offer a substantial advantage, since the temporal *and* spatial discretization can be adjusted in each integration step.

Spatial discretization within the Rothe method Among the conventional discretization strategies, finite volume methods are generally favorable for our problem, because conservation principles are inherently included. However, we aim at solving problems in higher dimensions than those that can be treated with conventional discretization techniques.

Sparse grids provide one option. But prior to the usage of sparse grids, a complete understanding of the "full" grid situation is desirable. Another option are meshfree methods, which in our view is the most promising approach to our problem.

The starting point of this research was the transfer of the TRAIL scheme to the general problem of ODEs with random initial values. TRAIL combines all desirable features mentioned earlier: error control based an adaptive Rothe scheme, and applicability to high-dimensional problems due to meshfree spatial discretization.

First studies of the sensitivity of pharmacokinetic models proved applicability of the general strategy, but revealed severe problems concerning the adaptive error control. The temporal error estimation is error-prone due to possible cancellation effects. Furthermore, the coupling of temporal discretization and spatial accuracy requires large numbers of basis functions to meet the accuracy condition—even for one-dimensional problems, and especially for functions with steep gradients. Initial spatial refinement typically necessitated further refinements, resulting in inefficiently large numbers of basis functions and extremely small time steps. Moreover, the propagated basis functions can become very *wide*, i.e. with large variance, so that the assumption underlying the linearized propagation is questionable.

Heuristic modifications addressing these problems were published in [91]. These modifications, however, improved the performance only to a minor extent. The lack of a theoretical basis to prove convergence of the spatial discretization scheme obstructed substantial improvements.

A novel approach The objective of this work is to develop a fully adaptive numerical scheme to solve ODEs with random initial values. Based on the PDE formulation, the problem is addressed using an adaptive Rothe method as introduced by Bornemann for the solution of parabolic PDEs [8, 9]. The key feature of this method is a *multiplicative error correction*, which realizes the computation of temporal error estimates in a multiplicative fashion to avoid numerical cancellation effects. The method also provides a framework for coupling temporal discretization and spatial accuracy (which will be modified in the course of this work to account for properties of the PDE considered herein).

The adaptive Rothe scheme is combined with *approximate approximations* to solve the stationary spatial problems. This novel approximation method developed by Maz'ya et al. [64, 66] is based on representing a function in an ansatz space spanned by a rather general class of basis functions. We consider approximate approximations with radial basis functions. In particular, the method provides a powerful convergence theory for Gaussian basis functions. It further allows for constructing basis functions that yield high approximation orders and thereby facilitate an efficient solution of the spatial discretization problem. The basis functions are centered around grid points covering the spatial domain. The grid points can be uniformly distributed or scattered. In this work we consider approximate approximations on uniform grids in order to fully understand theoretical and numerical properties of the overall discretization scheme within the adaptive Rothe setting. The perspective is to combine the Rothe method with approximate approximations on scattered points, i.e. in a meshfree setting.

Part II

A novel approach to adaptive density propagation

Chapter 3

A Rothe method with multiplicative error correction

In the previous chapter we concluded that in terms of adaptivity the Rothe method is advantageous for the numerical solution of time-dependent PDEs. In the following, we present an adaptive Rothe method that was introduced by Bornemann [8, 9] for the solution of parabolic PDEs. The method is applied to solve the first-order linear PDE

$$\frac{\partial}{\partial t}u = \mathcal{A}u = -\mathrm{div}(F \cdot u)\,, \qquad u(0,\cdot) = u_0\,, \tag{3.1}$$

which describes the evolution of a probability density function $u : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ under deterministic dynamics F. As demonstrated previously, in Section 2.2.1, the Rothe method first performs a semi-discretization in time, which yields stationary spatial problems. Let $u_{t+\tau}^{(k)} := R_\tau^{(k)} u_t$ denote the exact solution of the stationary problem imposed by the rational approximation $R_\tau^{(k)}$ to the strongly continuous semigroup of order k, i.e.

$$||\varepsilon_t(\tau)|| = \left|\left|u_{t+\tau} - u_{t+\tau}^{(k)}\right|\right| = \mathcal{O}(\tau^{k+1})\,, \quad \text{as} \quad \tau \to 0\,,$$

where $u_{t+\tau}^{(0)} := u_t$. Further let

$$\Delta u_{t+\tau}^{(k-1)} := u_{t+\tau}^{(k)} - u_{t+\tau}^{(k-1)}\,, \quad k \geq 1\,, \tag{3.2}$$

denote the difference or *correction* between two solutions of order k and $k-1$. Since the local errors of $u^{(k)}$ and $u^{(k-1)}$ decay with $\mathcal{O}(\tau^{k+1})$ and $\mathcal{O}(\tau^k)$, the correction will have the same asymptotic behavior as the true local error of $u^{(k-1)}$ (see Part A of the Appendix). Consequently, the local temporal error is estimated by

$$\left|\left|\varepsilon_t^{(k-1)}(\tau)\right|\right| = \left|\left|\Delta u_{t+\tau}^{(k-1)}\right|\right| = \mathcal{O}(\tau^k)\,, \quad \text{as } \tau \to 0\,. \tag{3.3}$$

In the adaptive setting, time steps τ_j are chosen in each integration step $t_j \in [0,T]$, $j = 0, 1, \ldots$, such that the estimate $\varepsilon_{t_j}^{(k-1)}$ remains below a specified tolerance. A source of problems with this procedure is the possibility of numerical cancellation in the above estimation (3.3). To prevent this, the corrections $\Delta u_{t_j}^{(i)}$ for $i = 0, \ldots, k-1$, are computed recursively, in a *multiplicative* fashion.

In the following we introduce the temporal semi-discretization scheme with multiplicative error correction as in [8]. Later, in Section 3.2, we discuss the influence of spatial perturbations (caused by the spatial discretization of $u^{(k)}$) to the adaptive selection of time steps and spatial discretization.

3.1 Semi-discretization in time

To construct temporal semi-discretization schemes of a certain order, we first consider a spatially unperturbed solution of the PDE on a time interval $[0, T]$ and a fixed time step $\tau > 0$. Without loss of generality, we assume that $n \cdot \tau = T$ for some $n \in \mathbb{N}$. The true solution is described by the semigroup of Frobenius-Perron operators $\{\mathcal{P}_t\}_{t\in[0,T]}$ via

$$u_t = \mathcal{P}_t u_0\,, \quad t \in [0, T]\,.$$

The following results by Bornemann [8] show how to recursively construct L- and A-stable rational approximations $R_\tau^{(k)}$ to $\mathcal{P}_\tau$ such that the discrete evolution at $t_j = j\cdot\tau$, $j = 1, \ldots, n$, defined by

$$u_{t_j}^{(k)} = \left(R_\tau^{(k)}\right)^j u_0 = R_\tau^{(k)} \left(\left(R_\tau^{(i)}\right)^{j-1} u_0\right)\,, \quad \left(R_\tau^{(i)}\right)^0 u_0 = u_0\,,\ j = 1, \ldots, n\,, \tag{3.4}$$

has consistency order[1] k. Note that by Theorem A.1.6 L-/A-stability and consistency of order k imply convergence of the discrete evolution to the analytical solution with order k, see also [11].

Theorem 3.1.1 (**A family of *L*-stable rational approximations**, [8, Lemma 2.1])**.** *Let*

$$L_i(x) = \frac{e^x}{i!}\,\frac{\partial^i}{\partial x^i}\left(x^i \cdot e^{-x}\right) \tag{3.5}$$

denote the Laguerre polynomial *of order i. Then, the discrete evolution defined in* (3.4) *with $R_\tau^{(k)} := r_L^{(k)}(\tau\mathcal{A})$ and*

$$r_L^{(k)}(z) = \frac{1}{1-z} \cdot \sum_{i=0}^{k} L_i(1) \left(\frac{z}{1-z}\right)^i \tag{3.6}$$

has at least consistency order k. Moreover,

(i) $r_L^{(k)}(z)$ is L-stable, and

(ii) $r_L^{(k)}(z)$ can be computed recursively if $L_i(1) \neq 0$ for $2 \leq i \leq k-1$:

$$\begin{aligned} r_L^{(1)}(z) &:= \frac{1}{1-z} \\ r_L^{(i)}(z) &:= r_L^{(i-1)}(z) + \vartheta_L^{(i-1)}(z), \quad i = 2, \ldots, k, \end{aligned} \tag{3.7}$$

where

$$\begin{aligned} \vartheta_L^{(1)}(z) &:= -\frac{z^2}{2\cdot(1-z)^2} \cdot r_L^{(1)}(z), \\ \vartheta_L^{(i-1)}(z) &:= -\gamma_L^{(i)} \cdot \frac{z}{1-z} \cdot \vartheta_L^{(i-1)}(z)\,, \quad \textit{with} \quad \gamma_L^{(i)} := \frac{L_{i+1}(1)}{L_i(1)}\,, \quad i = 2, \ldots, k\,. \end{aligned} \tag{3.8}$$

[1] see Appendix, Part A for definitions of L-/A-stability, consistency and convergence.

The above recursion allows us to compute the corrections

$$\Delta u_{t_j}^{(i)} = u_{t_j}^{(i+1)} - u_{t_j}^{(i)}, \qquad i = 0, \ldots, k-1,$$

multiplicatively, which, as mentioned previously, is favorable to avoid numerical cancellation. Constructing rational approximations $R_\tau^{(i)} = r_L^{(i)}(\tau\mathcal{A})$ according to Theorem 3.1.1 yields solutions

$$u_{t+\tau}^{(i)} = R_\tau^{(i)} u_t\,, \quad i = 1, \ldots, k\,. \tag{3.9}$$

By recursion (3.7) we have

$$\begin{aligned} u_{t+\tau}^{(i+1)} &= R_\tau^{(i+1)} u_t \\ &= \left(R_\tau^{(i)} + \vartheta_L^{(i)}(\tau\mathcal{A})\right)\, u_t\,, \quad i = 2, \ldots, k-1\,. \end{aligned} \tag{3.10}$$

Combining (3.9) and (3.10), the correction $\Delta u_{t+\tau}^{(i)}$ can be computed multiplicatively by

$$\Delta u_{t+\tau}^{(i)} = \vartheta_L^{(i)}(\tau\mathcal{A})\; u_t\,, \quad i = 2, \ldots, k-1\,, \tag{3.11}$$

and by recursion (3.8), this becomes

$$\Delta u_{t+\tau}^{(i)} = -\gamma_L^{(i)}\,(\mathrm{Id} - \tau\mathcal{A})^{-1}(\tau\mathcal{A})\; \Delta u_{t+\tau}^{(i-1)}\,, \quad i = 2, \ldots, k-1\,. \tag{3.12}$$

Semi-discretization in time yields stationary spatial problems. To illustrate their formulation, we consider the third-order L-stable solution $u_{t+\tau}^{(3)}$ obtained by the rational approximation $R_\tau^{(3)} = r_L^{(3)}(\tau\mathcal{A})$,

$$r_L^{(3)}(z) = \frac{1}{1-z} \cdot \left(1 - \frac{1}{2}\frac{z^2}{(1-z)^2} + \frac{2}{3}\frac{z^3}{(1-z)^3}\right)\,. \tag{3.13}$$

Using the recursion for $r_L^{(3)}$ we get

$$\begin{aligned} u_{t+\tau}^{(1)} &= (\mathrm{Id} - \tau\mathcal{A})^{-1}\; u_t \\ \Rightarrow \quad (\mathrm{Id} - \tau\mathcal{A})\; u_{t+\tau}^{(1)} &= u_t\,, \end{aligned} \tag{3.14}$$

(the implicit Euler approximation) and subsequently

$$\Delta u_{t+\tau}^{(0)} = u_{t+\tau}^{(1)} - u_t\,,$$

which allows us to compute the next corrections for $i = 1, 2$ by

$$\begin{aligned} \Delta u_{t+\tau}^{(1)} &= -\frac{1}{2}(\mathrm{Id} - \tau\mathcal{A})^{-2}(\tau\mathcal{A})(\mathrm{Id} - \tau\mathcal{A})^{-1}(\tau\mathcal{A})\; u_t \\ &= -\frac{1}{2}(\mathrm{Id} - \tau\mathcal{A})^{-2}(\tau\mathcal{A})\; \Delta u_{t+\tau}^{(0)} \\ \Rightarrow \quad (\mathrm{Id} - \tau\mathcal{A})^2\; \Delta u_{t+\tau}^{(1)} &= -\frac{1}{2}(\tau\mathcal{A})\; \Delta u_{t+\tau}^{(0)} \end{aligned} \tag{3.15}$$

and

$$\begin{aligned}\Delta u^{(2)}_{t+\tau} &= \frac{2}{3}(\mathrm{Id}-\tau\mathcal{A})^{-2}(\tau\mathcal{A})(\mathrm{Id}-\tau\mathcal{A})^{-2}(\tau\mathcal{A})^2\ u_t\\ &= -\frac{4}{3}(\mathrm{Id}-\tau\mathcal{A})^{-1}(\tau\mathcal{A})\ \Delta u^{(1)}_{t+\tau}\\ \Rightarrow\quad (\mathrm{Id}-\tau\mathcal{A})\ \Delta u^{(2)}_{t+\tau} &= -\frac{4}{3}(\tau\mathcal{A})\ \Delta u^{(1)}_{t+\tau}\,. \end{aligned} \tag{3.16}$$

Solution of the stationary spatial problems

$$\begin{aligned}(\mathrm{Id}-\tau\mathcal{A})\ u^{(1)}_{t+\tau} &= u_t\\ (\mathrm{Id}-\tau\mathcal{A})^2\ \Delta u^{(1)}_{t+\tau} &= -\frac{1}{2}(\tau\mathcal{A})\ \Delta u^{(0)}_{t+\tau}\\ (\mathrm{Id}-\tau\mathcal{A})\ \Delta u^{(2)}_{t+\tau} &= -\frac{4}{3}(\tau\mathcal{A})\ \Delta u^{(1)}_{t+\tau}\,.\end{aligned} \tag{3.17}$$

yields the corrections, which are used to compute the solution up to order $k = 3$ and which can subsequently be used for error estimation and the adaptive selection of time steps.

In case a second-order approximation is required, only $u^{(1)}_{t+\tau}$ and $\Delta u^{(1)}_{t+\tau}$ need to be solved for. Note that relation (3.15) requires the computation of $(\mathrm{Id}-\tau\mathcal{A})^2$, where

$$(\mathrm{Id}-\tau\mathcal{A})^2 = \mathrm{Id} - 2\cdot\tau\mathcal{A} + \tau^2\mathcal{A}^2$$

involves second derivatives of $\Delta u^{(2)}_{t+\tau}$, the computation of which can become very costly. Therefore we now consider an alternative semi-discretization scheme, which is A-stable and avoids the computation of $(\mathrm{Id}-\tau\mathcal{A})^2$. A similar result to Theorem 3.1.1 can be found for the construction of A-stable approximations:

Theorem 3.1.2 (**A family of *A*-stable rational approximations**,[9, Chapter 2])**.** *The discrete evolution defined in* (3.4) *with* $R^{(k)}_\tau := r^{(k)}_A(\tau\mathcal{A})$ *and*

$$r^{(k)}_A(z) = L_0(1) + \sum_{i=1}^{k} \frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}L_i(1)\left(\frac{z}{1-z}\right)^i\,, \tag{3.18}$$

has at least consistency order k. Moreover,

(i) $r^{(k)}_A(z)$ *is A-stable, and*

(ii) $r^{(k)}_A(z)$ *can be computed recursively if* $\frac{\mathrm{d}}{\mathrm{d}x}L_i(1) \neq 0$ *for* $2 \le i \le k-1$*:*

$$\begin{aligned} r^{(1)}_A(z) &:= \frac{1}{1-z}\\ r^{(i)}_A(z) &:= r^{(i-1)}_A(z) + \vartheta^{(i-1)}_A(z)\,,\quad i = 2,\dots,k\,,\end{aligned} \tag{3.19}$$

where

$$\begin{aligned} \vartheta^{(1)}_A(z) &:= -\frac{1}{2}\cdot\frac{z^2}{1-z}\cdot r^{(1)}_A(z),\\ \vartheta^{(i)}_A(z) &:= -\gamma^{(i)}_A\cdot\frac{z}{1-z}\cdot\vartheta^{(i-1)}_A(z)\,,\quad \textit{with}\quad \gamma^{(i)}_A := \frac{L_{i+1}(1)}{L_i(1)}\,,\quad i = 2,\dots,k\,.\end{aligned} \tag{3.20}$$

As previously for the L-stable scheme, we illustrate the derivation of the stationary spatial problems by means of the third-order A-stable solution $u^{(3)}$ obtained with $R_\tau := r_A^{(3)}$ with $r_A^{(A)}$ according to Theorem 3.1.2, i.e.

$$r_A^{(3)}(z) := \frac{1}{1-z} \cdot \left(1 - \frac{1}{2}\frac{z^2}{1-z} + \frac{1}{6}\frac{z^3}{(1-z)^2}\right) . \tag{3.21}$$

The first-order solution $u_{t+\tau}^{(1)}$ and its correction $\Delta u_{t+\tau}^{(0)}$ are computed as before. The corrections $\Delta u_{t+\tau}^{(i)}$, $i = 1, 2$, are computed recursively by

$$\begin{aligned} \Delta u_{t+\tau}^{(1)} &= -\frac{1}{2}(\mathrm{Id} - \tau\mathcal{A})^{-2}(\tau\mathcal{A})^2 \; u_t \\ &= -\frac{1}{2}(\mathrm{Id} - \tau\mathcal{A})^{-1}(\tau\mathcal{A}) \; \Delta u_{t+\tau}^{(0)} \\ \Rightarrow \quad (\mathrm{Id} - \tau\mathcal{A}) \; \Delta u_{t+\tau}^{(1)} &= -\frac{1}{2}(\tau\mathcal{A}) \; \Delta u_{t+\tau}^{(0)} \end{aligned} \tag{3.22}$$

and

$$\begin{aligned} \Delta u_{t+\tau}^{(2)} &= \frac{1}{6}(\mathrm{Id} - \tau\mathcal{A})^{-1}(\tau\mathcal{A})(\mathrm{Id} - \tau\mathcal{A})^{-2}(\tau\mathcal{A})^2 \; u_t \\ &= -\frac{1}{3}(\mathrm{Id} - \tau\mathcal{A})^{-1}(\tau\mathcal{A}) \; \Delta u_{t+\tau}^{(1)} \\ \Rightarrow \quad (\mathrm{Id} - \tau\mathcal{A}) \; \Delta u_{t+\tau}^{(2)} &= -\frac{1}{3}(\tau\mathcal{A}) \; \Delta u_{t+\tau}^{(1)} . \end{aligned} \tag{3.23}$$

Therefore, the stationary spatial problems are

$$\begin{aligned} (\mathrm{Id} - \tau\mathcal{A}) \; u_{t+\tau}^{(1)} &= u_t \\ (\mathrm{Id} - \tau\mathcal{A}) \; \Delta u_{t+\tau}^{(1)} &= -\frac{1}{2}(\tau\mathcal{A}) \; \Delta u_{t+\tau}^{(0)} \\ (\mathrm{Id} - \tau\mathcal{A}) \; \Delta u_{t+\tau}^{(2)} &= -\frac{1}{3}(\tau\mathcal{A}) \; \Delta u_{t+\tau}^{(1)} . \end{aligned} \tag{3.24}$$

In case a second-order approximation is required, only the first two problems of (3.24) need to be solved. Note that the A-stable methods defined by $r_A^{(2)}(z)$ and $r_A^{(3)}(z)$ do not require any computation of second-order derivatives.

The above results inform how to construct L- and A-stable semi-discretization schemes of a specified consistency order k, which imply convergence of the discrete evolution to the analytical solution of order k. Furthermore, the recursions given in Theorem 3.1.1 and 3.1.2 allow for a multiplicative computation of the corrections, i.e., the difference between two solutions of different order. In the following, the corrections will be used to adjust the time steps in each integration step.

3.2 Adaptive time step control & spatial perturbations

In the previous section we assumed $\tau > 0$ to be fixed. In the adaptive setting, τ_j is adjusted in each integration step t_j. The procedure is analogous to the adaptive time step selection

for ODEs, see Part A of the Appendix or [25, 30]. Ideally the local temporal error, i.e., the error made in the current integration step, remains below a predefined tolerance TOL_t, thus

$$\left|\left|\varepsilon_{t_j}(\tau_j)\right|\right| = \left|\left|u_{t_j+\tau_j} - u^{(k)}_{t_j+\tau_j}\right|\right| \leq TOL_t\,.$$

Since u_t is unknown for $t > 0$, local errors are estimated by comparing solutions of different consistency orders

$$\varepsilon^{(k-1)}_{t_j}(\tau_j) := u^{(k)}_{t_j+\tau_j} - u^{(k-1)}_{t_j+\tau_j} = \Delta u^{(k-1)}_{t_j+\tau_j}\,, \tag{3.25}$$

and an approximation $u^{(k)}_{t_j+\tau_j}$ is accepted if

$$\left|\left|\varepsilon^{(k-1)}_{t_j}(\tau_j)\right|\right| \leq TOL_t\,. \tag{3.26}$$

If $||\varepsilon^{(k-1)}_{t_j}(\tau_j)|| > TOL_t$, then τ_j is decreased in order to obtain a more accurate solution. Otherwise, a time step τ_{j+1} for the following integration step is suggested. In both cases a new time step τ^* is suggested by

$$\tau^* = \sqrt[k]{\frac{TOL_t}{||\varepsilon^{(k-1)}_{t_j}(\tau_j)||}} \cdot \tau_j\,. \tag{3.27}$$

Consequences of spatial perturbations

So far we have assumed exact solutions of the stationary spatial problems. Now we focus on how time steps have to be controlled in the presence of perturbations caused by spatial discretization. Let

$$\hat{u}^{(i)}_{t_j} = u^{(i)}_{t_j} + \delta^{(i)}_{t_j}\,, \quad i = 1, \ldots, k \tag{3.28}$$

denote the spatially perturbed solutions of order i. The spatial errors, $\delta^{(i)}_{t_j}$ comprise approximation errors as well as the propagation of previous approximation errors through the recursion (3.17) or (3.24), respectively. In Part B of the Appendix it is shown how the spatial error estimates $[\delta^{(i)}_{t_j}]$ can be derived from the error estimates provided by the spatial discretization scheme.

We aim at approximating the true solution such that the error in one integration step remains below a specified tolerance, i.e.

$$\left|\left|u_{t_j+\tau_j} - \hat{u}^{(k)}_{t_j+\tau_j}\right|\right| \leq TOL\,. \tag{3.29}$$

This error comprises the temporal and spatial error, i.e.

$$\begin{aligned}\left|\left|u_{t_j+\tau_j} - \hat{u}^{(k)}_{t_j+\tau_j}\right|\right| &\leq \left|\left|u_{t_j+\tau_j} - u^{(k)}_{t+\tau_j}\right|\right| + \left|\left|u^{(k)}_{t_j+\tau_j} - \hat{u}^{(k)}_{t_j+\tau_j}\right|\right|\,, \\ &= \left|\left|\varepsilon_{t_j}(\tau_j)\right|\right| + \left|\left|\delta^{(k)}_{t_j+\tau_j}\right|\right|\,. \end{aligned} \tag{3.30}$$

We demand that temporal and spatial errors each remain below tolerances TOL_t and TOL_x that satisfy

$$TOL_t + TOL_x \leq TOL\,.$$

With only spatially perturbed solutions available, the purely temporal error is not accessible. The spatially perturbed error estimates

$$\hat{\varepsilon}_{t_j}^{(i)}(\tau_j) = \left|\left|\hat{u}_{t_j+\tau_j}^{(i+1)} - \hat{u}_{t_j+\tau_j}^{(i)}\right|\right| , \qquad i = 0, \dots, k-1 , \tag{3.31}$$

differ from the unperturbed estimates by

$$\delta_{\varepsilon}^{(i)} := \hat{\varepsilon}_{t_j}^{(i)}(\tau_j) - \varepsilon_{t_j}^{(i)}(\tau_j) , \qquad i = 0, \dots, k-1 . \tag{3.32}$$

The spatial perturbations $\delta_{\varepsilon}^{(i)}$ of the temporal error estimates can also be estimated from the error estimates provided by the spatial discretization scheme, see Appendix, Part B. Therefore, although the unperturbed temporal error is inaccessible, the contribution of the spatial error can be monitored, such that the adaptive selection of time steps is not decisively compromised by the spatial errors.

Let $[\delta_{\varepsilon}^{(i)}]$, $i = 0, \dots, k-1$, denote the corresponding error estimates. The following algorithm gives constraints for the temporal and spatial accuracy in each integration step, which provide a basis for the decision how to adapt the temporal and spatial discretization.

Algorithm 3.2.1 (**Adaptive integration**, [8]). *Suppose a local tolerance $TOL > 0$ is specified. Semi-discretization in time is performed as described above with*

$$TOL_t = \rho \cdot TOL , \quad 0 < \rho < 1 \tag{3.33}$$

and a solution $\hat{u}_{t_j+\tau_j}^{(k)}$ is accepted, if the following two conditions are satisfied:

$$\left|\left|\hat{\varepsilon}_{t_j}^{(k-1)}(\tau_j)\right|\right| + \left|\left|[\delta_{t_j+\tau_j}^{(k)}]\right|\right| \leq TOL \tag{3.34}$$

$$\left|\left|[\delta_{\varepsilon}^{(k-1)}]\right|\right| \leq \frac{\left|\left|\hat{\varepsilon}_{t_j}^{(k-1)}(\tau_j)\right|\right|}{4} . \tag{3.35}$$

Remark 3.2.2 (**Choice of the parameter ρ**). *In [8], it is further suggested to determine the temporal and spatial tolerance TOL_t and TOL_x by choosing*

$$\rho = \frac{1}{d+1}$$

constant throughout integration. However, we will see later, in Chapter 6, that in the case considered herein, a coupling of TOL_x with the time step τ is necessary to ensure convergence of the overall numerical scheme.

Chapter 4

Approximate approximations

In this chapter we introduce the spatial discretization technique that will later be used to solve the stationary spatial problems within the adaptive Rothe scheme that was described in the previous chapter. Approximate approximations were first introduced by Maz'ya in the early 1990s [64, 63]. Since then they have found a number of applications, for example in the approximation of pseudodifferential operators [65], the solution of boundary [68] and time dependent initial-boundary value problems [47, 48] and the approximation of potentials [44]. A review of applications can be found in [81]. For a detailed survey on approximate approximations, see [66].

The concept is based on a basis expansion, where the basis functions are obtained by shifting and scaling a generating function that has to satisfy two conditions. Then, although the basis functions are not necessarily orthogonal, the coefficients are explicitly computable. As a trade-off, the approximation error does not fully vanish when the discretization is refined, but reaches a saturation value. However, a precise description of the saturation error is available and allows to select the scaling parameter such that the error becomes arbitrarily small and thus negligible in practical applications.

Generating functions can be constructed such that the action of differential operators on the approximate approximant can in many cases be computed analytically, see e.g. [47, 48, 65, 68]. Furthermore, a solid theory allows for the construction of approximants with high approximation order. These features make the method particularly attractive for the use in an adaptive Rothe context.

We start by presenting the general concept in Section 4.1 and show an alternative derivation by means of kernel regression in Section 4.2. Asymptotic properties of the approximation error are discussed in Section 4.3, where we will also see that a truncation of the summation in the basis expansion yields similar approximation quality. The construction of high-order approximants is then described in Section 4.4. Finally, in Section 4.5, we describe how the approximation error can be estimated and how this information can be used for an adaptive refinement of the discretization.

4.1 Sums of shifted and scaled basis functions

Suppose a function $u : \mathbb{R}^d \to \mathbb{R}$ has to be approximated. The idea of approximate approximations is based on a representation of the function u as the weighted sum of a shifted

and scaled generating function η, where η has to satisfy two conditions: fast decay and vanishing moments. The function u is then approximated by

$$u(x) \approx \mathcal{M}_{h,\mathcal{D}} u(x) := \mathcal{D}^{-d/2} \cdot \sum_{m \in \mathbb{Z}^d} u(mh) \cdot \eta\left(\frac{x - mh}{\sqrt{\mathcal{D}}h}\right), \quad \mathcal{D} > 0\,, \tag{4.1}$$

where $\{mh, m \in \mathbb{Z}^d\}$ forms a uniform grid on $\mathbb{R}^d$ with grid size $h > 0$. The shifted and scaled function η generates a basis

$$\left\{\eta\left(\frac{x - mh}{\sqrt{\mathcal{D}}h}\right),\ m \in \mathbb{Z}^d\right\}$$

for all approximants of the form (4.1). The parameter $\mathcal{D}$ scales the decay speed of the basis functions and thereby their *width*. (Consider e.g. Gaussian basis functions, where $\mathcal{D}$ corresponds to the variance.)

The approximation error is bounded by

$$||u - \mathcal{M}_{h,\mathcal{D}} u|| \leq c(||u||) \cdot h^M + \varepsilon_{\text{sat}}(||u||\,, \mathcal{D})\,, \quad \text{as} \quad h \to 0\,, \tag{4.2}$$

i.e., the error decays with order M and saturates at ε_{sat}. We will see later, in Section 4.3, that the approximation order M depends on the smoothness of the function u and the number of vanishing moments of η, whereas the saturation error is determined by the scaling parameter $\mathcal{D}$. Since the approximation error does not fully vanish, $\mathcal{M}_{h,\mathcal{D}} u$ is called *approximate* approximation of u.

To motivate the above approximation formula (4.1), let us consider $d = 1$ and a sum of Gaussian basis functions with variance $\sigma^2 = \mathcal{D}/2$, i.e.

$$f_{\mathcal{D}}(x) = \sum_{m \in \mathbb{Z}} e^{-(x-m)^2/\mathcal{D}}\,. \tag{4.3}$$

It can be shown that for all $x \in \mathbb{R}^d$, $f_{\mathcal{D}}(x) \approx \sqrt{\pi\mathcal{D}}$, oscillating around the value, and with growing $\mathcal{D}$ the amplitude decreases [66]. Consequently,

$$\bar{f}_{\mathcal{D}}(x) := \frac{1}{\sqrt{\pi\mathcal{D}}} \cdot f_{\mathcal{D}}(x) \approx 1\,, \qquad \forall\, x \in \mathbb{R}\,,$$

which is shown in Figure 4.1 for $\mathcal{D} = 0.4$ (left), $\mathcal{D} = 0.5$ (middle) and $\mathcal{D} = 2$ (right). It can be seen that the amplitude of the oscillations rapidly decreases as $\mathcal{D}$ is increased; for $\mathcal{D} = 2$, the function already appears constant.

A sum of smooth, nonnegative functions is called a *partition of unity* on $\mathbb{R}$, if for all $x \in \mathbb{R}$ the number of functions with support at x is finite and further the sum is identical to one. Partitions of unity can be used to globally describe functions that are defined or known only locally. Considering $\bar{f}_{\mathcal{D}}$, the support of the Gaussian basis functions is infinite. However, since for all $x \in \mathbb{R}$ and any $\delta > 0$, only a finite number of summands exceeds δ, and since further $\bar{f}_{\mathcal{D}}$ approximates one at any point $x \in \mathbb{R}$, it forms an *approximate* partition of unity on $\mathbb{R}$. As a consequence, the approximate approximant

$$\mathcal{M}_{h,\mathcal{D}} u(x) = \frac{1}{\sqrt{\pi\mathcal{D}}} \cdot \sum_{m \in \mathbb{Z}} u(mh) \cdot e^{(x-mh)^2/(\mathcal{D}h^2)}\,. \tag{4.4}$$

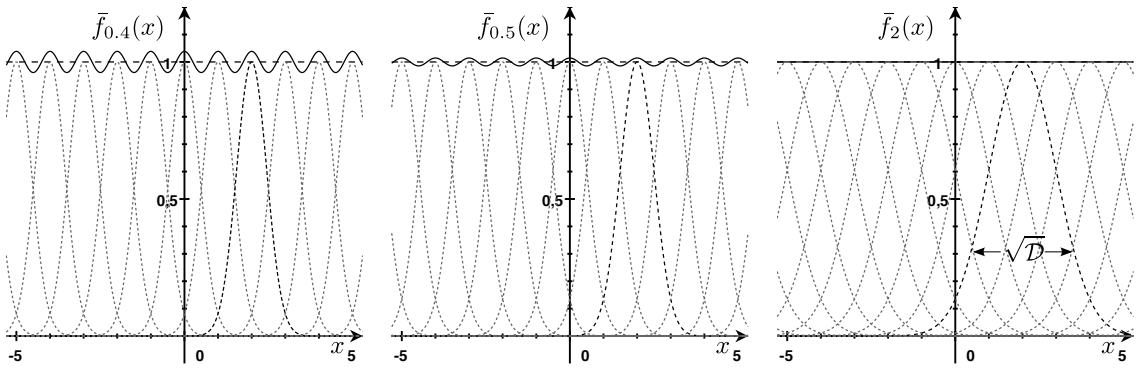

Figure 4.1: Approximate partition of unity: the function $\bar{f}_{\mathcal{D}}(x)$ (solid line), a sum of Gaussian basis functions (dotted grey), oscillates around one (dashed). The amplitude of the oscillations rapidly decreases as $\mathcal{D}$ is increased, shown for $\mathcal{D} = 0.4$, $\mathcal{D} = 0.5$ and $\mathcal{D} = 2$ (left to right), which determine the width of the basis functions.

yields a smooth approximation of the function u, without interpolating the values $u(mh)$, $m \in \mathbb{Z}$. Instead, the approximating function oscillates around the values $u(mh)$ and is therefore also referred to as an *approximate quasi-interpolation*.

An alternative, perhaps more intuitive, approach to deriving (4.4) is by considering the *Nadaraya-Watson* kernel regression estimator as shown next.

4.2 Derivation from kernel regression

Kernel regression aims at finding a functional dependence of two random variables X and Y and relies on the estimation of the corresponding density functions. An introduction to kernel regression can be found e.g. in [34], and to density estimation in [88, 89]. Since X and Y are random, dependence is captured by the *conditional expectation* of Y given X

$$\mathbb{E}[Y|X] = \int y \cdot f(y|X) \, \mathrm{d}y = u(X) \,, \tag{4.5}$$

where $f(y|X)$ denotes the probability density function of Y conditional on X. As $f(y|X)$ can be expressed by the joint probability density function $f(x, y)$ and the marginal density of X, $f_X(x)$, (4.5) becomes

$$\mathbb{E}[Y|X] = \int y \, \frac{f(x,y)}{f_X(x)} \, \mathrm{d}y \,. \tag{4.6}$$

Estimating the conditional expectation can be regarded as a function approximation from random data points. The marginal and joint probability density functions $f_X(x)$ and $f(x, y)$ are estimated and subsequently used to approximate the integral in (4.6).

Suppose a random sample $\{(x_1, y_1), \ldots, (x_N, y_N)\}$ drawn from the joint probability distribution of X and Y is given. A natural approach to estimating f_X at a point x_0 is to count the sample points in the vicinity of x_0 and divide this number by the total number N of sample points, i.e.

$$\hat{f}_X(x_0) = \frac{|\mathcal{N}_\lambda(x_0)|}{N\lambda} \,. \tag{4.7}$$

Here, $\hat{f}$ denotes the estimate of f, $\mathcal{N}_\lambda(x_0)$ the set of points in some *neighbourhood* of x_0 and $|\mathcal{N}_\lambda(x_0)|$ the number of points in $\mathcal{N}_\lambda(x_0)$. The size or *width* of the neighbourhood is specified by the parameter λ. Clearly, as λ is increased, $|\mathcal{N}(x_0)|$ approaches N for all x_0. Division by λ in (4.7) compensates for the whole term approaching one.

The estimator (4.7) is not continuous in x_0. As continuity is a desirable feature of density estimators, we replace the discontinuous definition of neighbourhood in (4.7) by a smooth one. Instead of counting the points in the vicinity of x_0 we specify a continuous *kernel* function $K_\lambda(x_0, x)$ of width λ that assigns a weight to any point x. A kernel is any positive function with $K_\lambda(x, y) = K_\lambda(y, x)$ for all x, y, and $\int K_\lambda(x_0, x)\mathrm{d}x = 1$ for all x_0.

Typically, $K_\lambda(x_0, x)$ only depends on the distance between x_0 and x. Those kernel functions $K_\lambda(x_0, x) = \varphi_\lambda(x - x_0)$ are called *radial* kernels or *radial basis functions*. Common kernel functions are shown in Figure 4.2.

Figure 4.2: From left to right: triangular, Epanechnikov, and Gaussian kernel.

The discontinuous estimator $\hat{f}_X$ in (4.7) can then be replaced by

$$\hat{f}_X(x_0) = \frac{1}{N\lambda} \cdot \sum_{i=1}^{N} \varphi_\lambda(x_0 - x_i) \ , \tag{4.8}$$

which is called the *Parzen* estimator [72]. Analogously, the joint probability density function $f(x, y)$ can be estimated by

$$\hat{f}(x, y) = \frac{1}{N\lambda^2} \cdot \sum_{i=1}^{N} \varphi_\lambda(x - x_i) \cdot \varphi_\lambda(y - y_i) \ . \tag{4.9}$$

Applying (4.8) and (4.9) to estimate the densities in (4.6) yields

$$\begin{aligned}
\int y \frac{\hat{f}(x, y)}{\hat{f}_X(x)} \mathrm{d}y &= \frac{N\lambda}{\sum_{i=1}^{N} \varphi_\lambda(x - x_i)} \cdot \frac{1}{N\lambda} \cdot \sum_{i=1}^{N} \varphi_\lambda(x - x_i) \cdot \int \frac{y}{\lambda} \cdot \varphi_\lambda(y - y_i) \ \mathrm{d}y \\
&= \frac{N\lambda}{\sum_{i=1}^{N} \varphi_\lambda(x - x_i)} \cdot \frac{1}{N\lambda} \cdot \sum_{i=1}^{N} \varphi_\lambda(x - x_i) \cdot \int (s\lambda + y_i) \cdot \eta(s) \ \mathrm{d}s \\
&= \frac{N\lambda}{\sum_{i=1}^{N} \varphi_\lambda(x - x_i)} \cdot \frac{1}{N\lambda} \cdot \sum_{i=1}^{N} \varphi_\lambda(x - x_i) \cdot y_i \, .
\end{aligned}$$

For the first equality, we substituted $s = \frac{y - y_i}{\lambda}$, and for the second equality, we used that φ is symmetric around 0 and integrates to 1. Finally, we obtain the *Nadaraya-Watson* estimator [69, 90]:

$$\hat{u}_\lambda(x) = \frac{N^{-1} \sum_{i=1}^{N} y_i \cdot \varphi_\lambda(x - x_i)}{N^{-1} \sum_{i=1}^{N} \varphi_\lambda(x - x_i)} \ . \tag{4.10}$$

We can identify the Nadaraya-Watson kernel regression estimator (4.10) with the approximate approximation (4.4), if:

1. The radial Kernel function φ denotes a Gaussian Kernel.
2. The width λ in (4.10) corresponds to $\mathcal{D}$ in (4.4).
3. The random sample $\{(x_i, y_i)\}_{i=1,\dots,N}$ is identified with the equally spaced points $\{(mh, u(mh)\}_{m\in\mathbb{Z}^d}$. While the number of sample points N is finite, the number of grid points in (4.4) is infinite. However, for the denominator in (4.10) the limit

$$\lim_{N\to\infty} \sum_{i=1}^{N} \varphi_\lambda (x - x_i)$$

exists for all x, oscillating in x around $\frac{1}{\sqrt{\pi\mathcal{D}}}$.

4.3 Asymptotics of the approximation error

In this section we consider asymptotics of the approximation error of approximate approximations with respect to h and $\mathcal{D}$. Before stating the corresponding theorems, we need to introduce some definitions and specify the two conditions on the generating function η. Then, in Section 4.3.1, error bounds are given for approximate approximants of the form (4.1). These error bounds account for approximate approximations with infinite sums. Section 4.3.2 gives similar error bounds for approximate approximations where the summation is truncated. The proofs of all theorems in this and the following section can be found in [66].

Let us denote by $\boldsymbol{\alpha} \in \mathbb{N}^d$ a multi-index of length $|\boldsymbol{\alpha}| := \alpha_1 + \ldots + \alpha_d$. Further we set $\boldsymbol{\alpha}! := \alpha_1 \cdot \ldots \cdot \alpha_d$, $x^{\boldsymbol{\alpha}} := x_1^{\alpha_1} \cdot \ldots \cdot x_d^{\alpha_d}$,

$$\partial^{\boldsymbol{\alpha}} u(x) := \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial x_1^{\alpha_1} \dots x_d^{\alpha_d}} u(x)\,,$$

and

$$\nabla_k u := (\partial^{\boldsymbol{\alpha}} u)_{|\boldsymbol{\alpha}|=k}\,.$$

Sobolev spaces denote classes of functions that are closely related to $\mathcal{L}_p$-spaces.

Definition 4.3.1 (**Sobolev spaces**). *The Sobolev space $W_p^L(\mathbb{R}^d)$, $L \in \mathbb{N}$, denotes the space of all functions $u \in \mathcal{L}_p(\mathbb{R}^d)$ whose* generalized *or* weak *derivatives $\partial^{\boldsymbol{\alpha}} u$ up to order $|\boldsymbol{\alpha}| = L$ also lie in $\mathcal{L}_p(\mathbb{R}^d)$, i.e.*

$$W_p^L(\mathbb{R}^d) := \left\{ u \in \mathcal{L}_p(\mathbb{R}^d) :\ \partial^{\boldsymbol{\alpha}} u \in \mathcal{L}_p(\mathbb{R}^d)\,,\ \textit{for all } \boldsymbol{\alpha} \textit{ with } 0 \le |\boldsymbol{\alpha}| \le L \right\}.$$

The conditions that the generating function η has to satisfy are the following:

Condition 4.3.2 (**Decay Condition**). *Suppose that for all $\boldsymbol{\alpha} \in \mathbb{N}^d$, $0 \le |\boldsymbol{\alpha}| \le \lfloor \frac{d}{2} \rfloor + 1$, the partial derivatives $\partial^{\boldsymbol{\alpha}}\eta$ are continuous. A function $\eta : \mathbb{R}^d \to \mathbb{R}$ is said to satisfy the decay condition, if there exists an $A > 0$ and $K > d$ such that*

$$|\partial^{\boldsymbol{\alpha}}\eta(x)| \le A \cdot (1 + |x|_2)^{-K}\,, \quad \forall x \in \mathbb{R}^d\,.$$

In other words, the Decay Condition requires η, and its continuous derivatives up to order $(\lfloor \frac{d}{2} \rfloor + 1)$, to decay faster than $A \cdot (1 + |x|)^{-K}$, for some constants A and K.

Condition 4.3.3 (**Moment Condition**). *A function $\eta : \mathbb{R}^d \to \mathbb{R}$ is said to satisfy the moment condition of order N, if*

$$(i) \qquad \int_{\mathbb{R}^d} \eta(x)\, \mathrm{d}x = 1\,, \quad \textit{and}$$

$$(ii) \qquad \int_{\mathbb{R}^d} x^{\boldsymbol{\alpha}} \cdot \eta(x)\, \mathrm{d}x = 0\,, \quad \forall \boldsymbol{\alpha},\ 1 \leq |\boldsymbol{\alpha}| < N\,.$$

Thus, all moments up to the $(N-1)$-th moment vanish except for the 0-th moment, which is equal to one.

Remark 4.3.4 (**Gaussian generating functions**). *A Gaussian generating function is infinitely continuously differentiable, and all its derivatives decay exponentially, i.e. faster than any order $K > 0$. Further, if the mean is zero, it satisfies the Moment Condition of order $N = 2$.*

4.3.1 The approximation error on infinite grids

We can now state a pointwise result for the approximation error of approximate approximants of the form (4.1).

Theorem 4.3.5 (**Pointwise error estimate for approximate approximations**, [66, Theorem 2.17]). *Assume $u \in W^L_\infty(\mathbb{R}^d)$. Let further η satisfy the Decay Condition with decay exponent K and the Moment Condition of order N, where $K > N + d$. Then for any $\varepsilon > 0$ there exists $\mathcal{D}' > 0$ such that for all $\mathcal{D} > \mathcal{D}'$ and $h > 0$ the approximation error can be bounded* pointwise *by*

$$|u(x) - \mathcal{M}_{h,\mathcal{D}} u(x)| \leq c_\eta \cdot \left(\sqrt{\mathcal{D}} h\right)^M \cdot ||\nabla_M u||_{\mathcal{L}_\infty} + \varepsilon \cdot \sum_{k=0}^{M-1} \left(\sqrt{\mathcal{D}} h\right)^k \cdot |\nabla_k u(x)|\,, \qquad (4.11)$$

where $M := \min(L, N)$, and the constant c_η is independent of u, h and $\mathcal{D}$.

The first term of the right hand side of (4.11) decays with $\mathcal{O}\left((\sqrt{\mathcal{D}}h)^M\right)$ as h vanishes and is thus of order M in h. The second term,

$$\varepsilon_{\text{sat}} := \varepsilon \cdot \sum_{k=0}^{M-1} \left(\sqrt{\mathcal{D}} h\right)^k \cdot ||\nabla_k u(x)||_{\mathcal{L}_2}\,,$$

is called the *saturation error*. It can be shown that the saturation error has the representation

$$\sum_{j=0}^{M-1} \left(\frac{i\sqrt{\mathcal{D}}h}{2\pi}\right)^j \cdot \sum_{|\boldsymbol{\alpha}|=j} \frac{\partial^{\boldsymbol{\alpha}} u}{\boldsymbol{\alpha}!} \cdot \sum_{m \in \mathbb{Z}^d \setminus 0} \partial^{\boldsymbol{\alpha}} \mathcal{F}\eta\left(\sqrt{\mathcal{D}} m\right) \cdot e^{\frac{2\pi i}{h} \cdot \langle x, m \rangle}\,,$$

where $\mathcal{F}\eta$ denotes the Fourier transform of η and $\langle \cdot, \cdot \rangle$ is the common scalar product. The inner sum consists of fast oscillating functions, which become arbitrarily small for $\mathcal{D}$ sufficiently large [66].

Theorem 4.3.5 conveys that for $\mathcal{D} > 0$ fixed and h vanishing, $\mathcal{M}_{h,\mathcal{D}}u(x)$ approximates any function value $u(x)$, $u \in W_\infty^L$, with order M up to the saturation error. Since for $\mathcal{D}$ sufficiently large, the saturation error becomes arbitrarily small, it means that effectively—i.e. in practical computations—$\mathcal{M}_{h,\mathcal{D}}u(x)$ approximates $u(x)$ with order M. The following theorem provides a similar bound for the global approximation error.

Theorem 4.3.6 (**Global error estimate for approximate approximations**, [66, Theorem 2.28])**.** *Let η satisfy the Decay Condition with decay exponent K and the Moment Condition of order N. Further, assume that $u \in W_p^L(\mathbb{R}^d)$, $1 \leq p \leq \infty$, with $d/p < L < K$. Then for any $\varepsilon > 0$ there exists $\mathcal{D}' > 0$ such that for all $\mathcal{D} > \mathcal{D}'$ and $h > 0$ the approximation error can be bounded by*

$$||u - \mathcal{M}_{h,\mathcal{D}}u||_{\mathcal{L}_p} \leq c_\eta \cdot \left(\sqrt{\mathcal{D}}h\right)^M \cdot ||\nabla_M u||_{\mathcal{L}_p} + \varepsilon \cdot \sum_{k=0}^{M-1} \left(\sqrt{\mathcal{D}}h\right)^k \cdot ||\nabla_k u||_{\mathcal{L}_p} , \tag{4.12}$$

where $M := \min(L, N)$, and the constant c_η is independent of u, h and $\mathcal{D}$.

Consequently, for $\mathcal{D} > 0$ fixed and vanishing h, the global approximation error exhibits similar asymptotics to the pointwise error: a decay with $\mathcal{O}\left((\sqrt{\mathcal{D}}h)^M\right)$ in h and saturation at a value that becomes arbitrarily small for $\mathcal{D}$ sufficiently large.

Remark 4.3.7 (**Upper bound for h**)**.** *In order to guarantee the convergence in the pointwise and global error, a natural bound on $h > 0$ arises as*

$$h < \frac{1}{\sqrt{\mathcal{D}}} ,$$

such that $\sqrt{\mathcal{D}}h < 1$.

4.3.2 Truncation of summation

The approximate approximant $\mathcal{M}_{h,\mathcal{D}}u(x)$ defined in (4.1) uses an infinite sum to approximate a function u at a point $x \in \mathbb{R}^d$. Since the support of the generating function η is generally unbounded, theoretically an infinite number of summands contributes to the value of $\mathcal{M}_{h,\mathcal{D}}u$ at any point $x \in \mathbb{R}^d$. However, in practical applications, the summation can be truncated, since the generating functions are chosen to decay fast [44].

Let $B_\kappa(x) := \left\{y \in \mathbb{R}^d, |y - x|_2 \leq \kappa\right\}$ be the set of points in a closed ball with radius κ around x. The truncated approximant

$$\mathcal{M}_{h,\mathcal{D}}^{(\kappa)}u(x) := \mathcal{D}^{-d/2} \cdot \sum_{\substack{m \in \mathbb{Z}^d \\ mh \in B_\kappa(x)}} u(mh) \cdot \eta\left(\frac{x - mh}{\sqrt{\mathcal{D}}h}\right) \tag{4.13}$$

only takes into account points mh within the neighborhood of x defined by $B_\kappa(x)$. Consequently, the uniform grid $\{mh, mh \in \mathcal{B}_\kappa(x)\}$ is *finite*. The difference of the approximant defined in (4.1) and the truncated approximant can be bounded by

$$\left|\mathcal{M}_{h,\mathcal{D}}^{\kappa}u(x) - \mathcal{M}_{h,\mathcal{D}}u(x)\right| \leq \sup_{\substack{m \in \mathbb{Z}^d \\ mh \notin B_\kappa(x)}} |u(mh)| \cdot \mathcal{D}^{-d/2} \cdot \sum_{\substack{m \in \mathbb{Z}^d \\ mh \notin B_\kappa(x)}} \left|\eta\left(\frac{x - mh}{\sqrt{\mathcal{D}}h}\right)\right| \tag{4.14}$$

$$\leq g_{\mathcal{D}}(\kappa/h, \eta) \cdot ||u||_{\mathcal{L}_\infty} \tag{4.15}$$

with

$$g_{\mathcal{D}}(\zeta,\eta) := \sup_{x\in\mathbb{R}^d} \mathcal{D}^{-d/2} \cdot \sum_{\substack{m\in\mathbb{Z}^d,\\ |x-m|>\zeta}} \left|\eta\left(\frac{x-m}{\sqrt{\mathcal{D}}}\right)\right| . \tag{4.16}$$

Since η satisfies the Decay Condition with constant A and decay order $K > d$, $g_{\mathcal{D}}$ can be bounded by

$$\begin{aligned} g_{\mathcal{D}}(\zeta,\eta) &\leq A \cdot \sup_{x\in\mathbb{R}^d} \mathcal{D}^{-d/2} \cdot \sum_{\substack{m\in\mathbb{Z}^d,\\ |x-m|>\zeta}} \left(1+\frac{|x-m|}{\sqrt{\mathcal{D}}}\right)^{-K} \\ &= A\cdot\mathcal{D}^{-d/2} \cdot \sup_{x\in\mathbb{R}^d} \sum_{\substack{m\in\mathbb{Z}^d,\\ |x-m|>\zeta}} \mathcal{D}^{K/2} \cdot \left(\sqrt{\mathcal{D}}+|x-m|\right)^{-K} \\ &\leq A\cdot\mathcal{D}^{(K-d)/2} \cdot \sup_{x\in\mathbb{R}^d} \sum_{\substack{m\in\mathbb{Z}^d,\\ |x-m|>\zeta}} |x-m|^{-K} \\ &= A\cdot\mathcal{D}^{(K-d)/2} \cdot \sup_{x\in\mathbb{R}^d} \sum_{j=0}^{\infty} \sum_{\substack{m\in\mathbb{Z}^d,\\ \zeta+j<|x-m|\leq\zeta+j+1}} |x-m|^{-K} . \end{aligned}$$

Since the number of integers $m \in \mathbb{Z}^d$, for which $j \leq |x-m| \leq j+1$, is bounded by $\hat{C}\cdot(j)^{d-1}$ with $\hat{C} > 0$, this becomes

$$\begin{aligned} g_{\mathcal{D}}(\zeta,\eta) &\leq A\cdot\mathcal{D}^{(K-d)/2}\cdot\hat{C}\cdot\sum_{j=0}^{\infty}(\zeta+j)^{d-1-K} \\ &\leq C\cdot\mathcal{D}^{(K-d)/2}\cdot\zeta^{d-K} = C\cdot\left(\frac{\zeta}{\sqrt{\mathcal{D}}}\right)^{d-K} \end{aligned}$$

for a constant $C > 0$, which depends on η and d. Hence, using (4.15) the difference between the truncated and the non-truncated approximant can be bounded by

$$\left|\mathcal{M}^{\kappa}_{h,\mathcal{D}}u(x) - \mathcal{M}_{h,\mathcal{D}}u(x)\right| \leq C\cdot\left(\frac{\sqrt{\mathcal{D}}h}{\kappa}\right)^{K-d} \cdot ||u||_{\mathcal{L}_\infty} . \tag{4.17}$$

If κ is proportional to h, so $\kappa = \nu h$ with $\nu > 0$, h cancels out in (4.17), and thus, the bound is independent of h. The truncated approximate approximant then only considers terms for which $|x/h - m| \leq \nu$, i.e. the number of summands is independent of h. The approximation error of such truncated approximate approximations can be bounded as follows.

Corollary 4.3.8 (**Pointwise error estimate for truncated approximate approximations**, [66, Corollary 2.20])**.** *Assume $u \in W^L_\infty$. Let η satisfy the Decay Condition and the Moment Condition of order N. Then for any $\varepsilon > 0$ there exist $\mathcal{D}' > 0$ and $\nu > 0$ such that for all $\mathcal{D} > \mathcal{D}'$, $h > 0$ and $\kappa \geq \nu h$*

$$\begin{aligned} \left|u(x) - \mathcal{M}^{(\kappa)}_{h,\mathcal{D}}u(x)\right| \leq\ & c_\eta\cdot\left(\sqrt{\mathcal{D}}h\right)^M \cdot ||\nabla_M u||_{\mathcal{L}_\infty} + \\ & \varepsilon\cdot\left(\sum_{k=0}^{M-1}\left(\sqrt{\mathcal{D}}h\right)^k\cdot|\nabla_k u(x)| + ||u||_{\mathcal{L}_\infty}\right) \end{aligned} \tag{4.18}$$

for all $x \in \mathbb{R}^d$, where $M := \min(L, N)$, and the constant c_η only depends on η.

As compared to the pointwise error (4.11) of the non-truncated approximant, the saturation error in (4.18) contains an extra term $\varepsilon \cdot ||u||_{\mathcal{L}_\infty}$. However, the first term of (4.18) still decays with order $\mathcal{O}\left((\sqrt{\mathcal{D}}h)^M\right)$ for $\mathcal{D}$ fixed and $h \to 0$.

A similar bound is given for the global error of the truncated approximant on a domain $\Omega \subset \mathbb{R}^d$:

Corollary 4.3.9 (**Global error estimate for truncated approximate approximations**, [66, Lemma 2.30])**.** *Let η satisfy the Decay Condition and the Moment Condition of order N. Furthermore, let $\Omega \subset \mathbb{R}^d$ and $u \in W_p^L(\Omega)$, $1 \leq p \leq \infty$ with $d/p < L$. Then for any $\varepsilon > 0$ there exists a $\mathcal{D}' > 0$ and $\kappa > 0$ such that for all $\mathcal{D} > \mathcal{D}'$ and $h > 0$*

$$\left|\left|u - \mathcal{M}_{h,\mathcal{D}}^{(\kappa)} u\right|\right|_{\mathcal{L}_p(\Omega_{\kappa h})} \leq c_\eta \cdot \left(\sqrt{\mathcal{D}}h\right)^M \cdot ||\nabla_M u||_{\mathcal{L}_p(\Omega)} + \underbrace{\varepsilon \cdot \sum_{k=0}^{M-1} \left(\sqrt{\mathcal{D}}h\right)^k ||\nabla_k u||_{\mathcal{L}_p(\Omega)}}_{=:\varepsilon_{\text{sat}}}, \quad (4.19)$$

where $\Omega_{\kappa h} := \{x,\ B_{\kappa h}(x) \subset \Omega\}$, and the constant c_η is independent of u, h and $\mathcal{D}$.

Note that the norms in the global bound (4.19) are restricted to the domain $\Omega \subset \mathbb{R}^d$. From Theorems 4.3.8 and 4.3.9 we conclude that the truncated approximate approximant (4.13) has similar approximation quality to the non-truncated approximant. In fact, the approximation order M is conserved (up to the saturation error).

4.4 Construction of high-order approximants

Functions that satisfy the Moment Condition 4.3.3 of arbitrary order can be constructed from other functions satisfying the Decay Condition 4.3.2. Using those functions as generating functions yields high-order approximate approximations. This is shown in detail in [66] and more briefly for radial generating functions in [28]. Here we concentrate on radial generating functions. The following theorem shows how to construct generating functions that yield high-order approximate approximations.

Theorem 4.4.1 (**Construction of generating functions for high-order approximants**, [66, Theorem 3.5])**.** *Let*

$$\mathscr{L}_k^\gamma(x) := \frac{x^{-\gamma}}{k!} \cdot e^x \cdot \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k \left(x^{k+\gamma} e^{-x}\right), \quad \gamma > -1\,, \quad (4.20)$$

denote the generalized Laguerre polynomials*. A d-dimensional approximant of the form* (4.1) *with generating function*

$$\eta^{(2M)}(x) := \pi^{-d/2} \cdot \mathscr{L}_{M-1}^{(d/2)}(|x|_2^2) \cdot e^{-|x|_2^2}\,, \quad (4.21)$$

has approximation order $2M$.

Example 4.4.2. *Suppose $x \in \mathbb{R}$, so $d = 1$. For $M = 1, 2, 3$, the generalized Laguerre polynomials are*

$$\begin{aligned}
\mathscr{L}_0^{1/2}(x) &\equiv 1\,, \\
\mathscr{L}_1^{1/2}(x) &= \frac{3}{2} - x\,, \\
\mathscr{L}_2^{1/2}(x) &= \frac{1}{2} \cdot \left(\frac{15}{4} - 5x + \frac{x^2}{2} \right).
\end{aligned}$$

The corresponding generating functions are given by

$$\eta^{(2)}(x) = \frac{1}{\sqrt{\pi}} \cdot e^{-x^2}\,, \tag{4.22}$$

$$\eta^{(4)}(x) = \frac{1}{\sqrt{\pi}} \cdot \left(\frac{3}{2} - x^2 \right) \cdot e^{-x^2}\,, \tag{4.23}$$

$$\eta^{(6)}(x) = \frac{1}{\sqrt{\pi}} \cdot \frac{1}{2} \cdot \left(\frac{15}{4} - 5x^2 + \frac{x^4}{2} \right) \cdot e^{-x^2}\,. \tag{4.24}$$

Note that $\eta^{(2)}$ is a Gaussian generating function with variance $\sigma^2 = 1/4$. The three generating functions are shown in Figure 4.3.

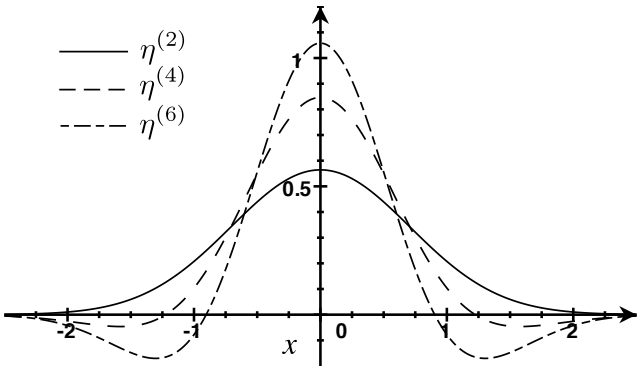

Figure 4.3: Generating functions for approximate approximations of order 2,4 and 6.

□

Besides high approximation orders, approximate approximations that are constructed according to Theorem 4.4.1 possess another desirable property. The structure of the generating function allows us to obtain analytical expressions for the action of the infinitesimal generator $\mathcal{A}$ of the semigroup of Frobenius-Perron operators on the approximate approximant. These are shown in the Appendix C (where also explicit formulas for $\eta^{(2)}$, $\eta^{(4)}$ and $\eta^{(6)}$ in the general case of $d \geq 1$ are given). This fact will be exploited later in Chapter 5, when approximate approximations are used to solve the stationary spatial problems within an adaptive Rothe scheme. Furthermore, we exploit that approximate approximations straightforwardly provide error estimates without further computations necessary, as will be shown next.

4.5 Readily available error estimates

In the previous sections we have discussed asymptotics of the approximation error. In this section we concentrate on error estimates that, in an adaptive setting, allow us to make use of the asymptotic properties to estimate optimal grid sizes.

Error estimation of many other approximation techniques (finite elements, finite differences etc.) relies on the comparison of two solutions with different approximation orders (similar to the estimation of temporal discretization errors, see Appendix, Part A.2). With approximate approximations we can avoid the computation of a higher order solution. It turns out that their feature of not interpolating points, but *quasi*-interpolating them, allows for particularly accessible error estimates, since

$$\begin{aligned} ||\delta(h)|| = ||u - \mathcal{M}_{h,\mathcal{D}}u||_{\mathcal{L}_p(\Omega)} &= \left(\int_\Omega |u(x) - \mathcal{M}_{h,\mathcal{D}}u(x)|^p \ \mathrm{d}x \right)^{1/p} \\ &\approx \left(h^d \cdot \sum_{\substack{m\in\mathbb{Z}^d, \\ mh\in\Omega}} |u(mh) - \mathcal{M}_{h,\mathcal{D}}u(mh)|^p \right)^{1/p} =: [\delta]\,(h)\,, \end{aligned} \tag{4.25}$$

where the factor h^d arises from the approximation of a d-dimensional integral. This means that the global error can be estimated by a comparison of the coefficients $u(mh)$ with the approximate approximant $\mathcal{M}_{h,\mathcal{D}}u(mh)$ evaluated at the grid points $mh \in \Omega$. Since these values are all readily available, we can avoid the comparison to a higher-order approximant, which is required for interpolation methods (since for those the above estimate is by definition zero).

From the previous sections we further know that for fixed $\mathcal{D} > 0$ and h decreasing, the pointwise and global error estimates exhibit the same asymptotics: a decay with order M and saturation at a value determined by $\mathcal{D}$. Consequently, for a prescribed accuracy $TOL > 0$, the estimate (4.25) allows us to use the asymptotics to estimate an optimal grid size by

$$h^* = \sigma \cdot \sqrt[M]{\frac{TOL}{[\delta](h)}} \cdot h\,, \quad 0 < \sigma < 1\,, \tag{4.26}$$

such that the error estimate satisfies

$$[\delta](h^*) \approx TOL\,.$$

Note that the above procedure is analogous to the adaptive choice of time steps as shown in Part A.2 of the Appendix.

Chapter 5

Adaptive density propagation: A Rothe method using approximate approximations

This and the following chapter constitute the main contribution of the thesis. In this chapter we present a method for the numerical solution of ODEs with random initial values. The evolution of the probability density function associated with the random state variable is described by the linear first-order PDE

$$\frac{\partial}{\partial t}u = \mathcal{A}u = -\mathrm{div}(F \cdot u)\,, \quad u(0,\cdot) = u_0\,. \tag{5.1}$$

The proposed method addresses the problem by numerically solving this PDE. Integration is performed adaptively in both time and space, using the Rothe scheme with multiplicative error correction and approximate approximations to solve the stationary spatial problems. For a given order k, the solutions $u_{t_j}^{(i)}$ of order i, $i = 1, \ldots, k$, are approximated at each discrete time point $t_j \in [0, T]$ by

$$\hat{u}_{t_j}^{(i)}(x) = \mathcal{M}_{h,\mathcal{D}} u_{t_j}^{(i)}(x) = \mathcal{D}^{-d/2} \cdot \sum_{x_n \in \mathcal{G}_h} u_{t_j}^{(i)}(x_n) \cdot \eta\left(\frac{x - x_n}{\sqrt{\mathcal{D}}h}\right)\,, \quad \forall x \in \Omega\,,$$

where $\mathcal{G}_h := \{mh \in \Omega,\ m \in \mathbb{Z}^d\}$ is a finite uniform grid with grid size h on the spatial domain $\Omega \subset \mathbb{R}^d$.

The generating function η is constructed according to (4.21) by the product of a Gaussian and a Laguerre polynomial of order $(M-1)$, which implies that the approximate approximations are of order $2M$. This choice of η allows us to compute the action of the differential operator $\mathcal{A}$ on η analytically for every $x \in \Omega$ and in particular for the grid points $x_n \in \mathcal{G}_h$. As a consequence, the discretized stationary spatial problems can be restated as systems of linear equations. Solution of the latter yields the coefficients of the approximate approximations.

In this chapter we introduce an algorithmic realization of the method. The flowchart in Figure 5.1 illustrates the algorithmic scheme. Roman numbers on the right indicate different stages of the algorithm, which will be discussed in the following sections. These are:

I Semi-discretization in time & solution of the stationary spatial problems, Section 5.1.

II Error estimation & adaptivity, Section 5.2.

III Movement of the boundaries of the discretization domain Ω, Section 5.3.

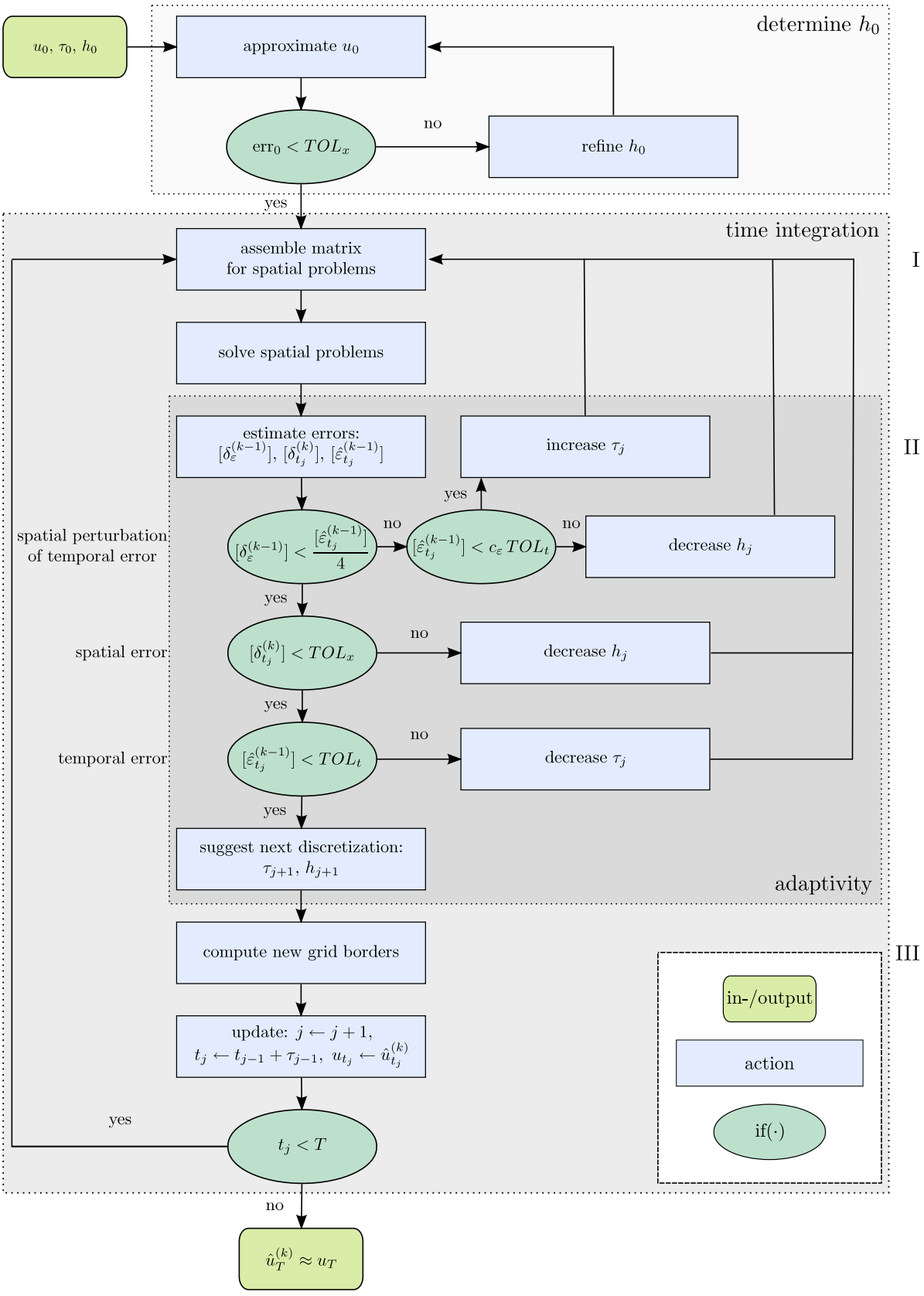

Figure 5.1: Adaptive density propagation: flowchart of the algorithm.

The effects of different parameters of the algorithm on the output and performance are discussed in Section 5.4. A convergence analysis and numerical simulations are postponed to Chapters 6 and 7.

5.1 Semi-discretization in time & solution of the stationary spatial problems

In each integration step $t_{j+1} = t_j + \tau_j$, $j = 0, 1, 2, \ldots$, a time step $\tau_j > 0$ and grid size h_j are given. As previously we denote by $u_t^{(i)}$ the solution of order i, and by $\Delta u_t^{(i)} = u_t^{(i+1)} - u_t^{(i)}$ the difference of two solutions of order $i+1$ and i, where the 0th-order solution is such that $u_{t+\tau}^{(0)} := u_t$ for all $\tau > 0$. Semi-discretization in time is carried out using the second- or third-order scheme with multiplicative error correction introduced in Chapter 3. Depending on the selected discretization method, the stationary spatial problems are given by (3.17) for the L-stable discretization method, i.e.,

$$\begin{aligned}
(\mathrm{Id} - \tau\mathcal{A})\ u_{t_{j+1}}^{(1)} &= u_{t_j} \\
(\mathrm{Id} - \tau\mathcal{A})^2\ \Delta u_{t_{j+1}}^{(1)} &= -\frac{1}{2}(\tau\mathcal{A})\ \Delta u_{t_{j+1}}^{(0)} \\
(\mathrm{Id} - \tau\mathcal{A})\ \Delta u_{t_{j+1}}^{(2)} &= -\frac{4}{3}(\tau\mathcal{A})\ \Delta u_{t_{j+1}}^{(1)}
\end{aligned}$$

or by (3.24) for the A-stable discretization method, i.e.,

$$\begin{aligned}
(\mathrm{Id} - \tau\mathcal{A})\ u_{t_{j+1}}^{(1)} &= u_{t_j} \\
(\mathrm{Id} - \tau\mathcal{A})\ \Delta u_{t_{j+1}}^{(1)} &= -\frac{1}{2}(\tau\mathcal{A})\ \Delta u_{t_{j+1}}^{(0)} \\
(\mathrm{Id} - \tau\mathcal{A})\ \Delta u_{t_{j+1}}^{(2)} &= -\frac{1}{3}(\tau\mathcal{A})\ \Delta u_{t_{j+1}}^{(1)}\,.
\end{aligned}$$

Spatial discretization of the stationary problems will be exemplified with the first-order implicit Euler approximation

$$(\mathrm{Id} - \tau\mathcal{A})\ u_{t_{j+1}}^{(1)} = u_{t_j}\,. \tag{5.2}$$

Given $h = h_j$, the approximate approximant of $u_{t_{j+1}}^{(1)}$ at the new time point t_{j+1} is

$$\mathcal{M}_{h,\mathcal{D}} u_{t_{j+1}}^{(1)}(x) = \mathcal{D}^{-d/2} \cdot \sum_{x_n \in \mathcal{G}_h} u_{t_{j+1}}^{(1)}(x_n) \cdot \eta\left(\frac{x - x_n}{\sqrt{\mathcal{D}}h}\right)\,. \tag{5.3}$$

The coefficients $u_{t_{j+1}}^{(1)}(x_n)$, $x_n \in \mathcal{G}_h$, are the unknowns to be determined. Inserting (5.3) into (5.2) yields

$$(\mathrm{Id} - \tau\mathcal{A})\ \left(\mathcal{D}^{-d/2} \cdot \sum_{x_n \in \mathcal{G}_h} u_{t_{j+1}}^{(1)}(x_n) \cdot \eta\left(\frac{x - x_n}{\sqrt{\mathcal{D}}h}\right)\right) \quad = \quad u_{t_j}(x)\,. \tag{5.4}$$

Since $\mathcal{A}$ is linear and the sum over x_n is finite, (5.4) is equivalent to

$$\mathcal{D}^{-d/2} \cdot \sum_{x_n \in \mathcal{G}_h} \left(u_{t_{j+1}}^{(1)}(x_n) \cdot (\mathrm{Id} - \tau\mathcal{A})\ \eta\left(\frac{x - x_n}{\sqrt{\mathcal{D}}h}\right)\right) = u_{t_j}(x)\,. \tag{5.5}$$

For the Laguerre-Gaussian generating function η defined by (4.21), $\mathcal{A}\eta(x)$ can be computed analytically for every $x \in \mathbb{R}^d$, see Appendix, Part C. Therefore, by evaluating (5.5) at all $N := |\mathcal{G}_h|$ grid points $x_n \in \mathcal{G}_h$, (5.2) can be restated as a system of linear equations

$$\mathbf{A} \cdot \mathbf{u}^{(1)} = \mathbf{u}^{(0)} \tag{5.6}$$

with $\mathbf{u}^{(0)} \in \mathbb{R}^N$, $\mathbf{A} \in \mathbb{R}^{N \times N}$ and $\mathbf{u}^{(1)} \in \mathbb{R}^N$ defined as

$$\mathbf{u}^{(0)} := \left(u_{t_j}(x_n)\right)_{x_n \in \mathcal{G}_h} \tag{5.7}$$

$$\mathbf{A} \quad := \mathcal{D}^{-d/2} \left((\mathrm{Id} - \tau_j \mathcal{A}) \ \eta\left(\frac{x_m - x_n}{\sqrt{\mathcal{D}}h}\right)\right)_{x_m, x_n \in \mathcal{G}_h} \tag{5.8}$$

$$\mathbf{u}^{(1)} := \left(u^{(1)}_{t_{j+1}}(x_n)\right)_{x_n \in \mathcal{G}_h} . \tag{5.9}$$

Solution of (5.6) yields the coefficients $\mathbf{u}^{(1)}_n = u^{(1)}_{t_{j+1}}(x_n)$, $n = 1, \ldots, N$, of the approximate approximant (5.3). Thus, the fully discrete first-order solution $\hat{u}^{(1)}_{t_{j+1}}$ is given by

$$\hat{u}^{(1)}_{t_{j+1}}(x) = \mathcal{M}_{h,\mathcal{D}} u^{(1)}_{t_{j+1}}(x) = \mathcal{D}^{-d/2} \cdot \sum_{x_n \in \mathcal{G}_h} \mathbf{u}^{(1)}_n \cdot \eta\left(\frac{x - x_n}{\sqrt{\mathcal{D}}h}\right) . \tag{5.10}$$

Subsequently, with

$$\Delta u^{(0)}_{t_{j+1}}(x_n) = u^{(1)}_{t_{j+1}}(x_n) - u_{t_j}(x_n), \quad x_n \in \mathcal{G}_h ,$$

the other spatial problems are solved analogously. Their solution yields the values $\Delta u^{(1)}_{t_{j+1}}(x_n)$ and $\Delta u^{(2)}_{t_{j+1}}(x_n)$, $x_n \in \mathcal{G}_h$, which are used to compute the coefficients of $\hat{u}^{(2)}_{t_{j+1}}$ and $\hat{u}^{(3)}_{t_{j+1}}$ by

$$u^{(2)}_{t_{j+1}}(x_n) = u^{(1)}_{t_{j+1}}(x_n) + \Delta u^{(1)}_{t_{j+1}}(x_n), \quad x_n \in \mathcal{G}_h \tag{5.11}$$

$$u^{(3)}_{t_{j+1}}(x_n) = u^{(2)}_{t_{j+1}}(x_n) + \Delta u^{(2)}_{t_{j+1}}(x_n), \quad x_n \in \mathcal{G}_h . \tag{5.12}$$

The fully discrete solutions to (5.1) of order $k = 2, 3$ are then given by

$$\hat{u}^{(2)}_{t_{j+1}}(x) = \mathcal{D}^{-d/2} \cdot \sum_{x_n \in \mathcal{G}_h} u^{(2)}_{t_{j+1}}(x_n) \cdot \eta\left(\frac{x - x_n}{\sqrt{\mathcal{D}}h}\right) \tag{5.13}$$

and

$$\hat{u}^{(3)}_{t_{j+1}}(x) = \mathcal{D}^{-d/2} \cdot \sum_{x_n \in \mathcal{G}_h} u^{(3)}_{t_{j+1}}(x_n) \cdot \eta\left(\frac{x - x_n}{\sqrt{\mathcal{D}}h}\right) . \tag{5.14}$$

Remark 5.1.1 (Structure of the matrix A).

1. *Note that for all but the second stationary problem in the L-stable discretization scheme, the matrix* $\mathbf{A}$ *is defined by* (5.8). *For the second stationary problem of the L-stable scheme, the matrix becomes*

$$\mathbf{A}^{(2)} = \left((\mathrm{Id} - \tau\mathcal{A})^2 \ \eta\left(\frac{x_m - x_n}{\sqrt{\mathcal{D}}h}\right)\right)_{x_m, x_n \in \mathcal{G}_h} .$$

This has two disadvantages: First, the calculation of second-order derivatives is computationally demanding, see Appendix, Part C. Second, if the matrix $\mathbf{A}$ *remains constant for all stationary spatial problems, the linear systems can be solved more efficiently by a previous decomposition of the matrix.*

2. *Further note that due to the results on approximate approximations with finite grids shown in Section 4.3.2, a basis function* $\eta\left(\frac{x-x_n}{\sqrt{\mathcal{D}}h}\right)$ *needs* not *to be evaluated at all grid points* $x_m \in \mathcal{G}_h$. *Since* η *is rapidly decaying, only points* $x_m \in B_\kappa(x_n)$ *in a neighborhood* $B_\kappa(x_n)$ *are considered, where* $\kappa = \nu \cdot h$ *is proportional to* h. *Then, the matrix* $\mathbf{A}$ *is* sparse, *i.e.*

$$\mathbf{A} = \begin{pmatrix} * & \cdots & * & & & \\ \vdots & * & \cdots & * & & \\ * & \vdots & \ddots & & \ddots & \\ & * & & * & \cdots & * \\ & & \ddots & \vdots & \ddots & \vdots \\ & & & * & \cdots & * \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \uparrow \\ \boldsymbol{\nu} \\ \downarrow \end{matrix}$$

$$\underbrace{\qquad\qquad}_{\leftarrow\ \boldsymbol{\nu}\ \rightarrow}$$

(assuming that the grid points are appropriately sorted). The parameter ν *determines the number of non-zero entries in each row and column—independent of* h. *Consequently, the linear systems* (5.6) *can be solved using efficient algorithms for sparse matrices.*

5.2 Error estimation & adaptivity

In the multiplicative error control, an approximation $\hat{u}_{t_{j+1}}^{(k)}$ is accepted, if the conditions

$$[\hat{\varepsilon}_{t_j}^{(k-1)}](\tau_j) + [\delta_{t_{j+1}}^{(k)}](h_j) \leq TOL \tag{5.15}$$

$$[\delta_{\varepsilon}^{(k-1)}] < \frac{[\hat{\varepsilon}_{t_j}^{(k-1)}](\tau_j)}{4} \tag{5.16}$$

are satisfied. The first condition is realized by requiring

$$[\hat{\varepsilon}_{t_j}^{(k-1)}](\tau_j) \leq TOL_t\,, \tag{5.17}$$

$$[\delta_{t_{j+1}}^{(k)}](h_j) \leq TOL_x\,, \tag{5.18}$$

where the temporal tolerance TOL_t and the spatial tolerance TOL_x are defined by

$$TOL_t := \rho \cdot TOL\,, \tag{5.19}$$

$$TOL_x := c_{\tau:x} \cdot \tau \cdot (1-\rho) \cdot TOL\,, \tag{5.20}$$

with constants ρ, $0 < \rho < 1$, and $c_{\tau:x} > 0$. Note that the definition of TOL_x differs by the factor $c_{\tau:x} \cdot \tau$ from the one suggested in [8] (compare Algorithm 3.2.1, where TOL_x is defined independent of τ). The choice of the spatial tolerance will be discussed in more detail in Chapter 6, where the convergence of the method is analyzed.

The grey shaded area in the flowchart in Figure 5.1 illustrates the order of verifying conditions (5.16), (5.17) and (5.18): First, the spatial error estimates $[\delta_{\varepsilon}^{(k-1)}]$ and $[\delta_{t_{j+1}}^{(k)}]$ have to

meet their accuracy conditions and subsequently the temporal error estimate $[\hat{\varepsilon}_{t_j}^{(k-1)}]$. This is realized as follows:

1. Violation of condition (5.16) can have two sources:
   (a) Spatial discretization is essentially accurate enough, but the temporal error estimate $[\hat{\varepsilon}_{t_j}^{(k-1)}]$ is considerably smaller than TOL_t, i.e.,

   $$[\hat{\varepsilon}_{t_j}^{(k-1)}](\tau_j) < c_\varepsilon \cdot TOL_t \quad \text{with} \quad 0 < c_\varepsilon \ll 1\,.$$

   To avoid computationally expensive refinement of the spatial discretization, the time step τ_j is increased in this case, and all steps described previously are repeated with a larger time step $\tau_j = \tau^*$.
   (b) The grid size h_j is too large such that the spatial perturbation $\delta_\varepsilon^{(k-1)}$ of the temporal error $\hat{\varepsilon}_{t_j}^{(k-1)}$ may impair time step selection. In this case, h_j is decreased and all previously described steps are repeated with the reduced $h_j = h^*$.
2. In case the spatial tolerance condition (5.18) does not hold, the grid size h_j is decreased and all previous steps are repeated with the reduced $h_j = h^*$.
3. Violation of the temporal tolerance condition (5.17) requires the decrease of the time step τ_j and subsequent repetition of all previous steps with the reduced $\tau_j = \tau^*$.

In case all accuracy conditions are satisfied, the approximate approximation $\hat{u}_{t_{j+1}}^{(k)}$ is accepted, i.e.,

$$\begin{aligned} j &\leftarrow j+1 \\ t_j &\leftarrow t_{j-1} + \tau_{j-1} \\ u_{t_j} &\leftarrow \hat{u}_{t_j}^{(k)} \end{aligned}$$

and a new grid size h^* as well as time step τ^* is suggested for the next integration step. In the following, we will treat the estimation of errors and the selection of h^* and τ^* in more detail.

5.2.1 Spatial error estimates & grid size selection

We have seen in the previous chapter that approximate approximations provide easily accessible error estimates by the difference of the coefficients to the quasi-interpolating values, i.e., the spatial discretization error in the solution of the first stationary problem can be estimated by

$$[\mathrm{err}^{(1)}] \;:=\; \left(h^d \cdot \sum_{x_n \in \mathcal{G}_h} |u_{t_j}^{(1)}(x_n) - \hat{u}_{t_j}^{(1)}(x_n)|^p \right)^{1/p} \approx \mathrm{err}^{(1)} \tag{5.21}$$

and for the second and third stationary problems by

$$[\mathrm{err}^{(i)}] \;:=\; \left(h^d \cdot \sum_{x_n \in \mathcal{G}_h} |\Delta u_{t_j}^{(i-1)}(x_n) - \Delta \hat{u}_{t_j}^{(i-1)}(x_n)|^p \right)^{1/p} \approx \mathrm{err}^{(i)}\,, \tag{5.22}$$

where $\mathrm{err}^{(i)}$ denotes the true approximation error of the i-th stationary problem. The spatial errors

$$\delta_{t_{j+1}}^{(i)}(h) = \hat{u}_{t_{j+1}}^{(i)} - u_{t_{j+1}}^{(i)}\,, \qquad i = 1, \ldots, k \tag{5.23}$$

$$\delta_{\varepsilon}^{(i)} = \hat{\varepsilon}_{t_j}^{(i)}(\tau_j) - \varepsilon_{t_j}^{(i)}(\tau_j)\,, \quad i = 0, \ldots, k-1 \tag{5.24}$$

as defined in (3.28) and (3.31) comprise of the approximation errors $\mathrm{err}^{(i)}$ as well as their propagation through the recursion (3.17) or (3.24). They can be estimated recursively from the $\mathrm{err}^{(i)}$ using relation (B.17) or (B.16), as shown in Part B of the Appendix. Let $[\delta_{t_{j+1}}^{(i)}]$ and $[\delta_{\varepsilon}^{(i)}]$ denote the estimates of the norms of (5.23) and (5.24). A new grid size h^* is then suggested by

$$h^* = \min \left\{ \sigma \cdot \underbrace{\sqrt[2M]{\frac{1}{4}\frac{[\varepsilon_{t_j}^{(k-1)}](\tau_j)}{[\delta_{\varepsilon}^{(k-1)}]}} \cdot h}_{(a)}\,,\; \sigma \cdot \underbrace{\sqrt[2M]{\frac{TOL_x}{[\delta_{t_{j+1}}^{(k)}](h)}} \cdot h}_{(b)}\,,\; c_h \cdot h\,,\; h_{\max} \right\}. \tag{5.25}$$

Here, the first term (a) ensures that h^* is decreased if condition (5.16) is not satisfied, and that h^* is adapted to the spatial perturbation $[\delta_{\varepsilon}^{(k-1)}]$ for the subsequent integration step, in case (5.16) is satisfied. The same holds for the second term (b) with condition (5.18) on the spatial error estimate $[\delta_{t_{j+1}}^{(k)}]$. The constant σ with $0 < \sigma < 1$, is a safety factor. The constant $c_h > 1$ in the third term prevents h^* from growing too quickly. Finally, $h_{\max} < 1/\sqrt{\mathcal{D}}$ ensures that h^* remains within a convergence range (see Remark 4.3.7).

Remark 5.2.1 (**Suggesting an initial grid size**). *The process of finding an initial grid size h_0 that satisfies the spatial tolerance condition* (5.18) *can be speeded up. Instead of solving the linear systems with the user-specified grid size h_0 and possibly repeating those steps until h_0 satisfies the conditions, the initial grid size can be refined previously such that the spatial tolerance condition is satisfied for u_0, i.e.*

$$||u_0 - \mathcal{M}_{h_0,\mathcal{D}} u_0|| \leq TOL_x(\tau_0)\,.$$

This procedure relies on the assumption that $u_0 \approx u_{\tau_0}$ for τ_0 sufficiently small.

5.2.2 Temporal error estimates & time step selection

As shown in Section 3.2 and Part A of the Appendix, the exact temporal error is estimated by the comparison of $\hat{u}_{t_{j+1}}^{(k)}$ with $\hat{u}_{t_{j+1}}^{(k-1)}$,

$$\hat{\varepsilon}_{t_j}^{(k-1)}(\tau_j) := \hat{u}_{t_j+\tau_j}^{(k)} - \hat{u}_{t_j+\tau_j}^{(k-1)} = \Delta \hat{u}_{t_j+\tau_j}^{(k-1)}\,. \tag{5.26}$$

The integral is approximated on the grid points $x_n \in \mathcal{G}_h$ and thus $\hat{\varepsilon}_{t_j}^{(k-1)}$ is estimated by

$$[\hat{\varepsilon}_{t_j}^{(k-1)}] \;:=\; \left(h^d \cdot \sum_{x_n \in \mathcal{G}_h} |\Delta \hat{u}_{t_j+\tau_j}^{(k-1)}(x_n)|^p \right)^{1/p} \approx \left|\left| \hat{\varepsilon}_{t_j}^{(k-1)}(\tau_j) \right|\right| . \tag{5.27}$$

As derived in Section 3.1 and analogously to the grid size selection (5.25), a new time step τ^* is then suggested as

$$\tau^* = \min \left\{ \sigma \cdot \underbrace{\sqrt[k]{\frac{TOL_t}{[\hat{\varepsilon}_{t_j}^{(k-1)}]}} \cdot \tau}_{(c)} \;,\; c_\tau \cdot \tau \;,\; \tau_{\max} \right\} . \tag{5.28}$$

The first term (c) ensures that τ^* is decreased if the temporal accuracy condition (5.17) is violated and that τ^* is adapted to $[\hat{\varepsilon}_{k-1}]$ if (5.17) is satisfied. The constant σ, $0 < \sigma < 1$ is a safety factor, and $c_\tau > 1$ in the second term prevents τ^* from growing too quickly. Finally, the constant $\tau_{\max}$ prevents τ^* from becoming too large. In contrast to $h_{\max}$, $\tau_{\max}$ needs to be specified in advance.

5.3 Moving the spatial domain

Moving the discretization region $\Omega = [x_{\min}, x_{\max}] \subset \mathbb{R}^d$, $x_{\min}, x_{\max} \in \mathbb{R}^d$, with the solution is a simple way to keep the number of basis functions small and hence reduce computational costs. In each integration step $t_j \in [0, T]$, new margins $x_{\min}$ and $x_{\max}$ are specified such that Ω spans important regions of u. A region is considered important, if the approximation $\hat{u}_{t_j}^{(k)}$ exceeds a certain threshold. We select this threshold proportional to the spatial tolerance TOL_x, i.e.

$$|\hat{u}_{t_j}^{(k)}| \geq \omega \cdot TOL_x \,, \quad \forall x \in \Omega \,, \tag{5.29}$$

where the constant ω, $0 < \omega < 1$, is specified in advance. Note that, although u is a probability density function and thus $u \geq 0$ for all $x \in \Omega$, $t > 0$, the numerical solution $\hat{u}_{t_j}^{(k)}$ may be negative at certain points $x \in \Omega$. Therefore, we require the *absolute value* to exceed $\omega \cdot TOL_x$ in (5.29). For the same reason, an additional safety constant $\zeta > 0$ is added/subtracted, such that for $d = 1$, new margins are chosen by

$$\begin{aligned} x_{\min} &:= \min \left\{ x_m \in \mathcal{G}_h \,,\; |\hat{u}_{t_j}^{(k)}(x_n)| \geq \omega \cdot TOL_x \right\} - \zeta \,, \\ x_{\max} &:= \max \left\{ x_m \in \mathcal{G}_h \,,\; |\hat{u}_{t_j}^{(k)}(x_n)| \geq \omega \cdot TOL_x \right\} + \zeta \,. \end{aligned} \tag{5.30}$$

For $d > 1$, $x_{\min}$ and $x_{\max}$ are selected such that (5.29) holds in each dimension.

5.4 Parameters & numerical aspects

In this section we present the relevant parameters of the algorithm and discuss their effect on the performance of the method. Default values are also given.

The generating function η: With the Laguerre-Gaussian generating function η defined in (4.21), the choice of η refers to the choice of M, where $2M$ is the approximation order of the approximate approximation $\mathcal{M}_{h,\mathcal{D}}$. By increasing M, the computational costs grow only due to the evaluation of the generalized Laguerre polynomial $\mathcal{L}_{M-1}$ of order $M-1$. In contrast to other conventional spatial discretization methods such as finite element or finite volume methods, an increase of the approximation order of approximate approximations does not require any additional grid points. Therefore, the growth of computational costs is negligible. Moreover, the linear systems are more likely to become ill-conditioned for low M, see Figure 5.2. We therefore recommend to use generating functions with high approximation order. *Default:* $M = 3$, which means $\eta = \eta^{(6)}$.

The parameter $\mathcal{D}$ of the approximate approximation: The constant $\mathcal{D} > 0$ scales the basis functions

$$\eta\left(\frac{x - x_m}{\sqrt{\mathcal{D}}h}\right), \quad x_m \in \mathcal{G}_h\,, \tag{5.31}$$

and hence determines their width. Increasing $\mathcal{D}$ results in a decrease of the saturation error ε_{sat}, see Theorems 4.3.6 and 4.3.9. As a consequence, more basis functions need to be considered for the evaluation of $\mathcal{M}_{h,\mathcal{D}}$ as shown in Section 4.3.2. This effects the sparseness of the matrix $\mathbf{A}$ in (5.6) and thus causes higher computational costs for the solution of the linear systems. Moreover, a large $\mathcal{D}$ denotes a large overlap of the basis functions, which can cause the linear systems to become ill-conditioned for decreasing grid sizes, see Figure 5.2. We therefore recommend to choose $\mathcal{D}$ as small as possible such that (for sufficiently smooth solutions u) the saturation error ε_{sat} is in the range of machine precision. *Default:* $\mathcal{D} = 3$.

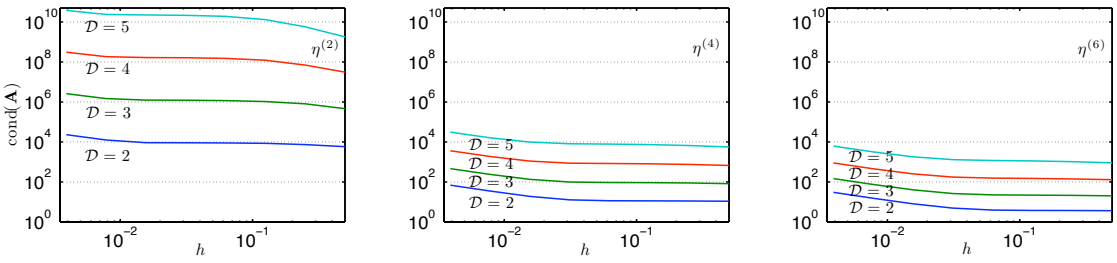

Figure 5.2: Condition of the matrix $\mathbf{A}$ depending on h and $\mathcal{D}$ for approximate approximations of order 2, 4, 6 (left to right).

***A*- or *L*-stability:** In Section 3.1 we introduced two semi-discretization methods, an L-stable (3.17) and an A-stable method (3.24). L-stability allows us to select larger time steps τ, which makes it favorable in terms of reducing the number of time steps. However, in view of Remark 5.1.1, the A-stable method allows for a more efficient computation. *Default:* A-stable.

The order k of semi-discretization in time: Due to a higher convergence rate of the temporal error, larger time steps can be chosen with a high discretization order k. For problems, where time steps become very small, the third-order scheme is recommendable. This is in particular true, if time steps largely vary along integration. For slow temporal dynamics with few variations, a lower-order scheme is favorable, because it requires the solution of less stationary spatial problems. *Default:* $k = 2, 3$.

Local tolerance TOL: The local tolerance determines the local accuracy of temporal and spatial discretization. Decreasing TOL results in smaller time steps τ and grid sizes h, thus in a higher number N of grid points $x_m \in \mathcal{G}_h$. The choice of TOL also depends on the accuracy required to study different problems. Computations shown in this thesis used values of TOL ranging from 10^{-7} to 10^{-1}.

Tolerance factor ρ: The factor ρ, $0 < \rho < 1$, is used to split the local tolerance into temporal TOL_t and spatial tolerance TOL_x, see (5.19) and (5.20). A small ρ accounts for high temporal accuracy and less spatial accuracy. We recommend to choose ρ close to one. *Default:* $\rho = 0.9$.

The spatial tolerance factor $c_{\tau:x}$: The role of the constant $c_{\tau:x}$ will be discussed in more detail in Chapter 6. It should be larger than the average time step. To realize this, we recommend to choose the reciprocal of the specified maximal time step. *Default:* $c_{\tau:x} = 1/\tau_{\max}$.

Safety factor σ: In the adaptive selection of grid sizes h^* (5.25) and time steps τ^* (5.28), σ serves as a safety factor. A small choice of σ accounts for a cautious choice of τ^* and h^* with higher computational costs as a consequence. *Default:* $\sigma = 0.8$ to 0.9.

Maximal step size $\tau_{\max}$: In general it is recommended to choose a large maximal time step and leave the choice of time steps to the adaptive scheme. However, due to $c_{\tau:x} = 1/\tau_{\max}$, a smaller value denotes less spatial discretization costs. We suggest to choose $\tau_{\max}$ large as long as few knowledge about the dynamics is available. *Default:* $\tau_{\max} = 0.5$.

Maximal grid size $h_{\max}$: As shown in Section 4.3, for fixed $\mathcal{D}$ and decreasing h, the approximation error of $\mathcal{M}_{h,\mathcal{D}}$ decays with $\mathcal{O}\left((\sqrt{\mathcal{D}}h)^{2M}\right)$. Thus, the maximal step size is determined by the choice of $\mathcal{D}$ such that $\sqrt{\mathcal{D}}h < 1$. *Default:* $h_{\max} < 1/\sqrt{\mathcal{D}}$.

Parameters ω and ζ for the grid movement: The factor ω, $0 < \omega < 1$, determines the cut-off value, at which the support of u is considered insignificant. Since the grid points at the boundary of the discretization region have less neighboring grid points, the boundary region yields a lower approximation quality. Large values of u_t close to the boundary can therefore impair the spatial accuracy. To avoid this, ω should be chosen small. In addition, the static value $\zeta > 0$ provides an extra safety margin. A large value ζ accounts for a cautious grid movement. However, when the grid size becomes small, the grid may then become computationally expensive, especially for $d > 1$. The choice of ζ depends on the problem under study. *Default:* $\omega = 0.1$, $\zeta = 0.05$ to 0.5.

Chapter 6

Convergence analysis

In this chapter we investigate the approximation error of the adaptive density propagation scheme proposed in Chapter 5. We are interested in a bound for the global approximation error

$$\left|\left| u_t - \hat{u}_t^{(k)} \right|\right| \tag{6.1}$$

in a compact interval $t \in [0, T]$. In each integration step $t_j \in [0, T]$, the numerical solution $\hat{u}_{t_j}^{(k)}$ is obtained by semi-discretization in time and subsequent spatial discretization by approximate approximation, i.e.

$$\hat{u}_{t_1}^{(k)} = \mathcal{M}_{h_0,\mathcal{D}} \left(R_{\tau_0} u_0 \right) , \quad \text{and} \quad \hat{u}_{t_{j+1}}^{(k)} = \mathcal{M}_{h_j,\mathcal{D}} \left(R_{\tau_j} \hat{u}_{t_j}^{(k)} \right) , \tag{6.2}$$

where $R_\tau := r(\tau \mathcal{A})$ denotes an A-stable rational approximation of order k to the strongly continuous semigroup describing the solution u_t, and $\mathcal{M}_{h,\mathcal{D}}$ is the approximate approximant defined in (4.13) with approximation order $2M$.

We restrict the analysis of the global error to sufficiently smooth functions $u \in \mathcal{U}$ with

$$\mathcal{U} := \left\{ u_t \in \mathcal{L}_1 \cap C^\infty, \ \forall t \in [0, T] : \quad \partial^i u_t \in \mathcal{L}_1, \, \forall i = 1, \ldots, 2M \right\} , \tag{6.3}$$

i.e., functions u that for each $t \geq 0$ are in $\mathcal{L}_1$, infinitely differentiable, and with spatial derivatives up to order $2M$ also in $\mathcal{L}_1$. Errors are considered in the $\mathcal{L}_1$-norm, i.e. $||\cdot|| := ||\cdot||_{\mathcal{L}_1}$ throughout the chapter.

First, we fix the time step τ and grid size h to analyze the properties of the global error. It is shown that the global error converges, if $\mathcal{D} = \mathcal{D}(\tau)$ and $h = h(\tau, \mathcal{D}(\tau))$ are chosen appropriately. This result allows us then to derive implications for the adaptive method, where τ and h are adjusted in each integration step such that local errors remain below a predefined tolerance TOL. To control both temporal and spatial local errors, the tolerance is split into a temporal and spatial tolerance, where

$$TOL_t + TOL_x \leq TOL \, .$$

We show that a coupling between the spatial tolerance TOL_x and τ is necessary to guarantee convergence of the adaptive method up to an error that is caused by the saturation error of the approximate approximations. Last we discuss the advantages of approximate approximations in comparison to classical discretization techniques.

6.1 Global approximation error with fixed discretization

Let $\tau > 0$ and $h > 0$ be fixed. Then by (6.2), for any discrete time point $t_j = j \cdot \tau$, $j = 1, \ldots, n$, the numerical solution $\hat{u}_{t_j}^{(k)}$ is given by

$$\hat{u}_{t_j}^{(k)} = (\mathcal{M}_{h,\mathcal{D}} R_\tau)^j \, u_0 = \mathcal{M}_{h,\mathcal{D}} \left(R_\tau \hat{u}_{t_{j-1}}^{(k)} \right) . \tag{6.4}$$

The global error at $t = t_n$ can be bounded by

$$\begin{aligned} \left|\left| u_{t_n} - \hat{u}_{t_n}^{(k)} \right|\right| &\leq \left|\left| \mathcal{P}_\tau u_{t_{n-1}} - \mathcal{P}_\tau \hat{u}_{t_{n-1}}^{(k)} \right|\right| + \left|\left| \mathcal{P}_\tau \hat{u}_{t_{n-1}}^{(k)} - \hat{u}_{t_n}^{(k)} \right|\right| \\ &= \left|\left| \mathcal{P}_\tau \left(u_{t_{n-1}} - \hat{u}_{t_{n-1}}^{(k)} \right) \right|\right| + \left|\left| \mathcal{P}_\tau \hat{u}_{t_{n-1}}^{(k)} - \hat{u}_{t_n}^{(k)} \right|\right| , \end{aligned} \tag{6.5}$$

where $\mathcal{P}_\tau$ denotes the Frobenius-Perron operator describing the analytical solution. Since $\mathcal{P}_\tau$ is a Markov operator, i.e. $||\mathcal{P}_\tau u|| = ||u||$ for all $u \in \mathcal{L}_1$, this becomes

$$\left|\left| u_{t_n} - \hat{u}_{t_n}^{(k)} \right|\right| \leq \left|\left| u_{t_{n-1}} - \hat{u}_{t_{n-1}}^{(k)} \right|\right| + \left|\left| \mathcal{P}_\tau \hat{u}_{t_{n-1}}^{(k)} - \hat{u}_{t_n}^{(k)} \right|\right| . \tag{6.6}$$

Repeating the above steps for t_j, $j = n-1, n-2, \ldots, 1$, yields the estimate

$$\left|\left| u_{t_n} - \hat{u}_{t_n}^{(k)} \right|\right| \leq \sum_{j=0}^{n-1} \left|\left| \mathcal{P}_\tau \hat{u}_{t_j}^{(k)} - \hat{u}_{t_{j+1}}^{(k)} \right|\right| . \tag{6.7}$$

Thus, the global error is bounded by the sum of the local errors of each integration step. To obtain an explicit error bound we closer investigate the local error. Using definition (6.4), the error can be bounded by

$$\begin{aligned} \left|\left| \mathcal{P}_\tau \hat{u}_{t_j}^{(k)} - \hat{u}_{t_{j+1}}^{(k)} \right|\right| &= \left|\left| \mathcal{P}_\tau \hat{u}_{t_j}^{(k)} - \mathcal{M}_{h,\mathcal{D}} \left(R_\tau \hat{u}_{t_j}^{(k)} \right) \right|\right| \\ &\leq \left|\left| \mathcal{P}_\tau \hat{u}_{t_j}^{(k)} - R_\tau \hat{u}_{t_j}^{(k)} \right|\right| + \left|\left| R_\tau \hat{u}_{t_j}^{(k)} - \mathcal{M}_{h,\mathcal{D}} \left(R_\tau \hat{u}_{t_j}^{(k)} \right) \right|\right| \\ &\leq || \underbrace{\left(\mathcal{P}_\tau - R_\tau \right) \hat{u}_{t_j}^{(k)}}_{\varepsilon_{t_j}(\tau)} || + || \underbrace{\left(\mathrm{Id} - \mathcal{M}_{h,\mathcal{D}} \right) \left(R_\tau \hat{u}_{t_j}^{(k)} \right)}_{\delta_{t_{j+1}}^{(k)}} || , \end{aligned} \tag{6.8}$$

where $\varepsilon_{t_j}(\tau)$ denotes the temporal error as defined in (A.4), and $\delta_{t_{j+1}}^{(k)}$ the spatial error defined in (3.28). Hence, the local error is bounded by the sum of the temporal and the spatial error.

By Theorem A.1.6, R_τ is of order k, and the temporal error in (6.8) is bounded by

$$\left|\left| \varepsilon_{t_j}(\tau) \right|\right| = \left|\left| \left(\mathcal{P}_\tau - R_\tau \right) \hat{u}_{t_j}^{(k)} \right|\right| \leq c_r \cdot \tau^{k+1} \cdot \left|\left| \mathcal{A}^{k+1} \hat{u}_{t_j}^{(k)} \right|\right| , \quad 0 < c_r < \infty. \tag{6.9}$$

Since $\hat{u}_t^{(k)}$ is a finite sum of Gaussians multiplied with a polynomial, $||\mathcal{A}^{k+1} \hat{u}_t^{(k)}|| < C < \infty$ for all $t \in [0\,,\, T]$.

By Theorem 4.3.9, the spatial error can be estimated by

$$\left|\left| \delta_{t_{j+1}}^{(k)} \right|\right| \leq c_\eta \cdot \left(\sqrt{\mathcal{D}} h \right)^{2M} \cdot \left|\left| \nabla_{2M} \left(R_\tau \hat{u}_{t_j}^{(k)} \right) \right|\right| + \varepsilon_{\mathrm{sat}}(t_{j+1}) . \tag{6.10}$$

Here, $\varepsilon_{\text{sat}}(t_{j+1})$ denotes the saturation error at t_{j+1}, which depends on derivatives of $R_\tau \hat{u}_{t_j}^{(k)}$. Note that by Theorem 4.3.9, ε_{sat} becomes arbitrarily small for well-behaved functions $u_t \in W_1^{2M} \subset \mathcal{U}$, $t \in [0,T]$, and sufficiently large $\mathcal{D} > 0$.

Combining (6.8), (6.9) and (6.10) yields the local error bound

$$\left|\left|\mathcal{P}_\tau \hat{u}_{t_j}^{(k)} - \hat{u}_{t_{j+1}}^{(k)}\right|\right| \leq c_r \cdot \tau^{k+1} \cdot \left|\left|\mathcal{A}^{k+1} \hat{u}_{t_j}^{(k)}\right|\right| + c_\eta \cdot \left(\sqrt{\mathcal{D}} h\right)^{2M} \cdot \left|\left|\nabla_{2M}\left(R_\tau \hat{u}_{t_j}^{(k)}\right)\right|\right| + \varepsilon_{\text{sat}}(t_{j+1})\,. \tag{6.11}$$

The above bound implies that, for decreasing τ, the local error decays with order $k+1$ until the spatial error is reached, and for decreasing h with order $2M$ until the temporal plus saturation error is reached. This is illustrated in the following example.

Example 6.1.1 (Decay of the local approximation error). *Consider $d = 1$ and a linear ODE*

$$\dot{x} = F(x) = \alpha \cdot x\,, \quad \alpha \in \mathbb{R}\,.$$

With an initial Gaussian distribution u_0, the analytical solution is computable and given by (1.14). For given $\mathcal{D}$, τ and h, we compute one time step

$$\hat{u}_\tau^{(k)} = \mathcal{M}_{h,\mathcal{D}}\left(R_\tau u_0\right)\,,$$

where $\mathcal{M}_{h,\mathcal{D}}$ is of approximation order $2M = 6$. The numerical solution $\hat{u}_\tau^{(k)}$ is then compared to the analytical solution u_τ. In Figure 6.1 and 6.2 local errors are shown for two different scenarios:

1. *Figure 6.1: $\mathcal{D} = 3$ and $h = 0.002$ constant, but such that the contribution of the spatial errors is expected to be small. The local error is shown for orders $k = 1, 2, 3$ (left to right), and plotted against τ. As predicted, errors (blue solid line) and their estimates (black dashed lines) decay with $\mathcal{O}(\tau^{k+1})$, indicated by red dotted lines, until the spatial error is reached, which for this choice of $\mathcal{D}$ and h is close to machine precision.*

   The second- and third-order solutions (middle and right panel) exhibit an unexpected loss of convergence order for time steps smaller than $\approx 10^{-4}$ (k=2) and $\approx 10^{-3}$ (k=3). So far we cannot explain this behavior, although we suspect that it is due to error propagations through the recursion.

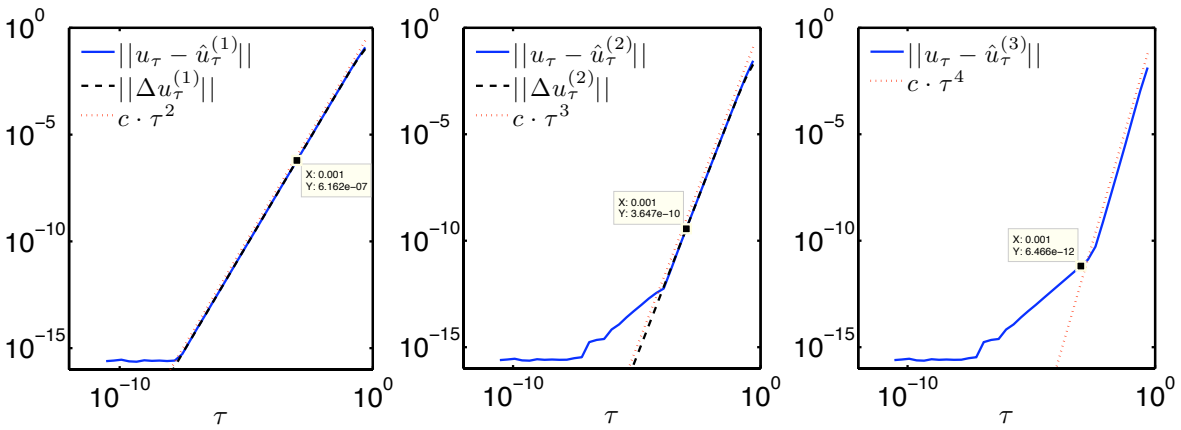

Figure 6.1: Errors of the solutions $\hat{u}_\tau^{(i)}$ with $i = 1$ (left), $i = 2$ (middle), and $i = 3$ (right) with $\mathcal{D} = 3$ for $h = 0.002$ for $\tau \to 0$.

2. *Figure 6.2:* $\mathcal{D} = 3$, $\tau = 0.01$ *(upper panel) and* $\tau = 0.001$ *(lower panel) constant. The local error is shown for orders* $k = 1, 2, 3$ *(left to right), and plotted against* h*. The error decays until the temporal error is reached indicated by highlighted coordinates, which can be identified with the coordinates in Figure 6.1. The error estimates (black dashed lines), defined in* (3.28)*, only estimate the spatial error and therefore, do not stagnate at the temporal error, but continue decaying until the saturation error is reached* $< 10^{-10}$*. The red dotted line shows that the error decays with order* 6*.*

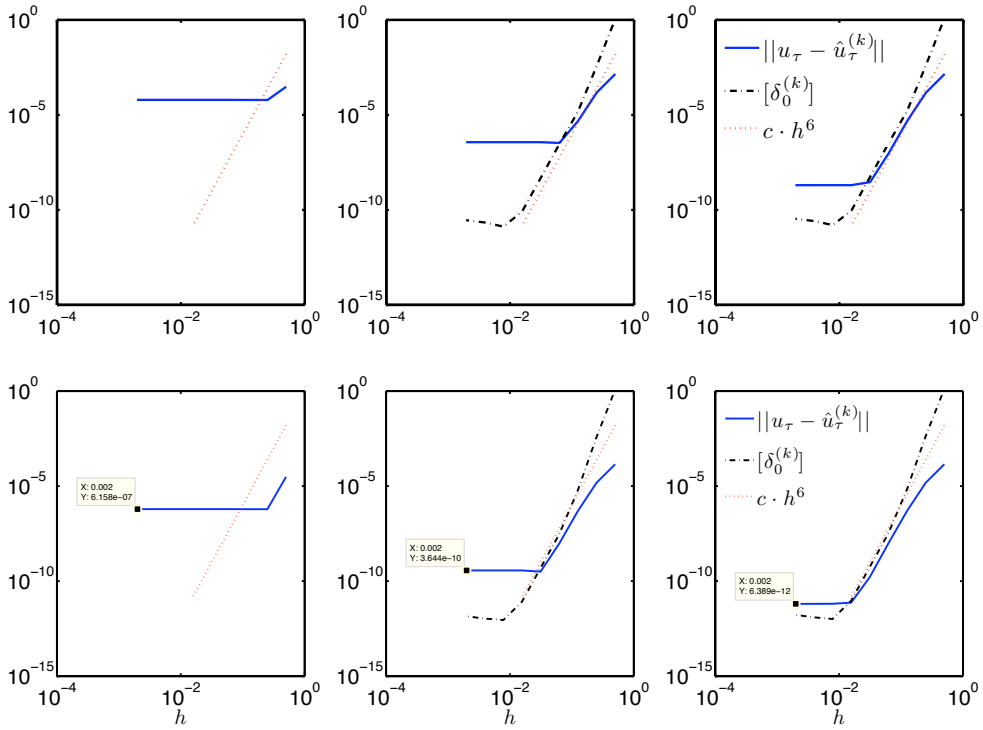

Figure 6.2: Errors of the solutions $\hat{u}_\tau^{(i)}$ of order $i = 1$ (left), $i = 2$ (middle), and $i = 3$ (right) for $\mathcal{D} = 3$ and $\tau = 0.01$ (upper panel), and $\tau = 0.001$ (lower panel) for $h \to 0$.

□

As a consequence of the local error bound (6.11), the global error, which, as in (6.7), is composed of the sum of the local errors, is bounded by

$$\begin{aligned}\left|\left|u_{t_n} - \hat{u}_{t_n}^{(k)}\right|\right| \leq \sum_{j=0}^{n-1} \Big(& c_r \cdot \tau^{k+1} \cdot \left|\left|\mathcal{A}^{k+1} \hat{u}_{t_j}^{(k)}\right|\right| + \\ & c_\eta \cdot \left(\sqrt{\mathcal{D}} h\right)^{2M} \cdot \left|\left|\nabla_{2M}\left(R_\tau \hat{u}_{t_j}^{(k)}\right)\right|\right| + \varepsilon_{\text{sat}}(t_{j+1}) \Big) \\ \leq n \cdot \max_{j=0,\ldots,n-1} \Big(& c_r \cdot \tau^{k+1} \cdot \left|\left|\mathcal{A}^{k+1} \hat{u}^{(k)}(t_j)\right|\right| + \\ & c_\eta \cdot \left(\sqrt{\mathcal{D}} h\right)^{2M} \cdot \left|\left|\nabla_{2M}\left(R_\tau \hat{u}_{t_j}^{(k)}\right)\right|\right| + \\ & \varepsilon_{\text{sat}}(t_{j+1}) \Big) \,. \end{aligned} \tag{6.12}$$

Since the number of integration steps at $t_n = T$ is $n = \frac{T}{\tau}$, this becomes

$$\boxed{\left|\left|u_T - \hat{u}_T^{(k)}\right|\right| \leq T \cdot \left(\hat{c}_r \cdot \tau^k + \frac{\hat{c}_\eta}{\tau} \cdot \left(\sqrt{\mathcal{D}} h\right)^{2M} + \frac{\hat{\varepsilon}_{\text{sat}}}{\tau} \right), \tag{6.13}}$$

where the constants $\hat{c}_\eta$ and $\hat{c}_r$ depend on the norms at the time point $t_{\max}$ maximizing (6.12) and $\hat{\varepsilon}_{\text{sat}} := \varepsilon_{\text{sat}}(t_{\max})$. Note that for a fixed spatial discretization, the spatial errors build up when τ is decreased. Intuitively, this is understandable since then the number of time steps grows, and so does the number of times that the spatial problems have to be solved. The error bound in (6.13) allows us to state the following convergence result.

Theorem 6.1.2 (**Convergence of the global approximation error**). *Let $u \in \mathcal{U}$ and $\hat{u}_T^{(k)} = (\mathcal{M}_{h,\mathcal{D}} R_\tau)^n u_0$, where $R_\tau = r(\tau\mathcal{A})$ denotes an A-stable rational approximation of order k to the strongly continuous semigroup describing the solution u_t, and $\mathcal{M}_{h,\mathcal{D}}$ is the approximate approximant of order $2M$ as defined in (4.13). Then, for any given $C > 0$ there exist $\tau > 0$, $\mathcal{D} > 0$ and $h > 0$, such that the global approximation error at T is bounded by*

$$\left|\left|u_T - \hat{u}_T^{(k)}\right|\right| \leq C \,. \tag{6.14}$$

Hence, the global error converges in τ, $\mathcal{D}$ and h.

Proof: We examine the different parts of the global error bound (6.13) separately. The first term, $\hat{c}_r \tau^k$, refers to the temporal error and vanishes as τ approaches 0. More precisely, for $C_t = \frac{C}{3T}$ there is a $\tau' > 0$ such that for all $\tau < \tau'$

$$\hat{c}_r \cdot \tau^k \leq C_t \,. \tag{6.15}$$

The second and third term are associated with the spatial errors that accumulate during integration. As for the accumulated saturation errors, Theorem 4.3.9 ensures that for any given $\tau > 0$ and $C_{\text{sat}} = \frac{C}{3T}$, there exists a $\mathcal{D} = \mathcal{D}(\tau) > 0$ such that

$$\frac{\hat{\varepsilon}_{\text{sat}}}{\tau} \leq C_{\text{sat}} \,. \tag{6.16}$$

Theorem 4.3.9 further implies that for any given $\tau > 0$, $\mathcal{D} > 0$ and $C_x = \frac{C}{3T}$, there is an $h' = h'(\tau, \mathcal{D}(\tau))$ such that for all $h < h'$ the remaining accumulated spatial errors are bounded by

$$\frac{\hat{c}_\eta}{\tau} \cdot \left(\sqrt{\mathcal{D}} h\right)^{2M} \leq C_x \,. \tag{6.17}$$

Hence, according to inequality (6.13) we have

$$\left|\left| u_T - \hat{u}_T^{(k)} \right|\right| \leq T \cdot (C_t + C_x + C_{\text{sat}}) = C \,.$$

□

Remark 6.1.3. *As a consequence of the accumulated spatial errors, convergence is only guaranteed if $\mathcal{D}$ and h depend on τ. Furthermore, h must satisfy*

$$h < h_{\max} := \frac{1}{\sqrt{\mathcal{D}}} \,. \tag{6.18}$$

In the algorithm in Chapter 5, the parameter $\mathcal{D}$ is constant throughout integration and chosen independently of τ. This implies that the error C_{sat} inevitably grows as τ approaches 0. However, for $\mathcal{D}$ sufficiently large, the saturation error $\hat{\varepsilon}_{\text{sat}}$ can be considered orders of magnitudes smaller than the temporal error and the remaining spatial error. Thus, in practical applications C_{sat} is expected to be considerably smaller than C_t and C_x.

Due to the structure of the error bound (6.13) and the above remark, for $\mathcal{D} > 0$ fixed and τ and $h(\tau)$ decreasing, the global approximation error is expected to decay until, after a possible transition phase, the error saturates, see Figure 6.3. We next closer examine the order of decay *before* the saturation phase.

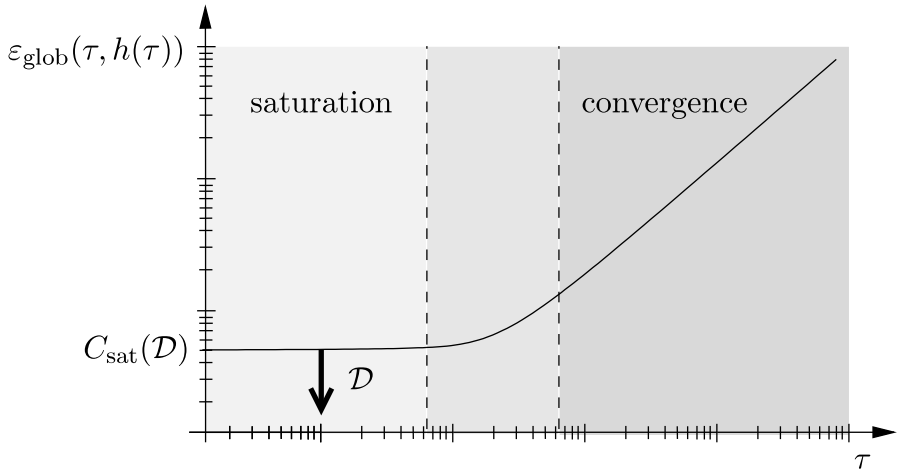

Figure 6.3: Behavior of the global approximation error for decreasing τ and $h(\tau)$.

Theorem 6.1.4 (Decay order of the global approximation error). *Let the same conditions as in Theorem 6.1.2 be satisfied. If $\mathcal{D} > 0$ is fixed and if*

$$h(\tau) = c \cdot \tau^{k+1/2M} \,, \quad \textit{with} \quad c > 0 \,, \tag{6.19}$$

then the global approximation error decays in τ with order k until the saturation error C_{sat} is reached, i.e.

$$\left|\left| u_T - \hat{u}_T^{(k)} \right|\right| = \mathcal{O}(\tau^k) + T \cdot C_{\text{sat}}(\mathcal{D}, \tau) \,, \quad \textit{as} \quad \tau \to 0 \,. \tag{6.20}$$

Proof: Using the global error bound (6.13) with $h = h(\tau)$ according to relation (6.19) yields

$$\left\| u_T - \hat{u}_T^{(k)} \right\| \leq T \cdot \left(\hat{c}_r \cdot \tau^k \; + \; \hat{c}_\eta \cdot \tau^k \cdot \left(c \cdot \sqrt{\mathcal{D}} \right)^{2M} \; + \; \frac{\hat{\varepsilon}_{\text{sat}}}{\tau} \right) .$$

Then with

$$C_{\text{sat}}(\mathcal{D}, \tau) := \frac{\hat{\varepsilon}_{\text{sat}}}{\tau} ,$$

the claim follows.

□

We illustrate this result in the following example, where two scenarios are considered: (1) error growth as τ and $h \to 0$, when τ and h are decreased independently, and (2) decay of order k, when τ and h are decreased according to Theorem 6.1.4.

Example 6.1.5 (Global approximation error using a fixed discretization). *Consider the same scenario as in Example 6.1.1: $d = 1$, F linear and an initial Gaussian distribution.*

We compute the second- and third-order solutions $\hat{u}_T^{(2)}$ and $\hat{u}_T^{(3)}$ with fixed temporal and spatial discretization for $T = 1$, and compare them to the analytical solution. This is repeated for different choices of τ and h. Figures 6.4 and 6.5 show the global approximation error for decreasing time steps and grid sizes plotted against the time steps.

1. *Figure 6.4: time steps and grid sizes are decreased independently; τ is always halved, and h is decreased linearly. As predicted by the global error bound (6.13), errors build up with $\mathcal{O}(\tau^{-1})$, indicated by the red dotted line, although both τ and h are decreased.*

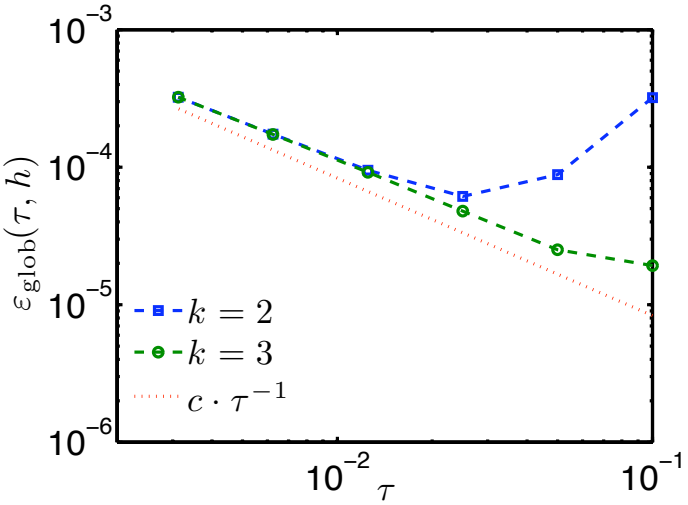

Figure 6.4: Growing approximation error for h and τ decreased independently.

2. *Figure 6.5: τ is decreased and h is determined via relation (6.19). Results show that, in accordance with Theorem 6.1.4, the global approximation error decays with $\mathcal{O}(\tau^k)$, indicated by the red dotted lines.*

□

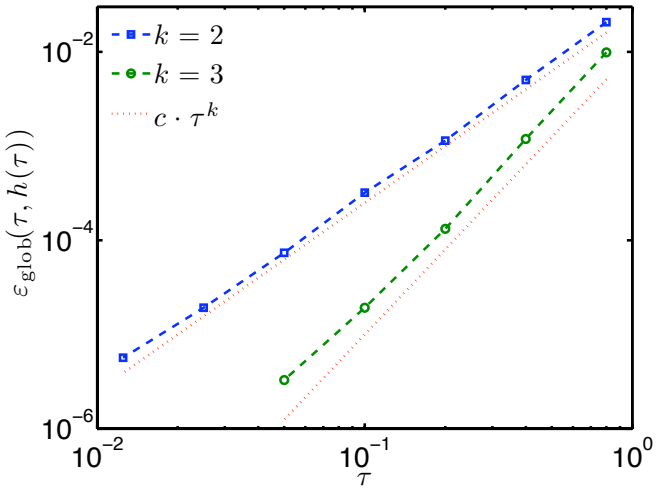

Figure 6.5: The global approximation error decays with order k for h as in (6.19).

6.2 Global approximation error of the adaptive method

From the previous results we inferred how to choose a grid size h such that the convergence order of the temporal discretization scheme is maintained. In the adaptive setting, τ_j and h_j are adjusted in each integration step t_j, $j = 0, \dots, n-1$, such that the local error estimates remain below their specified tolerances TOL_t and TOL_x, respectively. We investigate how the spatial accuracy TOL_x must be chosen to guarantee a global error decay with respect to a specified local tolerance $TOL > 0$. Ideally, the error should decay with the same order as expected in the absence of spatial errors. For the latter case, the following lemma describes the decay of the global approximation error.

Lemma 6.2.1 (**Global approximation error without spatial perturbations**). *Assume $u \in \mathcal{U}$, and let $R_\tau = r(\tau\mathcal{A})$ denote an A-stable approximation of order k to the strongly continuous semigroup describing the solution u_t. Suppose a local tolerance $TOL > 0$ is given. Let the discrete evolution of u at time points $t_{j+1} = t_j + \tau_j$ be defined by*

$$u_{t_1}^{(k)} = R_{\tau_0} u_0 \,, \quad u_{t_{j+1}}^{(k)} = R_{\tau_j} u_{t_j}^{(k)} \,, \quad j = 1, \dots, n-1 \,. \tag{6.21}$$

Assume that the local error estimates adequately approximate the true local errors, i.e.

$$\frac{||\varepsilon_{t_j}(\tau_j)||}{||\varepsilon_{t_j}^{(k-1)}(\tau_j)||} \leq \theta \,, \qquad \theta \leq 1, \ j = 0, \dots, n-1 \,. \tag{6.22}$$

If the time steps τ_j are chosen optimally, i.e.

$$\left|\left|\varepsilon_{t_j}^{(k-1)}(\tau_j)\right|\right| = TOL \,, \qquad j = 0, \dots, n-1 \,, \tag{6.23}$$

then there is a constant $C > 0$ such that the global approximation error at $T = t_n$ is bounded by

$$\left|\left|u_T - u_T^{(k)}\right|\right| \leq C \cdot T \cdot TOL^{\frac{k-1}{k}} \,. \tag{6.24}$$

Proof: As before in the non-adaptive setting, the global approximation error is bounded by the sum of the local errors, i.e.

$$\left|\left|u_T - u_T^{(k)}\right|\right| \;\leq\; \sum_{j=0}^{n-1} \left|\left|\varepsilon_{t_j}(\tau_j)\right|\right| \;\leq\; \theta \cdot \sum_{j=0}^{n-1} \left|\left|\varepsilon_{t_j}^{(k-1)}(\tau_j)\right|\right| \;=\; n \cdot \theta \cdot TOL\,. \tag{6.25}$$

The number of integration steps depends on the accuracy TOL. To replace n, we consider the dependency between the time steps and TOL. Since R_τ is A-stable and of order k, we know that $||\varepsilon_{t_j}(\tau)|| = \mathcal{O}(\tau^{k+1})$. The estimates $\varepsilon_{t_j}^{(k-1)}$, computed by the difference of solutions of order k and $k-1$, then decay with $\mathcal{O}(\tau^k)$. Together with assumption (6.23) this implies

$$\begin{aligned} TOL = \left|\left|\varepsilon_{t_j}^{(k-1)}(\tau_j)\right|\right| &\leq c_j \cdot \tau_j^k\,, \quad c_j > 0\,,\ j = 0, \ldots, n-1\,, \\ \Rightarrow \qquad \tau_j^{-1} &\leq TOL^{-\frac{1}{k}} \cdot c_j^{\frac{1}{k}}\,. \end{aligned}$$

Summation and division by n yields

$$\frac{1}{n} \cdot \sum_{j=0}^{n-1} \tau_j^{-1} \leq TOL^{-\frac{1}{k}} \cdot \left(\frac{1}{n} \cdot \sum_{j=0}^{n-1} c_j^{\frac{1}{k}} \right)\,. \tag{6.26}$$

We define the mean time step $\bar{\tau}$ as

$$\bar{\tau} = \frac{1}{n} \cdot \sum_{j=0}^{n-1} \tau_j \quad \Rightarrow \quad \bar{\tau} = \frac{T}{n}\,.$$

Since the function $\tau \mapsto \tau^{-1}$ is convex, Jensen's inequality can be applied, i.e.

$$\bar{\tau}^{-1} \leq \frac{1}{n} \cdot \sum_{j=0} \tau_j^{-1}\,.$$

Combining the above inequality with (6.26), it then follows that

$$\bar{\tau}^{-1} \leq C_1 \cdot TOL^{-\frac{1}{k}}\,, \quad \text{with} \quad C_1 := \left(\frac{1}{n} \cdot \sum_{j=0}^{n-1} c_j^{\frac{1}{k}} \right)\,. \tag{6.27}$$

Finally, replacing $n = {}^{T}/_{\bar{\tau}}$ in (6.25) yields

$$\begin{aligned} \left|\left|u_T - u_T^{(k)}\right|\right| &\leq \theta \cdot T \cdot \frac{TOL}{\bar{\tau}} \\ &\leq C \cdot T \cdot TOL^{\frac{k-1}{k}}\,, \quad \text{with} \quad C := \theta \cdot C_1\,. \end{aligned}$$

□

Remark 6.2.2. *The different convergence orders of $\varepsilon_t(\tau)$ and $\varepsilon_t^{(k-1)}(\tau)$ justify the assumption that $\theta \leq 1$. If $1 < \theta < \infty$, the claim still holds.*

Therefore, in the best case, without spatial errors, the global approximation error decays with $\mathcal{O}(TOL^{\frac{k-1}{k}})$ as TOL is decreased. In the presence of spatial errors, the adaptive scheme chooses a grid size h_j in each integration step t_j such that the spatial errors remain below the spatial tolerance, i.e.

$$\left|\left|\delta_{t_j}^{(k)}(h_j)\right|\right| \leq TOL_x\,.$$

In view of the results from the previous section, the spatial accuracy may depend on the current time step, i.e. $TOL_x = TOL_x(\tau)$. Imposing a local tolerance TOL that accounts for temporal and spatial errors, i.e.

$$TOL_t + TOL_x \leq TOL\,,$$

we now want to derive a constraint on TOL_x that guarantees the same decay order of the global error in TOL (before the saturation phase is reached, see Figure 6.3) as previously derived in the absence of spatial errors. We will assume that $\mathcal{D} > 0$ is fixed, but sufficiently large, such that the saturation errors $\varepsilon_{\text{sat}}(t_j)$, which are part of $\delta_{t_j}^{(k)}$, are negligible as compared to the remaining spatial errors. As in previous chapters, let $\delta_\varepsilon^{(k)}(h) := \hat{\varepsilon}_t^{(k)} - \varepsilon_t^{(k)}$ denote the spatial perturbation of the temporal error estimate. The following theorem provides conditions for the error estimates such that the global error decays with the same order as in the absence of spatial errors.

Theorem 6.2.3 (**Global approximation error of the fully adaptive scheme**). *Assume $u \in \mathcal{U}$. Further let $R_\tau = r(\tau\mathcal{A})$ denote an A-stable approximation to the strongly continuous semigroup of order k, and $\mathcal{M}_{h,\mathcal{D}}$ the approximate approximant of order $2M$ as defined in* (4.13). *Suppose a local tolerance $TOL > 0$ is given, then let*

$$TOL_t \;=\; \rho \cdot TOL\,, \quad 0 < \rho < 1\,,$$

and let $TOL_x = TOL_x(\tau) > 0$.

For constant but sufficiently large $\mathcal{D} > 0$, let the discrete evolution of u be defined by

$$\hat{u}_{t_1}^{(k)} = \mathcal{M}_{h_0,\mathcal{D}}\left(R_{\tau_0} u_0\right)\,, \quad \hat{u}_{t_{j+1}}^{(k)} = \mathcal{M}_{h_j,\mathcal{D}}\left(R_{\tau_j}\hat{u}_{t_j}^{(k)}\right) \tag{6.28}$$

at time points $t_0 = 0 < t_1 < \ldots < t_n = T$. Assume the unperturbed temporal error estimates provide adequate approximations to the true errors, i.e.

$$\frac{||\varepsilon_{t_j}(\tau_j)||}{||\varepsilon_{t_j}^{(k-1)}(\tau_j)||} \leq \theta\,, \qquad \theta \leq 1,\ j = 0,\ldots,n-1\,. \tag{6.29}$$

If the time steps $\tau_j = t_{j+1} - t_j$ and the sequence of grid sizes $h_j > 0$ satisfy

$$\begin{aligned}
&(i) && ||\hat{\varepsilon}_{t_j}^{(k-1)}(\tau_j)|| = TOL_t \\
&(ii) && ||\delta_{t_{j+1}}^{(k)}(h_j)|| \leq TOL_x(\tau_j) \\
&(iii) && ||\delta_\varepsilon^{(k-1)}(h_j)|| \leq \frac{||\hat{\varepsilon}_{t_j}^{(k-1)}(\tau_j)||}{4}
\end{aligned}$$

for all $j = 0, 1, \ldots, n-1$, *and if further*

$$(iv) \quad \frac{1}{n} \cdot \sum_{j=0}^{n-1} TOL_x(\tau_j) \leq (1-\rho) \cdot TOL\,,$$

then there is a constant $C > 0$ *such that the global approximation error is bounded by*

$$\left|\left| u_T - \hat{u}_T^{(k)} \right|\right| \leq C \cdot T \cdot TOL^{\frac{k-1}{k}}\,. \tag{6.30}$$

Proof: The global approximation error at $T = t_n$ is bounded by the sum of the local errors, i.e.

$$\begin{aligned}
\left|\left| u_T - \hat{u}_T^{(k)} \right|\right| &\leq \sum_{j=0}^{n-1} \left(\left|\left| \varepsilon_{t_j}(\tau_j) \right|\right| + \left|\left| \delta_{t_{j+1}}^{(k)}(h_j) \right|\right| \right) \\
&\leq \sum_{j=0}^{n-1} \left(\theta \cdot \left|\left| \varepsilon_{t_j}^{(k-1)}(\tau_j) \right|\right| + \left|\left| \delta_{t_{j+1}}^{(k)}(h_j) \right|\right| \right) \\
&\leq \sum_{j=0}^{n-1} \left(\left|\left| \varepsilon_{t_j}^{(k-1)}(\tau_j) \right|\right| + \left|\left| \delta_{t_{j+1}}^{(k)}(h_j) \right|\right| \right)\,.
\end{aligned} \tag{6.31}$$

Condition (iii) is equivalent to

$$\frac{4}{5} \cdot \left|\left| \varepsilon_{t_j}^{(k-1)}(\tau_j) \right|\right| \leq \left|\left| \hat{\varepsilon}_{t_j}^{(k-1)}(\tau_j) \right|\right| \leq \frac{4}{3} \cdot \left|\left| \varepsilon_{t_j}^{(k-1)}(\tau_j) \right|\right|\,. \tag{6.32}$$

Combining (6.31) and (6.32) yields

$$\left|\left| u_T - \hat{u}_T^{(k)} \right|\right| \leq \sum_{j=0}^{n-1} \left(\frac{5}{4} \cdot \left|\left| \hat{\varepsilon}_{t_j}^{(k-1)}(\tau_j) \right|\right| + \left|\left| \delta_{t_{j+1}}^{(k)}(h_j) \right|\right| \right)\,,$$

which together with (i), (ii) and (iv) becomes

$$\begin{aligned}
\left|\left| u_T - \hat{u}_T^{(k)} \right|\right| &\leq \frac{5}{4} \cdot \sum_{j=0}^{n-1} \left(TOL_t + TOL_x(\tau_j) \right) \\
&= \frac{5}{4} \cdot \left(n \cdot TOL_t + \sum_{j=0}^{n-1} TOL_x(\tau_j) \right) \\
&\leq \frac{5}{4} \cdot n \cdot (\rho \cdot TOL + (1-\rho) \cdot TOL) = \frac{5}{4} \cdot n \cdot TOL\,.
\end{aligned} \tag{6.33}$$

As in the proof of Lemma 6.2.1, n can be replaced, since R_τ is A-stable and of order k, which together with condition (i) implies

$$\bar{\tau}^{-1} \leq C_1 \cdot TOL^{-\frac{1}{k}}\,, \quad \text{with} \quad C_1 := \rho^{-\frac{1}{k}} \cdot \left(\frac{1}{n} \cdot \sum_{j=0}^{n-1} c_j^{\frac{1}{k}} \right)\,. \tag{6.34}$$

The constants c_j are determined via

$$\left|\left|\hat{\varepsilon}_{t_j}(\tau_j)\right|\right| \leq c_j \cdot \tau_j \,, \qquad j = 0, \ldots, n-1 \,.$$

Finally, replacing $n = {}^{T}/_{\bar{\tau}}$ in (6.33) and combining it with (6.34) yields

$$\begin{aligned} \left|\left| u_T - \hat{u}_T^{(k)} \right|\right| &\leq \frac{5}{4} \cdot n \cdot TOL = \frac{5}{4} \cdot T \cdot \frac{TOL}{\bar{\tau}} \\ &\leq C \cdot T \cdot TOL^{\frac{k-1}{k}} \,, \qquad\qquad C := \frac{5}{4} \cdot C_1 \,. \end{aligned}$$

□

The simplest way to realize the spatial accuracy constraint is by setting $TOL_x(\tau) \equiv (1-\rho) \cdot TOL$ constant, which implies that condition (iv) is satisfied with equality. However, in Theorem 6.2.3 it is assumed that the time steps are chosen optimally, i.e.

$$\hat{\varepsilon}_{t_j}^{(k-1)} = TOL_t \,, \qquad j = 0, \ldots, n-1 \,.$$

In practice, we only demand that

$$\hat{\varepsilon}_{t_j}^{(k-1)} \leq TOL_t \,, \qquad j = 0, \ldots, n-1,$$

i.e., the actual realization of time steps is generally smaller than the optimal sequence of time steps. A smaller choice of time steps will in this case cause an additional accumulation of spatial errors, as observed in the previous section for a fixed discretization (see Figure 6.4). Accumulation of the spatial errors and the consequent loss of decay order for a constant choice of TOL_x are illustrated in the following example.

Example 6.2.4 (Loss of decay order by neglecting the impact of sub-optimal time steps). *Consider the same scenario as in the previous examples: $d = 1$, F linear and an initial Gaussian distribution.*

We adaptively compute the second- and third-order solutions $\hat{u}_T^{(2)}$ and $\hat{u}_T^{(3)}$, where condition (i) in Theorem 6.2.3 is violated, i.e., integration takes more time steps than necessary. The spatial tolerance is constant throughout integration with $TOL_x \equiv (1-\rho) \cdot TOL$. The numerical solutions are then compared to the analytical solution.

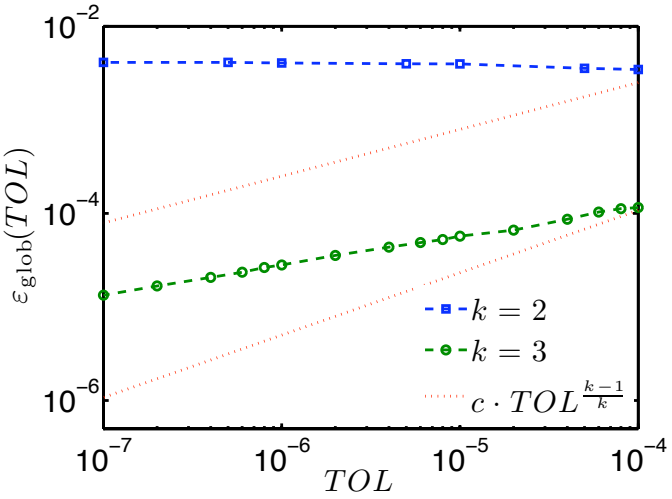

Figure 6.6: Loss of decay order for $TOL_x(\tau) \equiv (1-\rho) \cdot TOL$.

Figure 6.6 shows the global approximation errors of the second- (blue) and third-order solutions (green). Comparison with the red dotted lines indicates a loss of decay order.

□

In order to account for possible additional accumulation of spatial errors, due to sub-optimally chosen time steps τ, we suggest the following choice of TOL_x:

$$TOL_x(\tau) := c_{\tau:x} \cdot \tau \cdot (1-\rho) \cdot TOL\,, \qquad c_{\tau:x} := \frac{1}{\tau_{\max}}\,, \tag{6.35}$$

where $\tau_{\max}$ denotes the prescribed maximal step size. The spatial accuracy constraint (iv) in Theorem 6.2.3 is satisfied, since $\tau_j/\tau_{\max} \leq 1$ for all $j = 0, \ldots, n-1$. Furthermore, this choice of TOL_x accounts for possible accumulations of the spatial errors in the order of $\mathcal{O}(\tau^{-1})$—as observed for the non-adaptive integration scheme in the previous section—by requiring the contributions of the spatial errors to the global error to be equidistributed over the interval $[0, T]$, independent of the actual choice of time steps. An error bound of the global approximation error is then as follows.

Corollary 6.2.5 (**Spatial tolerance depending on τ**). *Assume the same conditions are satisfied as in Theorem 6.2.3, and let TOL_x be defined as in* (6.35)*. Then there are constants $C_t > 0,\ C_x > 0$ such that the global approximation error can be bounded by*

$$\left|\left|u_T - \hat{u}_T^{(k)}\right|\right| \leq T \cdot \left(C_t \cdot TOL^{\frac{k-1}{k}} + C_x \cdot TOL\right)\,. \tag{6.36}$$

Proof: The proof is analogous to the proof of Theorem 6.2.3, only the contributions of the spatial errors change, i.e.

$$\begin{aligned}
\left|\left|u_T - \hat{u}_T^{(k)}\right|\right| &\leq \frac{5}{4} \cdot \sum_{j=0}^{n-1} (TOL_t + TOL_x(\tau_j)) \\
&= \frac{5}{4} \cdot \left(n \cdot \rho \cdot TOL + c_{\tau:x} \cdot (1-\rho) \cdot TOL \cdot \sum_{j=0}^{n-1} \tau_j \right) \\
&= \frac{5}{4} \cdot n \cdot (\rho \cdot TOL + c_{\tau:x} \cdot (1-\rho) \cdot TOL \cdot \bar{\tau})\,.
\end{aligned}$$

Replacing n by $T/\bar{\tau}$ and using the bound (6.34) for $\bar{\tau}^{-1}$ yields

$$\left|\left|u_T - \hat{u}_T^{(k)}\right|\right| \leq \frac{5}{4} \cdot T \cdot C_1 \cdot \left(\rho \cdot TOL^{\frac{k-1}{k}} + c_{\tau:x} \cdot (1-\rho) \cdot TOL\right)$$

The claim follows with $C_t := \frac{5}{4} \cdot C_1 \cdot \rho$ and $C_x := \frac{5}{4} \cdot C_1 \cdot c_{\tau:x} \cdot (1-\rho)$.

□

The above corollary guarantees that, using a more careful choice of TOL_x that depends linearly on the time steps, the decay order of the global approximation error preserves the convergence order of the spatial discretization scheme, even if time steps are chosen sub-optimally. This is illustrated in the following example.

Example 6.2.6 (Decay of the global error in the adaptive scheme). *Consider the same scenario as in the previous examples: $d = 1$, F linear and an initial Gaussian distribution.*

We adaptively compute the second- and third-order solutions $\hat{u}_T^{(2)}$ and $\hat{u}_T^{(3)}$, and compare them to the analytical solution. As previously in Example 6.2.4, time steps are chosen sub-optimally, i.e., condition (i) in Theorem 6.2.3 is violated. The spatial tolerance is given by relation (6.35), where we choose $\tau_{\max} = 0.1$ for the second-order scheme, and $\tau_{\max} = 0.25$ for the third-order scheme, because the latter yields larger time steps.

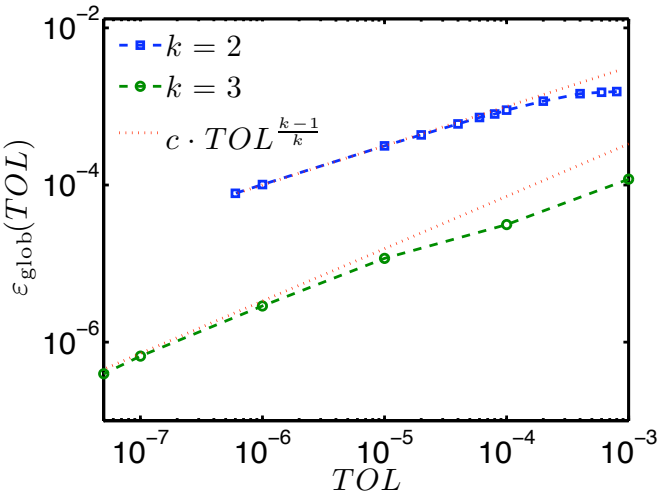

Figure 6.7: Global approximation error of the adaptive scheme.

Figure 6.7 shows the global approximation error at $T = 1$ for different local tolerances TOL.

- *Comparison with the red dotted lines indicates that, after an initial phase ($k = 2$: $TOL > 10^{-4}$, $k = 3$: $TOL > 10^{-5}$) of slightly different decay, the global approximation error decays with the expected order $\frac{k-1}{k}$.*
- *Figure 6.8 depicts the "mean" discretization chosen by the adaptive schemes: the mean time step $\bar{\tau}$ (grey), mean grid size $\bar{h}$ (black) and the mean number of grid points $\bar{N}$ (green). We observe that the grid sizes in the second-order scheme (left) are generally larger than those of the third-order scheme (right). This can be explained by the larger choice of $\tau_{\max}$, since $c_{\tau:x} = \tau_{\max}^{-1}$ determines TOL_x. However, since the time steps decrease faster in the second order scheme, the third-order scheme is less restrictive (asymptotically).*
- *Further note that the mean time steps decay with $O(TOL^{\frac{1}{k}})$, indicated by the red dotted lines, which we expect if equality holds in the upper bound (6.34) for $\bar{\tau}^{-1}$.*

□

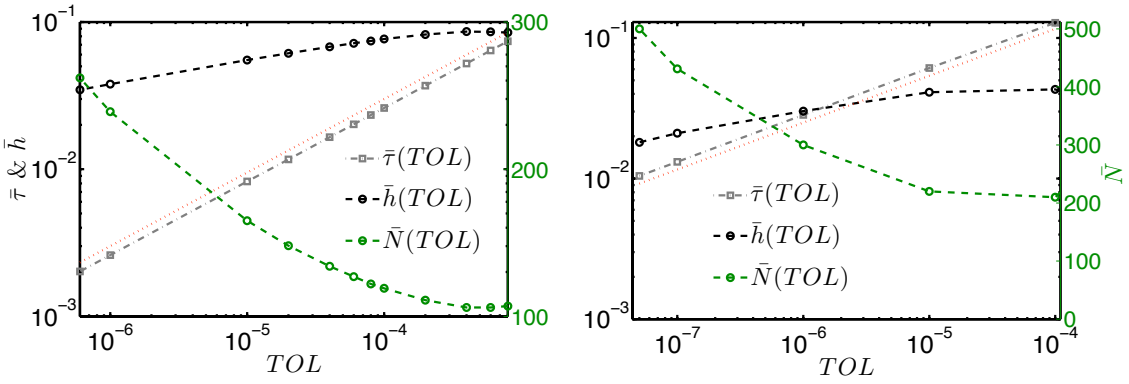

Figure 6.8: Average time steps (grey), grid sizes (black) and number of grid points (green) for $k = 2$ (left) and $k = 3$ (right) as they were chosen by the adaptive scheme.

6.3 Discussion of the results

In this chapter we proved convergence of the general integration scheme. Convergence is guaranteed, provided that the grid size h and the scaling parameter $\mathcal{D}$ of the approximate approximations are decreased dependent on the time steps. In practice, the parameter $\mathcal{D}$ is fixed, which implies that the saturation errors obtained in each integration step inevitably build up as time steps decrease. In the adaptive setting, this means that the local tolerance cannot be chosen arbitrarily small; the spatial tolerance must remain larger than the saturation error. However, for sufficiently large $\mathcal{D}$, the saturation errors are expected to be orders of magnitudes smaller than the remaining errors (compare with Figure 6.1, where the error stagnates at about 10^{-15}). Their contribution to the local and global errors is thus only noticeable for very small choices of τ and h, or TOL, respectively. Furthermore, we think that it is possible to include $\mathcal{D}$ in the spatial adaptivity to guarantee that the saturation errors remain below the spatial tolerance. The grid size h and the scaling parameter $\mathcal{D}$ could then be coupled such that the spatial accuracy conditions are satisfied, constrained to minimizing the condition of the discretized stationary spatial problems (compare with Figure 5.2).

We further investigated the speed of the global error decay before saturation errors become dominant. It was shown how the spatial discretization must be adapted in order to maintain the decay order of the temporal discretization scheme. Since the spatial accuracy constraints impose a dependency on the current time step, they can become prohibitively restrictive for small time steps. For an efficient solution of the spatial problems, satisfying these constraints, a high approximation order of the spatial discretization scheme is required.

The results on the global approximation error of the adaptive scheme required few assumptions about the spatial discretization scheme. The spatial accuracy constraint must be attainable, and the saturation errors must be small as compared to the remaining local errors. Therefore, the results also hold in the case of no saturation errors as, e.g., for classical spatial discretization techniques such as finite element or volume methods. Compared to these methods, approximate approximations possess properties that make them favorable despite the presence of saturation errors:

In the algorithm introduced in Chapter 5 we exploit the fact that the action of the differential operator $\mathcal{A}$ on the approximate approximants can be computed analytically, see also Part C of the Appendix. Furthermore, we showed in Section 4.5, how approximation errors can be estimated by comparing the approximant (a linear combination of basis functions) to its coefficients, i.e., using only information that is readily available. Error estimates for classical discretization methods are typically based on comparing two solutions of different approximation order, analogously to the estimation of temporal errors. With approximate approximations, the computation of additional solutions can be avoided. Moreover, using Theorem 4.4.1 we can construct approximate approximations of high approximation order. In contrast to most classical methods, an increase of the approximation order does not require the inclusion of additional grid points. Additional grid points would severely impair the computational efficiency of the method, because the computation of the coefficients requires the solution of a linear system, the cost of which grows quadratically in the number of grid points. In the case of approximate approximations, computational costs for higher approximation orders increase due to the evaluation of a polynomial of a higher order, cf. (4.21). Compared to the addition of grid points, this can be regarded a minor increase. In summary, approximate approximations allow for an efficient solution of the spatial accuracy constraints.

Most importantly, we prefer approximate approximations, because they provide the scope of extending the adaptive framework to higher-dimensional problems. In high dimensions, the main objective is to reduce the number of grid points. In view of that, a high approximation order is beneficial, since grid sizes can be chosen larger as compared to low approximation orders. However this alone is not sufficient for an efficient solution of the spatial problems in high dimensions. As concluded in Chapter 2, a meshfree setting is most suitable to tackle high-dimensional problems. Approximate approximations based on a meshfree discretization, or scattered grids, are a subject of current research [27, 44, 54]. Also a combination with sparse grids seems feasible and promising. A combination of the adaptive scheme suggested herein with approximate approximations on sparse or scattered grids is promising to efficiently extend the applicability to higher dimensions.

Chapter 7

Numerical examples

In the previous chapter, we illustrated the theoretical results with numerical examples for a one-dimensional linear ODE, where the analytical solution was known. In the following we demonstrate that the proposed method also yields good results for nonlinear ODEs even when the solution gives rise to locally steep gradients or bimodal structures. Since no analytical solution is available for the considered systems, we compare the numerical solutions to solutions obtained via the method of characteristics. The initial spatial discretization for the method of characteristics, see Section 1.2.3, is chosen to be the final grid size of the adaptive solution, and ODEs are solved using the MATLAB solver `ode15s`.

7.1 Michaelis-Menten kinetics (steep gradients close to the boundary)

Michaelis-Menten kinetics are common kinetics used to model a saturable enzymatic degradation of a substance. Consider the substance $\mathbf{X}$ being metabolized by an enzyme $\mathbf{E}$ according to

$$\mathbf{X} + \mathbf{E} \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftarrows}} \mathbf{X}:\mathbf{E} \xrightarrow{k_{\text{cat}}} \mathbf{M} + \mathbf{E}\,,$$

where $\mathbf{X}:\mathbf{E}$ denotes the substance-enzyme complex, $\mathbf{M}$ denotes the metabolite of the enzymatic reaction, and k_{on}, k_{off}, and k_{cat} denote the corresponding rate constants. Let $x \in \mathbb{R}^+$ denote the concentration of the substance $\mathbf{X}$, and $\mathbf{E}_{\text{tot}}$ the total concentration of bound and unbound enzymes. Following the widely used *Michaelis–Menten–approximation*, see, e.g., [18, 49], the temporal evolution of x is described by

$$\dot{x} \;=\; -\frac{V_{\max}}{K_m + x} \cdot x\,, \tag{7.1}$$

where the parameter $V_{\max} = k_{\text{cat}} \cdot \mathbf{E}_{\text{tot}}$ is the chemical flux at saturation, while the Michaelis-Menten constant $K_m = (k_{\text{off}} + k_{\text{cat}})/k_{\text{on}}$ denotes the concentration corresponding to the half-maximal flux $V_{\max}/2$.

We consider an initial Gaussian distribution propagated through the nonlinear dynamics. Figure 7.1 shows the initial distribution with mean $\mu = 2$ and variance $\sigma^2 = 1/4$, as well as the right hand side of (7.1) for $K_m = 1$ and $V_{\max} = 2$. As the concentrations cannot become negative, we expect a skew output distribution.

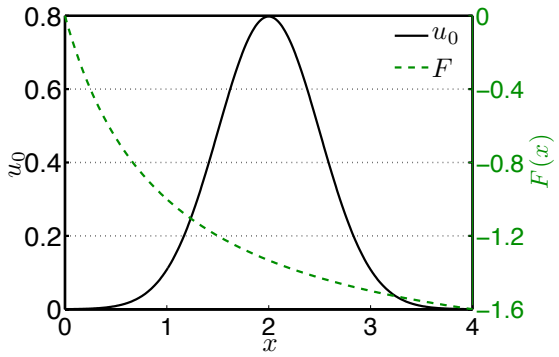

Figure 7.1: Michaelis-Menten model: initial density and right hand side of the ODE.

Figure 7.2 (left) shows the second-order solution at $t = 2$ for a local tolerance of $TOL = 10^{-4}$, a factor $\rho = 0.9$ determining the proportion of temporal and spatial tolerance, and a maximal time step $\tau_{\max} = 0.1$. The solution is in good agreement with the solution obtained via the method of characteristics. Although the border of the discretization domain is close to the steep region of the solution, the numerical solution shows an accurate resolution of the steep front. We further notice that the conservation of probability mass is almost perfect. Figure 7.2 (right) shows the evolution of time steps and grid sizes and the local

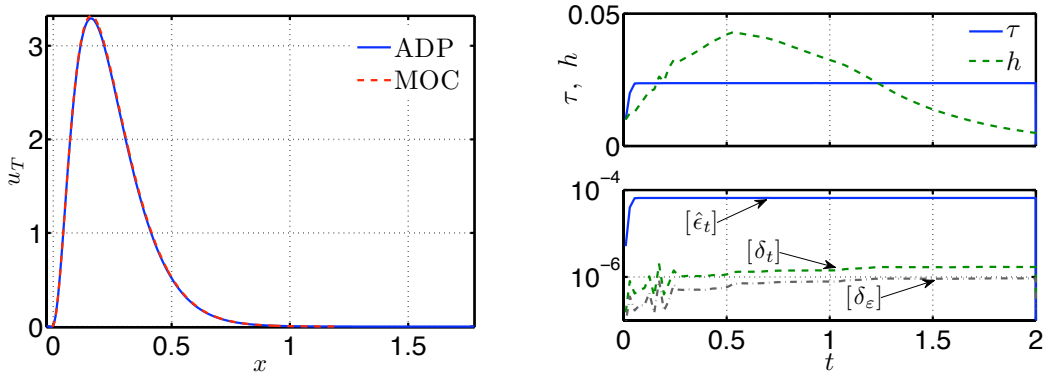

Figure 7.2: Michaelis-Menten model: $T = 2$, $k = 3$, $TOL = 10^{-4}$, $\rho = 0.9$, $\tau_{\max} = 0.1$. Left: final density. Right: Evolution of time steps and grid sizes (upper panel), and of the local error estimates (lower panel).

error estimates in each integration step. While time steps stabilize quickly, the grid size first grows, because of a comparatively simple structure of the solution, and decreases significantly when the solution develops the steep gradient close to the origin. It can be seen that the initially prescribed values for τ and h were smaller than necessary. As soon as the discretization is fully determined by the adaptive scheme (independent of initial choices), the local error estimates stabilize close to the temporal and spatial tolerance.

7.2 Hill kinetics (bimodality)

Hill kinetics are closely related to Michaelis-Menten kinetics. They arise when the enzyme has several binding sites for the substrate, see e.g. [18]. The temporal change of the concentration of the substance is then described by the ODE

$$\dot{x} = -\frac{V_{\max}}{K_h^n + x^n} \cdot x^n \,, \tag{7.2}$$

where $n \in \mathbb{N}$, $V_{\max}$ is the chemical flux at saturation, and K_h the concentration corresponding to $V_{\max}/2$. We want to see how the adaptive density propagation scheme performs when the problem gives rise to a bimodal solution. To do so, we artificially choose $V_{\max} = -0.5$ together with $K_h = 2$ and $n = 10$. Figure 7.3 depicts the right hand side of (7.2) for the chosen parameter values. We consider an initial Gaussian distribution with mean $\mu = 2$ and variance $\sigma^2 = 0.2$, also shown in Figure 7.3.

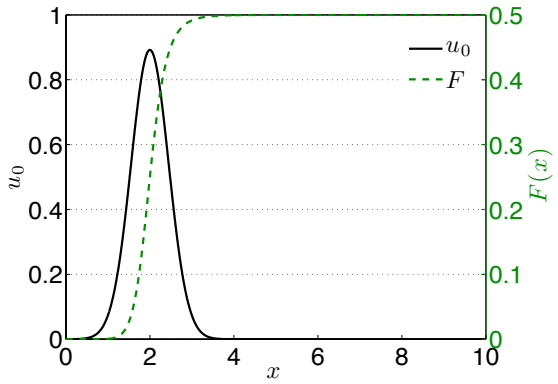

Figure 7.3: Hill model: initial density and right hand side of the ODE.

Figure 7.4 shows the third-order solution at time points $t = 3.3$, 6.6 and 10 computed with $TOL = 10^{-5}$, $\rho = 0.9$ and $\tau_{\max} = 0.1$ in comparison with the solution obtained via the method of characteristics. At all shown time points, both solutions coincide well except for a slight deviation that develops next to the lower mode. This deviation is due to a sparse coverage of grid points for the method of characteristics, which results in an impaired interpolation (since the method of characteristics only yields a pointwise representation of the final density). The steep front as well as the bimodality are accurately captured by the solution obtained by the proposed scheme. The evolution of the time steps and grid sizes selected within the adaptive scheme together with the estimated local errors are shown in Figure 7.5. Both the discretization and the error estimates stabilize quickly and remain roughly constant throughout integration.

7.3 A subcritical model (locally steep gradients)

The third example is a subcritical system that was analyzed in [60]. The system is described by the ODE

$$\dot{x} = x \cdot (\alpha + 2x^2 - x^4) \,, \qquad \alpha \in \mathbb{R} \,. \tag{7.3}$$

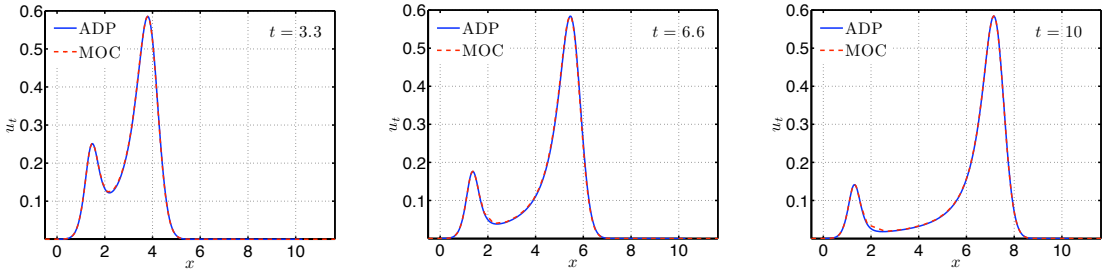

Figure 7.4: Hill model: density at $t = 3.3,\ 6.6,\ 10$ computed with $k = 3$, $TOL = 10^{-5}$, $\rho = 0.9$, $\tau_{\max} = 0.1$.

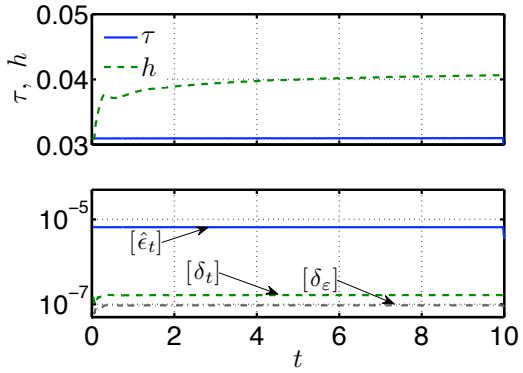

Figure 7.5: Hill model: evolution of time steps and grid sizes (upper panel), and of the local error estimates (lower panel).

For $\alpha = -1/2$, the system has two stable fixed points and three unstable fixed points. These are indicated by dots and circles in Figure 7.6, where the right hand side of (7.3) is shown. We choose a Gaussian initial distribution centered around the unstable fixed point at $x = 0$ and with variance $\sigma^2 = 1/5$, wide enough to cover all other fixed points.

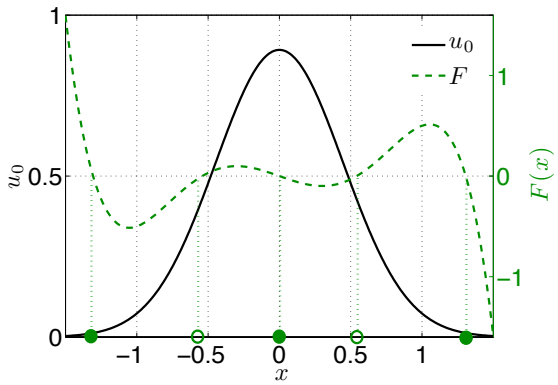

Figure 7.6: Subcritical model: initial density and right hand side of the ODE. Filled dots along the x-axis indicate the stable fixed points and circles the unstable fixed points.

The subcritical model with this initial distribution yields a solution with locally steep gra-

dients, which are challenging conditions for the adaptive scheme. Figure 7.7 (left) shows the third-order solution at $T = 0.5$ for a local tolerance of $TOL = 5 \cdot 10^{-6}$, a temporal tolerance factor $\rho = 0.9$, and a maximal time step $\tau_{\max} = 0.04$. To ensure volume conservation, the density is re-normalized after each integration step. The solution matches with that of the method of characteristics. The steep gradients at the outer stable fixed points are accurately resolved as well as the structure of the solution in between and beyond these points. Figure 7.7 (right) depicts the evolution of time steps and grid sizes (left), as well as of the

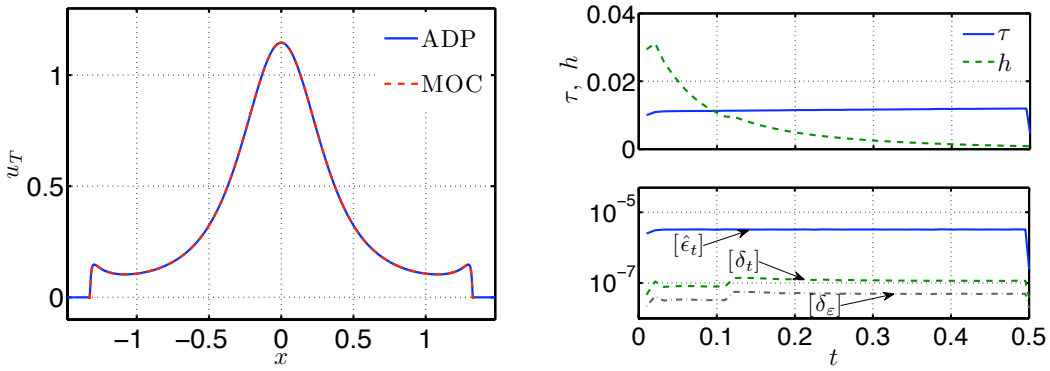

Figure 7.7: Subcritical model: $T = 0.5$, $k = 3$, $TOL = 5 \cdot 10^{-6}$, $\rho = 0.9$, $\tau_{\max} = 0.04$. Left: final density. Right: Evolution of time steps and grid sizes (upper panel), and of the local error estimates (lower panel).

local error estimates (right). Time steps stabilize rapidly due to low temporal dynamics, whereas the grid size continuously decreases as the structure of the solution becomes more challenging.

7.4 Michaelis-Menten kinetics with extended state space (two dimensions)

Last, we reconsider the first example of Michaelis-Menten kinetics. To demonstrate the extension of the method to two-dimensional problems, we consider $V_{\max}$ as an uncertain parameter; the state space is extended by $V_{\max}$. Biologically, it makes sense to consider $V_{\max}$ variable, since with $V_{\max} = k_{\text{cat}} \cdot \mathbf{E}_{\text{tot}}$ one can account for variability in the total enzyme concentration. Assuming that $V_{\max}$ is variable but constant in time, the extended ODE is given by

$$\begin{aligned} \dot{x} \quad &= -\frac{V_{\max}}{K_m + x} \cdot x \\ \dot{V}_{\max} &= 0\,. \end{aligned} \tag{7.4}$$

Setting $K_m = 1$, we start with an initial Gaussian distribution with mean and covariance matrix chosen as

$$\mu = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 1/8 & 0 \\ 0 & 1/40 \end{pmatrix}.$$

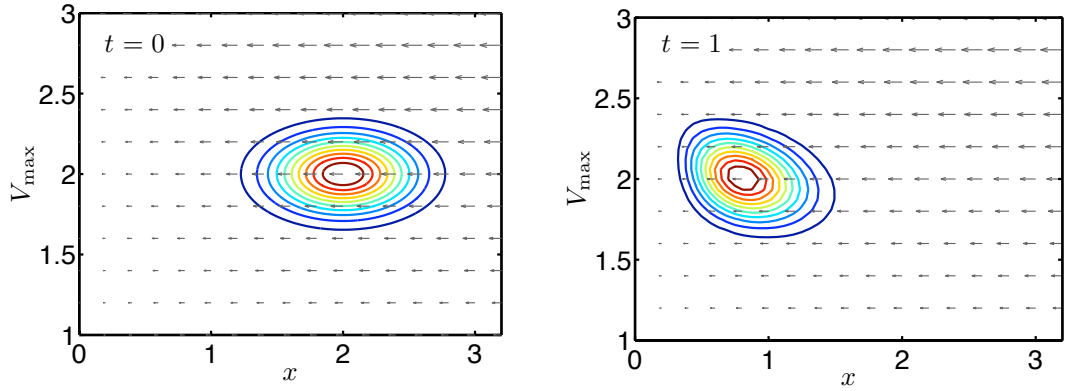

Figure 7.8: Extended Michaelis-Menten model: contour lines denote the initial (left) & final density (right) at $T = 1$, computed with $k = 2$, $TOL = 0.05$, $\rho = 0.9$ and $\tau_{\max} = 0.1$. The vector field of the ODE is shown with arrows.

Figure 7.8 (left) depicts the initial distribution by means of contour lines. The arrows indicate the vector field imposed by the right hand side of (7.4). The second-order solution is computed using $TOL = 0.05$, $\rho = 0.9$ and a maximal time step $\tau_{\max} = 0.1$. The final density at $T = 1$ is shown in Figure 7.8 (right). Since for larger maximal fluxes $V_{\max}$, the substance $\mathbf{X}$ is degraded faster, the solution develops asymmetries.

In Figure 7.9, the solution is compared to the solution obtained via the method of characteristics. The three-dimensional plot illustrates that the adaptive solution captures the

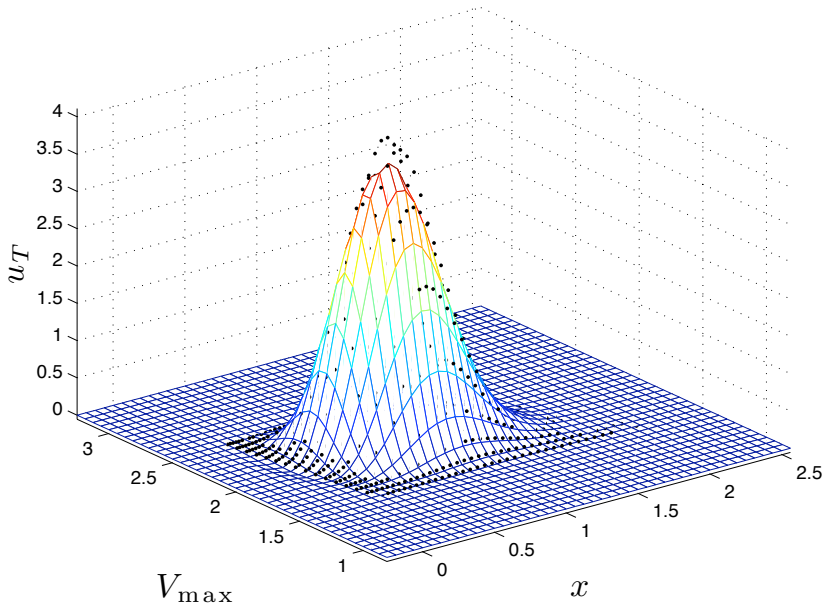

Figure 7.9: Extended Michaelis-Menten model: comparison of the solutions obtained via adaptive density propagation and the method of characteristics (indicated by black dots).

structure of the solution well, but differs slightly around the mode of the distribution and along the upcoming steep gradient. To quantify the difference, we evaluate our solution at the points given by the method of characteristics. The difference of the two solutions is illustrated by the heatmap in Figure 7.10 (left). The pointwise error remains below 0.4

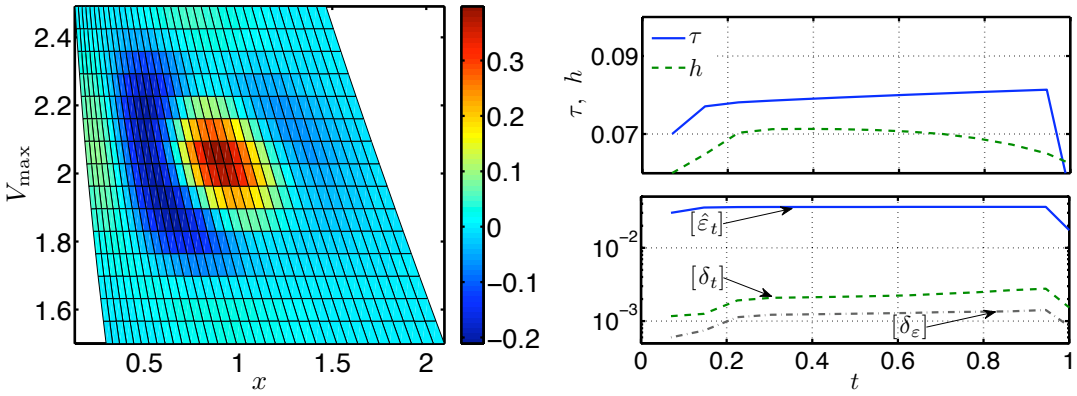

Figure 7.10: Extended Michaelis-Menten model. Left: heatmap of the difference between the two solutions. Right: evolution of time steps and grid sizes (upper panel), and of the local error estimates (lower panel).

and is considerably smaller at most of the points. Considering that the local tolerance is $TOL = 0.05$, pointwise errors of that order or magnitude in the final density are well acceptable.

The evolution of the time steps and grid sizes chosen by the adaptive scheme, as well as the estimated local errors are shown in Figure 7.10 (right). The time steps continuously increase, because temporal dynamics become slower as x decreases. Although this implies less constraints on the spatial accuracy, grid sizes decrease in the course of integration due to the development of steeper gradients as the distribution approaches the y-axis.

Concluding remarks In this chapter we examined nonlinear models that give rise to locally steep and bimodal distributions. The method has provided highly satisfactory results. Moreover, we note that in all the considered cases, the adaptive selection of time steps and grid sizes was very stable in time. We thus conclude that the adaptive framework can be applied to low-dimensional problems with high accuracy. An efficient extension of the applicability to higher-dimensional problems requires a reduction of spatial discretization costs, which will be a focus of upcoming work.

Part III

Summary & Outlook

Summary & Outlook

Summary In this thesis we developed a novel method for the global sensitivity analysis of ODEs. Assuming that the uncertainty & variability in the model input is captured by a known initial probability distribution, the problem can be recast as an ODE with random initial values. In this setting, the evolution of the probability density function associated with the random state variable is described by a first-order linear PDE. We exploited the PDE-based formulation, which gives access to a solid theory and methodology, to develop an error-controlled approach to sensitivity analysis. The presented method solves the PDE by combining an adaptive Rothe scheme with approximate approximations for spatial discretization. The Rothe scheme provides a framework for accurate error estimation and an adaptive choice of temporal and spatial discretization.

Many numerical approaches have been developed for global sensitivity analysis of ODEs. For higher-dimensional problems, these basically reduce to MC-based approaches or methods based on representations of the density in terms of heuristic Gaussian approximations. Both approaches suffer from the lack of reliable error estimates to perform error control. Our approach includes two main novelties: (1) The adaptive density propagation, i.e., an error-controlled solution of the related PDE, constitutes a new approach to the global sensitivity analysis of ODEs. (2) For the first time, approximate approximations were used to solve a time-dependent PDE in an adaptive and error-controlled Rothe context. The theoretical results obtained in this work clearly indicate how to implement the method efficiently.

We established and implemented a framework for adaptive density propagation with approximate approximations and studied its asymptotic properties. The method was shown to converge. Numerical examples in one and two space dimensions illustrated the theoretical results and showed that the method is applicable to nonlinear problems as well as problems that give rise to solutions with steep gradients or bimodal structure.

Our analysis further revealed dependencies between temporal and spatial discretization, imposing strong constraints on the spatial accuracy. An efficient solution of these constraints necessitates a high approximation order of the spatial discretization scheme. Compared to classical discretization methods such as finite element or finite volume methods, approximate approximations offer three substantial advantages:

1. Error estimates are readily available (avoiding computations of solutions of different approximation orders).
2. The approximation order can be increased at feasible computational costs, which allows for an efficient solution of the spatio-temporal accuracy constraints.
3. Although in this work we considered approximate approximations with basis functions positioned on a uniform grid, the concept is not restricted to those; it can be extended to transformations of uniform grids as well as unstructured, scattered grids.

This justifies our hope that the framework presented herein may constitute a first step towards error-controlled sensitivity analysis of higher-dimensional models.

Outlook To make the error-controlled approach competitive with common (ODE-based) sensitivity methods, even in higher-dimensions, costs of the spatial discretization have to be reduced, i.e. the number of basis functions. One possibility to reduce costs, while maintaining a uniform positioning of the basis functions, is a partitioning of the spatial domain, as illustrated in Figure 7.11 (left & middle). Adaptive density propagation may then be performed on each of the sub-domains with additional boundary conditions to account for the flow of probability between two neighboring sub-domains. Alternatively,

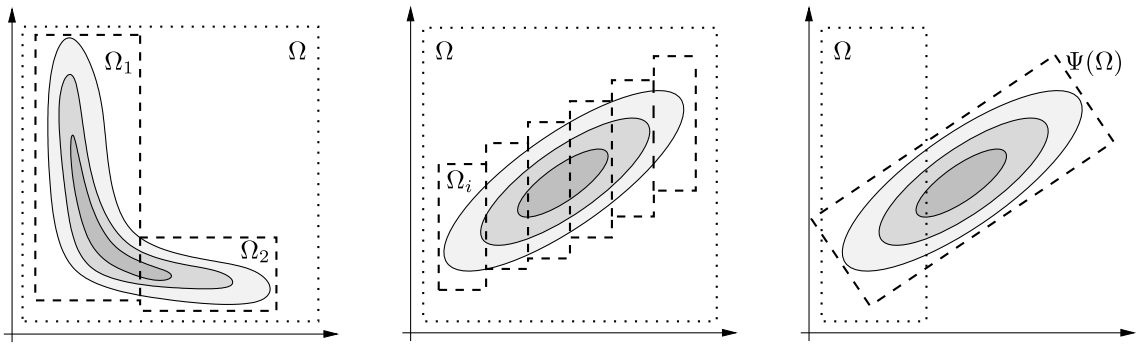

Figure 7.11: Partitioning and transformation of unifrom grids. Level lines represent the distribution, and the dotted box the discretization domain Ω. Left & middle: alternative approximation of the distribution on the sub-domains $\Omega_i \subset \Omega$, indicated by the dashed boxes. Right: a linearly transformed uniform grid $\Psi(\Omega)$.

costs can be reduced using basis functions that are positioned at a transformation of a uniform grid, see Figure 7.11 (right). Convergence results for approximate approximations based on transformed uniform grids are available, see [66]. However a serious attempt to extend the framework to high-dimensional problems can in our view only be based on sparse or scattered grids (in a meshfree setting). The scattered grids may be more dense in regions with fast dynamics or steep gradients, and sparse elsewhere. Such an approach requires reliable error estimates of approximate approximations with scattered grids. The derivation of those is a focus of ongoing research [44, 54].

Although the method has been developed in the context of global sensitivity analysis, it may also be amenable for an application in model assessment or model selection, see e.g. [34, 45, 59]. These tasks are generally complicated with deterministic models, since the model output either coincides with experimental data or not. An exact match of the model with the data is however unlikely, and a quantification of the mismatch remains a critical problem. For probabilistic models, the likelihood, i.e., the probability of the data under the specified model, is used to quantify a match or mismatch. Based on the likelihood, a broad range of statistical methods is available for model assessment and selection, see above references. The density propagation approach renders ODE models in a probabilistic context and yields an estimate of the likelihood function. Therefore, the method provides a link to the well-established methodology of statistical decision making.

Uncertainty and variability in ODE models is a general problem, and there is demand for numerical solutions that are applicable in high dimensions while maintaining a prescribed approximation quality. The thesis provides a theoretical framework as well as a fundamental understanding of adaptive density propagation based on a coupling of a Rothe method with approximate approximations. We believe that this is a sound basis to proceed towards an extension to higher-dimensional problems.

Appendix

Appendix A

Semi-discretization in time

In this chapter we introduce concepts for semi-discretization in time, which are based on considering the time-dependent PDE as an ODE in a function space. This view allows for applying the same discretization techniques as for ODEs. First, in Section A.1, we discuss properties, which the discrete solution has to satisfy to guarantee convergence of the discrete solution to the analytical solution. Then, in Section A.2, we show how temporal errors can be estimated and how time steps can be adapted accordingly. Throughout the chapter we assume spatially unperturbed solutions.

A.1 Approximation of the strongly continuous semigroup

We consider $u : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and a PDE of the form

$$\frac{\partial}{\partial t} u = \mathcal{A}\, u\,, \qquad u(0,\cdot) = u_0\,, \tag{A.1}$$

where $\mathcal{A}$ denotes a differential operator involving only spatial derivatives of u. Assume the analytical solution $u_t = u(t,\cdot)$ is given by $u_t = \mathcal{P}_t u_0$, where $\{\mathcal{P}_t\}_{t\geq 0}$ denotes the semigroup of evolution operators——in our case the semigroup of Frobenius-Perron operators. Analogously to the discretization of ODEs, the evolution operator $\mathcal{P}_t$ is approximated for a small time step $t = \tau > 0$ using approximations $r(z)$ to the exponential function, i.e.

$$\mathcal{P}_\tau \approx R_\tau := r(\tau\mathcal{A})\,, \tag{A.2}$$

and the approximation quality of $r(z) \approx e^z$ allows for conclusions on the approximation quality of $R_\tau \approx \mathcal{P}_\tau$. A discrete solution is then obtained by a recursive application of the discrete evolution operator R_τ via

$$R_\tau^j u_0 = R_\tau \left(R_\tau^{j-1} u_0\right) \approx \mathcal{P}_\tau^j u_0\,, \quad R_\tau^0 u_0 := u_0\,, \qquad j = 1, \ldots, {}^T\!/\!_\tau\,. \tag{A.3}$$

As τ vanishes, the discrete solution should converge to the analytical solution, where convergence is understood on two levels: convergence of the errors made in each time steps, the *local* errors, and convergence of the iterated application of R_τ, the *global* discretization error. Ideally, these errors converge with a certain speed. To facilitate the following definitions, we introduce the Landau symbol $\mathcal{O}$:

Definition A.1.1 (**$\mathcal{O}$-Notation**). *A function ν is said to be*

$$\nu(z) = \mathcal{O}(g(z)) \quad as \quad z \to z^*\,,$$

if for all $\varepsilon > 0$ there exists a constant $c > 0$ and z' such that for all z, $|z - z^*| \leq |z' - z^*|$

$$\nu(z) \leq c \cdot g(z)\,.$$

Generally, asymptotics for $z^* = 0$ or $z^* = \infty$ are of interest.

Definition A.1.2 (**Consistency**). *The local error*

$$\varepsilon_t(u,\tau) := \mathcal{P}_\tau u_t - R_\tau u_t \tag{A.4}$$

is called consistency error. *The discrete evolution R_τ^j, $j = 1,\ldots,{}^T\!/_\tau$, is called* consistent, *if for all $t \in [0,T]$*

$$\lim_{\tau\to 0} ||\varepsilon_t(u,\tau)|| = 0\,. \tag{A.5}$$

Furthermore, the discrete evolution has consistency order *k, if for all $t \in [0,T]$*

$$||\varepsilon_t(u,\tau)|| = \mathcal{O}(\tau^{k+1}) \quad as \quad \tau \to 0\,. \tag{A.6}$$

We will write $\varepsilon_t(\tau)$ instead of $\varepsilon_t(u,\tau)$, whenever the function u is clear from the context.

Definition A.1.3 (**Convergence**). *The discrete evolution $R_\tau^j u(x,0)$, $j = 1,\ldots,{}^T\!/_\tau$ is called* convergent *if the global approximation error vanishes, i.e.*

$$\max_{j=1,\ldots,{}^T\!/_\tau} \left(\lim_{\tau\to 0} \left|\left|\mathcal{P}_{t_j} u_0 - R_\tau^j u_0\right|\right|\right) = 0\,, \quad t_j = j\cdot\tau\,. \tag{A.7}$$

Furthermore, the discrete evolution has convergence order *k, if*

$$\max_{j=1,\ldots,{}^T\!/_\tau} \left|\left|\mathcal{P}_{t_j} u_0 - R_\tau^j u_0\right|\right| = \mathcal{O}(\tau^k) \quad as \quad \tau \to 0\,. \tag{A.8}$$

Thus, consistency of the discrete evolution refers to convergence of the local errors, whereas convergence refers to the convergence of the global discretization error.

Definition A.1.4 (**Discretization order**). *If the discrete evolution $R_\tau^j u_0$, $j = 1,\ldots,{}^T\!/_\tau$, is consistent and convergent with order k, the semi-discretization scheme R_τ is said to have* discretization order *k; we then denote the discrete solution by*

$$u_{t_{j+1}}^{(k)} := R_\tau u_{t_j}^{(k)}\,, \qquad u_0^{(k)} := u_0\,. \tag{A.9}$$

Definition A.1.5 (***A*-stability**). *An approximation $r(z)$ to e^z is called* A-stable, *if its stability region $\{z \in \mathbb{C},\ r(z) \leq 1\}$ contains the stability region of of the exponential function, i.e.*

$$|r(z)| \leq 1\,, \quad \forall\, z \in \mathbb{C},\ \Re(z) \leq 0\,. \tag{A.10}$$

The semi-discretization scheme R_τ defined by $r(t\mathcal{A})$ is called A-stable, if r is A-stable. Analogously, the discrete evolution $R_\tau^j u_0$ is called A-stable, if r is A-stable.

The following result by Brenner & Thomée [11] ensures convergence of the discrete evolution, if r is consistent and A-stable.

Theorem A.1.6 (**Rational approximation of semigroups**, [11, Theorem 3])**.** *Let $\mathcal{A}$ generate a* strongly continuous semigroup $\mathcal{P}_t = e^{t\mathcal{A}}$, $t > 0$, *i.e.*

$$\lim_{t\to 0} ||\mathcal{P}_t u - \mathcal{P}_0 u|| = 0\,, \quad \forall\, u \in D(\mathcal{P}_t)\,, \tag{A.11}$$

and let further

$$||\mathcal{P}_t|| \leq 1\,, \quad \forall\, t \geq 0\,. \tag{A.12}$$

*Then for any A-*stable *rational approximation $r(z)$ to e^z* of consistency order k *there is a constant $c > 0$ such that for all $u \in D(\mathcal{A}^{k+1})$*

$$||\mathcal{P}_{t_n} u - R_\tau^n u|| \leq c \cdot t_n\, \tau^k \left|\left|\mathcal{A}^{k+1} u\right|\right|, \quad t_n = n \cdot \tau,\, \tau > 0,\, n \in \mathbb{N}\,. \tag{A.13}$$

Proof can be found in [11] (note that condition (A.12) refers to [11, Eq. (7)'] with $C_0 = 1$ and $\omega = 0$). The semigroup of Frobenius-Perron operators $\mathcal{P}_t$, $t \geq 0$, is strongly continuous, see [55, Remark 7.6.2], and further condition (A.12) is satisfied, because $\mathcal{P}_t$ is a Markov operator. Therefore, the above theorem ensures that a discrete solution converges with order k to the analytical solution, if the rational function satisfies two conditions: consistency of order k and A-stability.

Definition A.1.7 (***L*-stability**)**.** *An A-stable approximation $r(z)$ to e^z that satisfies*

$$\lim_{z\to -\infty} r(z) = 0 \tag{A.14}$$

is called L-stable*, and a semi-discretization scheme defined by $R_\tau = r(\tau\mathcal{A})$ is called L-stable, if r is L-stable.*

While A-stability ensures that the discrete solution $R_\tau^j u_0$ inherits properties of the analytical solution $\mathcal{P}_{\tau^n} u_0$ for a small time step $\tau > 0$, L-stability ensures that asymptotic properties of the true solution in *one, large* time step are inherited. Therefore, L-stable semi-discretization schemes allow for larger time steps.

A.2 Adaptive time step selection

So far we have considered a discrete evolution with a constant time step $\tau > 0$. Now we consider discrete solutions $u^{(k)}$ of order k that are computed using variable time steps $\tau_j > 0$ in each integration step t_j, $j = 1, \ldots, n$, i.e.

$$u_{t_{j+1}}^{(k)} := R_{\tau_j} u_{t_j}^{(k)}\,, \qquad u_0^{(k)} := u_0\,.$$

Ideally, the time steps τ_j are chosen such that the local errors satisfy

$$\left|\left|\varepsilon_{t_j}(u^{(k)}, \tau_j)\right|\right| = \left|\left|\mathcal{P}_{\tau_j} u_{t_j}^{(k)} - R_{\tau_j} u_{t_j}^{(k)}\right|\right| \leq TOL_t\,, \qquad j = 0, \ldots, n-1\,, \tag{A.15}$$

where $TOL_t > 0$ denotes a specified temporal accuracy or tolerance. Since $\mathcal{P}_{\tau_j}$ is unknown, the local errors have to be estimated, and the tolerance condition can only be satisfied for the estimates. In the following we show how to estimate the local errors and adapt the time steps during integration such that the estimates satisfy condition (A.15). We consider the

standard time step selection strategy as presented in [30, Chapter II.4] and [25, Chapter 5].

Let $u^{(k-1)}$ denote a discrete solution of order $k-1$ defined by

$$u_{t_{j+1}}^{(k-1)} := Q_{\tau_j} u_{t_j}^{(k)} ,$$

where Q_τ denotes a rational approximation to the strongly continuous semigroup of order $k-1$. Then the local discretization errors of $u^{(k-1)}$ are estimated by the difference of the two discrete solutions of different discretization order, i.e.

$$\varepsilon_{t_j}^{(k-1)}(\tau_j) := u_{t_{j+1}}^{(k)} - u_{t_{j+1}}^{(k-1)} \tag{A.16}$$

In terms of computational costs, a time step τ^* is considered optimal with respect with to the tolerance TOL_t, if it satisfies condition (A.15) and

$$\left|\left|\varepsilon_{t_j}^{(k-1)}(\tau^*)\right|\right| \approx TOL_t \,. \tag{A.17}$$

Since $u^{(k)}$ and $u^{(k-1)}$ are consistent with order k and $k-1$, the error estimate (A.16) decays with $\mathcal{O}(\tau^k)$ as τ vanishes (same as the true local error of $u^{(k-1)}$), which implies

$$\frac{c \cdot \tau^{*k}}{c \cdot \tau_j^k} \approx \frac{TOL_t}{||\varepsilon_{t_j}^{(k-1)}(\tau_j)||} ,$$

for some constant $c > 0$. Consequently, the optimal time step τ^* is given by

$$\tau^* = \sqrt[k]{\frac{TOL_t}{||\varepsilon_{t_j}^{(k-1)}(\tau_j)||}} \cdot \tau_j \,, \tag{A.18}$$

where in practice, (A.18) is multiplied by a safety factor $0 < \sigma < 1$.

In case τ_j does not satisfy the tolerance condition (A.15), the integration step is repeated using $\tau_j = \tau^*$. In case the tolerance condition is satisfied, τ_j is accepted, and integration proceeds with the next integration step using the time step $\tau_{j+1} = \tau^*$.

Appendix B

Derivation of spatial error estimates

In this chapter we show how the local spatial error of a solution $\hat{u}_t^{(k)}$ can be estimated using the error estimates that are given by the spatial discretization scheme used to solve each of the stationary spatial problems. Using the corrections $\Delta u_t^{(k)}$, the solutions of order k are given by

$$u_t^{(k)} \;=\; u_t^{(k-1)} + \Delta u_t^{(k-1)} \,. \tag{B.1}$$

We first consider the A-stable scheme, where the corrections up to $k-1=2$ are computed multiplicatively by the recursion

$$\begin{aligned} (\mathrm{Id} - \tau\mathcal{A})\ \Delta u_t^{(0)} &= (\tau\mathcal{A})\ u_t^{(0)} \\ (\mathrm{Id} - \tau\mathcal{A})\ \Delta u_t^{(1)} &= -\frac{1}{2}(\tau\mathcal{A})\ \Delta u_t^{(0)} \\ (\mathrm{Id} - \tau\mathcal{A})\ \Delta u_t^{(2)} &= -\frac{1}{3}(\tau\mathcal{A})\ \Delta u_t^{(1)} \,. \end{aligned} \tag{B.2}$$

The spatially perturbed solution $\hat{u}_t^{(k+1)}$ of order $k+1$ can be written as

$$\hat{u}_t^{(k+1)} \;=\; \hat{u}_t^{(k)} + \Delta\hat{u}_t^{(k)} = u_t^{(k+1)} + \delta_t^{(k+1)} \,. \tag{B.3}$$

Furthermore, the spatial perturbation of the local temporal error estimate is given by

$$\delta_\varepsilon^{(k)}(t) \;:=\; \hat{\varepsilon}_t^{(k)} - \varepsilon_t^{(k)} = \Delta\hat{u}_t^{(k)} - \Delta u_t^{(k)} =: \Delta\delta_t^{(k)} \,. \tag{B.4}$$

(We denote the difference of the corrections by $\Delta\delta_t^{(k)}$, since $\Delta\hat{u}_t^{(k)} - \Delta u_t^{(k)} = \delta_t^{(k+1)} - \delta_t^{(k)}$.) Hence, we estimate $\delta_\varepsilon(t)$ by

$$[\delta_\varepsilon^{(k)}](t) = \Delta\delta_t^{(k)} \,. \tag{B.5}$$

Using the above relations, we can state the following identities for the perturbed correction $\Delta\hat{u}_t^{(k)}$ and $k = 1, 2$:

$$\begin{aligned}
\Delta\hat{u}_t^{(0)} &= (\mathrm{Id} - \tau\mathcal{A})^{-1}\ u_t^{(0)} + \mathrm{err}^{(1)} = \Delta u_t^{(0)} + \Delta\delta_t^{(0)} && \text{(B.6)}\\
\Delta\hat{u}_t^{(1)} &= -\frac{1}{2}(\tau\mathcal{A})(\mathrm{Id} - \tau\mathcal{A})^{-1}\ \Delta\hat{u}_t^{(0)} + \mathrm{err}^{(2)} && \text{(B.7)}\\
&= -\frac{1}{2}(\tau\mathcal{A})^2(\mathrm{Id} - \tau\mathcal{A})^{-2}\ u_t^{(0)}\\
&\quad -\frac{1}{2}(\tau\mathcal{A})(\mathrm{Id} - \tau\mathcal{A})^{-1}\ \Delta\delta_t^{(0)} + \mathrm{err}^{(2)}\\
&\overset{\text{by (B.2)}}{=} \Delta u_t^{(1)} \underbrace{-\frac{1}{2}}_{\gamma_A^{(1)}}(\tau\mathcal{A})(\mathrm{Id} - \tau\mathcal{A})^{-1}\ \Delta\delta_t^{(0)} + \mathrm{err}^{(2)}\\
&= \Delta u_t^{(1)} + \Delta\delta_t^{(1)}\,,
\end{aligned}$$

where $\mathrm{err}^{(k)}$ denotes the approximation error in the solution of the k-th spatial problem. Spatially perturbed corrections of order $k \geq 2$ can be computed recursively by

$$\begin{aligned}
\Delta\hat{u}_t^{(k)} &= \Delta u_t^{(k)} + \Delta\delta_t^{(k)} && \text{(B.8)}\\
&= \gamma_A^{(k)}(\tau\mathcal{A})(\mathrm{Id} - \tau\mathcal{A})^{-1}\ \Delta\hat{u}_t^{(k-1)}\\
&\quad +\mathrm{err}^{(k+1)} && \text{(B.9)}\\
\Rightarrow \Delta\delta_t^{(k)} &\overset{\text{by (B.2)}}{=} \gamma_A^{(k)}(\tau\mathcal{A})(\mathrm{Id} - \tau\mathcal{A})^{-1}\ (\Delta\hat{u}_t^{(k-1)} - \Delta u_t^{(k-1)})\\
&\quad +\mathrm{err}^{(k+1)}\,, && \text{(B.10)}
\end{aligned}$$

which yields

$$\boxed{\Delta\delta_t^{(k)} = \gamma_A^{(k)} \cdot \left[(\tau\mathcal{A})(\mathrm{Id} - \tau\mathcal{A})^{-1}\right]\ \Delta\delta_t^{(k-1)}\mathrm{err}^{(k+1)}\,.} \qquad \text{(B.11)}$$

Consequently, we obtain the following recursion for the estimates of $\delta_\varepsilon^{(k)}$

$$\boxed{[\delta_\varepsilon^{(k+1)}] = |\gamma_A^{(k+1)}| \cdot [\delta_\varepsilon^{(k)}] + [\mathrm{err}^{(k+2)}]\,, \quad [\delta_\varepsilon^{(0)}] = \mathrm{err}^{(1)}\,.} \qquad \text{(B.12)}$$

Furthermore, the spatial errors $\delta_t^{(k+1)}$ can be computed recursively by

$$\begin{aligned}
\delta_t^{(k+1)} &\overset{\text{by (B.3)}}{=} \hat{u}_t^{(k)} - u_t^{(k+1)} + \Delta\hat{u}_t^{(k)} && \text{(B.13)}\\
&= u_t^{(k)} + \delta_t^{(k)} - (u_t^{(k)} + \Delta u_t^{(k)}) + \Delta\hat{u}_t^{(k)}\\
&= \delta_t^{(k)} + (\Delta\hat{u}_t^{(k)} - \Delta u_t^{(k)})\,,
\end{aligned}$$

such that

$$\boxed{\delta_t^{(k+1)} = \delta_t^{(k)} + \Delta\delta_t^{(k)}\,, \quad \delta_t^{(0)} = 0.} \qquad \text{(B.14)}$$

and

$$\boxed{[\delta_t^{(k+1)}] = [\delta_t^{(k)}] + [\delta_\varepsilon^{(k)}]\,, \quad [\delta_t^{(0)}] = 0.} \qquad \text{(B.15)}$$

Combining (B.12) and (B.15), explicit error estimates for $k = 1, 2, 3$ are as follows.

Estimators explicitly for A-stable method:

$$\begin{aligned}
\delta_\varepsilon^{(0)} &= \mathrm{err}^{(1)} & \qquad \delta_t^{(1)} &= \mathrm{err}^{(1)} \\
\delta_\varepsilon^{(1)} &= \tfrac{1}{2}\cdot \mathrm{err}^{(1)} + \mathrm{err}^{(2)} & \qquad \delta_t^{(2)} &= \tfrac{3}{2}\cdot \mathrm{err}^{(1)} + \mathrm{err}^{(2)} \\
\delta_\varepsilon^{(2)} &= \tfrac{1}{6}\cdot \mathrm{err}^{(1)} + \tfrac{1}{3}\cdot \mathrm{err}^{(2)} + \mathrm{err}^{(3)} & \qquad \delta_t^{(3)} &= \tfrac{5}{3}\cdot \mathrm{err}^{(1)} + \tfrac{4}{3}\cdot \mathrm{err}^{(2)} + \mathrm{err}^{(3)}\,.
\end{aligned} \tag{B.16}$$

The derivation of error estimates within the L-stable scheme is analogous and thus omitted. For $k = 1, 2, 3$, the error estimates are explicitly given by

Estimators explicitly for L-stable method:

$$\begin{aligned}
\delta_\varepsilon^{(0)} &= \mathrm{err}^{(1)} & \qquad \delta_t^{(1)} &= \mathrm{err}^{(1)} \\
\delta_\varepsilon^{(1)} &= \tfrac{1}{2}\cdot \mathrm{err}^{(1)} + \mathrm{err}^{(2)} & \qquad \delta_t^{(2)} &= \tfrac{3}{2}\cdot \mathrm{err}^{(1)} + \mathrm{err}^{(2)} \\
\delta_\varepsilon^{(2)} &= \tfrac{2}{3}\cdot \mathrm{err}^{(1)} + \tfrac{4}{3}\cdot \mathrm{err}^{(2)} + \mathrm{err}^{(3)} & \qquad \delta_t^{(3)} &= \tfrac{13}{6}\cdot \mathrm{err}^{(1)} + \tfrac{7}{3}\cdot \mathrm{err}^{(2)} + \mathrm{err}^{(3)}\,.
\end{aligned} \tag{B.17}$$

Appendix C

Derivatives of the generating functions

The generating functions $\eta^{(2)}(x)$, $\eta^{(4)}(x)$ and $\eta^{(6)}(x)$ can be computed from Theorem (4.4.1) as

$$\eta^{(2)}(x) = \pi^{-d/2} \cdot e^{-|x|_2^2}, \tag{C.1}$$

$$\eta^{(4)}(x) = \pi^{-d/2} \cdot \left(\left(1 + \frac{d}{2} \right) - |x|_2^2 \right) \cdot e^{-|x|_2^2}, \tag{C.2}$$

$$\eta^{(6)}(x) = \pi^{-d/2} \cdot \left(\frac{1}{2} \left(2 + \frac{d}{2} \right) \left(1 + \frac{d}{2} \right) - \left(2 + \frac{d}{2} \right) |x|_2^2 + \frac{1}{2} |x|_2^2 \right) \cdot e^{-|x|_2^2}. \tag{C.3}$$

From the definition of the generator $\mathcal{A}$ in (1.18), the following identities can be derived

$$\mathcal{A}\eta = -\operatorname{div}(F) \cdot \eta - \langle F, \nabla\eta \rangle, \tag{C.4}$$

$$\begin{aligned} \mathcal{A}^2\eta = \quad & \operatorname{div}(F)^2 \cdot \eta + 2 \cdot \operatorname{div}(F) \cdot \langle F, \nabla\eta \rangle \\ & - \langle F, DF^T \nabla\eta \rangle + \langle F, \operatorname{hess}(\eta) \cdot F \rangle, \end{aligned} \tag{C.5}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product, $\operatorname{div}(F)$ the divergence of F, $\nabla\eta$ is the gradient of $\eta(x)$, $\operatorname{hess}(\eta)$ the Hessian matrix of $\eta(x)$, and DF the Jacobian of $F(x)$. The action of the generator $\mathcal{A}$ on the generating functions can be computed by using the above relations. The first order partial derivatives of the generating functions are given by

$$\frac{\partial}{\partial x_i} \eta^{(2)}(x) = -2x_i \cdot \eta^{(2)}(x), \tag{C.6}$$

$$\frac{\partial}{\partial x_i} \eta^{(4)}(x) = -2x_i \cdot \left(\pi^{-d/2} e^{-|x|_2^2} + \eta^{(4)}(x) \right), \tag{C.7}$$

$$\frac{\partial}{\partial x_i} \eta^{(6)}(x) = -2x_i \cdot \left(\pi^{-d/2} e^{-|x|_2^2} \left(2 + \frac{d}{2} - |x|_2^2 \right) + \eta^{(6)}(x) \right). \tag{C.8}$$

These derivatives can be used to compute $\mathcal{A}\eta$ according to (1.18). The computation of $\mathcal{A}^2\eta$ requires the second-order mixed derivatives of the generating functions, which are given

by:

$$\frac{\partial^2}{\partial x_j \partial x_i}\eta^{(2)}(x) = -\delta_{ij} \cdot 2 \cdot \eta^{(2)}(x) + 4x_i x_j \cdot \eta^{(2)}(x) \tag{C.9}$$

$$\begin{aligned}\frac{\partial^2}{\partial x_j \partial x_i}\eta^{(4)}(x) = &-\delta_{ij} \cdot 2 \cdot \left(\pi^{-d/2} e^{-|x|_2^2} + \eta^{(4)}(x)\right) \\ &+ 4x_i x_j \cdot \left(2 \cdot \pi^{-d/2} e^{-|x|_2^2} + \eta^{(4)}(x)\right)\end{aligned} \tag{C.10}$$

$$\begin{aligned}\frac{\partial^2}{\partial x_j \partial x_i}\eta^{(6)}(x) = &-\delta_{ij} \cdot 2 \cdot \left(\pi^{-d/2} e^{-|x|_2^2} \left(2 + \frac{d}{2} - |x|_2^2\right) + \eta^{(6)}\right) \\ &+ 4x_i x_j \cdot \left(\pi^{-d/2} e^{-|x|_2^2} \cdot \left(5 + d - 2 \cdot |x|_2^2\right) + \eta^{(6)}(x)\right) ,\end{aligned} \tag{C.11}$$

where $\delta_{ij} = 1$, if $i = j$ and 0 otherwise.

Zusammenfassung

Gewöhnliche Differentialgleichungen nehmen eine essentielle Stellung in der mathematischen Modellierung ein. Als Voraussetzung für zuverlässige Resultate muss sowohl in der Modellbildung als auch in der Analyse des Modells der Einfluss von Unsicherheit und/oder Variabilität in den Eingabedaten berücksichtigt werden. Mit Hilfe von Sensitivitätsanalyse wird untersucht, wie sich Unsicherheit und Variabilität durch die Modelldynamik ausbreiten und sich somit auf die Ausgabedaten auswirken. Globale Sensitivitätsanalyse untersucht die Auswirkungen von Abweichungen in den Eingabedaten, die sich möglicherweise über den gesamten Zustandsraum erstrecken. Zwei Probleme, die die globale Analyse erschweren, sind hohe Dimensionen und eine Kontrolle der Genauigkeit, mit der die Ausgabeunsicherheit geschätzt wird. Die meisten numerischen Ansätze konzentrieren sich derzeitig darauf, die Analyse von hoch-dimensionalen Problemen effizienter zu gestalten. Inwiefern die geschätzte Ausgabeunsicherheit dabei der tatsächlichen Ausgabeunsicherheit entspricht, bleibt jedoch meist unklar.

In dieser Arbeit wird ein neuer Ansatz zur globalen Sensitivitätsanalyse von gewöhnlichen Differentialgleichungen vorgestellt. Hauptmerkmal dieses Ansatzes ist eine adaptive Schätzung der Ausgabeunsicherheit, bei der der Approximationsfehler automatisch kontrolliert wird. Dafür bedienen wir uns einer äquivalenten Formulierung des Problems, in der die zeitliche Entwicklung der Wahrscheinlichkeitsdichte der unsicheren Zustandsvariablen durch eine partielle Differentialgleichung beschrieben wird. Zur Lösung dieser Differentialgleichung kombinieren wir neue Ansätze aus Numerik und Approximationstheorie. Die hier vorgestellte Methode kontrolliert den Approximationsfehler, indem sowohl die Zeit- als auch die Ortsdiskretisierung angepasst wird. Wir verwenden ein Rothe-Verfahren, das einen angemessenen Kontext für die separate Schätzung von Zeit- und Ortsfehlern schafft, so dass die Diskretisierung entsprechend adaptiert werden kann. Für die Ortsdiskretisierung verwenden wir *Approximate Approximations*, eine neu eingeführte Approximationsmethode, die hier zum ersten Mal im Rahmen eines adaptiven Rothe-Verfahrens eingesetzt wird.

Wir analysieren die Konvergenz des Verfahrens und untersuchen, wie sich *Approximate Approximations* für die adaptive Lösung der Ortsprobleme eignen. Wir zeigen, dass das Verfahren konvergiert. Darüber hinaus geben die theoretischen Resultate direkt Aufschluss darüber, wie eine effiziente Implementierung realisiert werden kann. Die Ergebnisse werden anhand von numerischen Beispielen illustriert, die auch zeigen, dass das Verfahren eine hohe Genauigkeit bei der Schätzung der Ausgabeunsicherheiten erzielt. Desweiteren erweisen sich *Approximate Approximations* als vorteilhaft innerhalb des adaptiven Verfahrens, da sowohl Fehlerschätzer als auch Approximationen hoher Ordnung zu vertretbaren Rechenzeiten verfügbar sind. Aktuelle Fortschritte in der Theorie von *Approximate Approximations*, beruhend auf einer gitterfreien Diskretisierung, lassen außerdem darauf hoffen, dass sich das in dieser Arbeit vorgestellte Konzept auch auf höher-dimensionale Probleme übertragen lässt.

List of Figures

Bibliography

[1] S. Adjerid, J. E. Flaherty, and I. Babuška. A posteriori error estimation for the finite element method-of-lines solution of parabolic problems. In *Mathematical Models and Methods in Applied Sciences*, pages 261–286, 1999.

[2] V. I. Arnold. *Ordinary Differential Equations*. Springer, Berlin, 3rd edition, 2006.

[3] I. Babuška, U. Banerjee, and J. E. Osborn. Generalized finite element methods—main ideas, results and perspective. *International Journal of Computational Methods*, 1(1):67–103, 2004.

[4] H. A. Barton, W. A. Chiu, R. W. Setzer, M. E. Andersen, A. J. Bailer, F. Y. Bois, R. S. DeWoskin, S. Hays, G. Johanson, N. Jones, G. Loizou, R. C. MacPhail, C. J. Portier, M. Spendiff, and Y.-M. Tan. Characterizing uncertainty and variability in physiologically based pharmacokinetic models: State of the science and needs for research and implementation. *Toxicological Sciences*, 99(2):395–402, 2007.

[5] P. Bernillon and F. Y. Bois. Statistical issues in toxicokinetic modeling: A Bayesian perspective. *Environmental Health Perspectives*, 108 Suppl 5:883–893, Oct 2000.

[6] A. Beuter, L. Glass, M. C. Mackey, and M. S. Titcombe, editors. *Nonlinear Dynamics in Physiology and Medicine*. Interdisciplinary Applied Mathematics. Springer, 2003.

[7] F. Y. Bois. Applications of population approaches in toxicology. *Toxicology Letters*, 120:385–394, 2001.

[8] F. A. Bornemann. An adaptive multilevel approach to parabolic equations. *IMPACT of Computing and Science in Engineering*, 3(2):93–122, 1991.

[9] F. A. Bornemann. *An Adaptive Multilevel Approach to Parabolic Equations in Two Space Dimensions*. PhD thesis, Freie Universität Berlin, Departement of Mathematics & Computer Science, 1991.

[10] G. E. P. Box and N. R. Draper. *Response Surfaces, Mixtures, and Ridge Analyses; 2nd ed.* Wiley, Hoboken, NJ, 2007.

[11] P. Brenner and V. Thomée. On rational approximation of semigroups. *SIAM Journal on Numerical Analysis*, 16(4):683–694, 1979.

[12] H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004.

[13] F. Campolongo, A. Saltelli, and S. Tarantola. Sensitivity anaysis as an ingredient of modeling. *Statistical Science*, 15(4):377–395, 2000.

[14] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods: Evolution to Complex Geometries and Applications to Fluid Dynamics*. Scientific Computation. Springer, Berlin-Heidelberg, 2006.

[15] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods: Fundamentals in Single Domains*. Scientific Computation. Springer, Berlin-Heidelberg, 2007.

[16] C. Chicone. *Ordinary Differential Equations with Applications*. Springer, 1999.

[17] K. Cho, S. Shin, W. Kolch, and O. Wolkenhauer. Experimental design in systems biology, based on parameter sensitivity analysis using a Monte Carlo method: A case study for the TNF alpha-mediated NF-kappa B signal transduction pathway. *SIMULATION–Transactions of The Society for Modeling and Simulation International*, 79(12):726–739, 2003.

[18] A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Portland Press, third edition, 2004.

[19] V. Costanza and J. H. Seinfeld. Stochastic sensitivity analysis in chemical kinetics. *Journal of Chemical Physics*, 74(7):3852–3858, 1981.

[20] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume II. Wiley-Interscience, 1962.

[21] K. Cowles and B. P. Carlin. Markov Chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, 91:883–904, 1996.

[22] R. I. Cukier, C. M. Fortuin, K. E. Shuler, A. G. Petschek, and J. H. Schaibly. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *Journal of Chemical Physics*, 59(8), 1973.

[23] X. Darzacq, Y. Shav-Tal, V. de Turris, Y. Brody, S. M. Shenoy, R. D. Phair, and R. H. Singer. In vivo dynamics of RNA polymerase II transcription. *Nature Structural & Molelucar Biology*, 14(9):796–806, 09 2007.

[24] H. de Jong. Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, 9(1):67–103, 2002.

[25] P. Deuflhard and F. A. Bornemann. *Scientific Computing with Ordinary Differential Equations*. Springer, 2002.

[26] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 4th edition, 1998.

[27] G. E. Fasshauer. Toward approximate moving least squares approximation with irregularly spaced centers. *Computer Methods in Applied Mechanics & Engineering*, 193:1231–1243, 2004.

[28] G. E. Fasshauer. *Meshfree Approximation Methods with MATLAB*. World Scientific Publishing Co Pte Ltd, 2007.

[29] C. Grossmann and H.-G. Roos. *Numerical Treatment of Partial Differential Equations*. Springer, Berlin-Heidelberg, 2007.

[30] E. Hairer, S. Nørsett, and G. Wanner. *Solving ordinary differential equations I. Nonstiff problems.* Springer, Berlin, 1987.

[31] E. Hairer and G. Wanner. *Solving ordinary differential equations II. Stiff and Differential-Algebraic Problems.* Springer, Berlin, 1991.

[32] J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods.* Methuen, London and John Wiley & Sons, New York, 1964.

[33] J. Hartinger and R. Kainhofer. Non-uniform low-discrepancy sequence generation and integration of singular integrands. In Niederreiter and Talay [71].

[34] T. Hastie, R. Tibshirani, and J. H. Friedman. *The Elements of Statistical Learning.* Springer, August 2001.

[35] R. Heinrich and S. Schuster. *The Regulation Of Cellular Systems.* Springer, 2nd edition, 2006.

[36] J. C. Helton and F. J. Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering and System Safety*, 81:23–69, 2003.

[37] W. J. Hill and W. G. Hunter. A review of response surface methodology: A literature survey. *Technometrics*, 8(4):571–590, 1966.

[38] E. Hlawka. Funktionen beschränkter Variation in der Theorie der Gleichverteilung. *Annali di Matematica Pura ed Applicata*, 54(1):325–333, 1961.

[39] I. Horenko. *Modeling and Numerical Simulation of Quantum Effects in Molecular Dynamics.* PhD thesis, Freie Universität Berlin, Department of Mathematics & Computer Science, 2003.

[40] I. Horenko, S. Lorenz, C. Schütte, and W. Huisinga. Adaptive approach for nonlinear sensitivity analysis of reaction kinetics. *Journal of Computational Chemistry*, 26(9):941–948, 2005.

[41] I. Horenko and M. Weiser. Adaptive integration of molecular dynamics. *Journal of Computational Chemistr*, 24:1921–1929, 2003.

[42] I. Horenko, M. Weiser, B. Schmidt, and C. Schütte. Fully adaptive propagation of the quantum-classical liouville equation. *J Chem Phys*, 120(19):8913–8923, May 2004.

[43] N. V. Hritonenko and Y. P. Yatsenko. *Mathematical Modeling in Economics, Ecology and the Environment (Applied Optimization).* Springer, 1999.

[44] T. Ivanov, V. Maz'ya, and G. Schmidt. Boundary layer approximate approximations and cubature of potentials in domains. *Advances in Computational Mathematics*, 10:311–342, 1999.

[45] E. T. Jaynes. *Probability Theory, The Logic of Science.* Cambridge University Press, 2003.

[46] J. Kačur. *Method of Rothe in Evolution Equations.* Teubner, Leipzig, 1985.

[47] V. Karlin and V. Maz'ya. Time-marching algorithms for initial-boundary value problems based upon "approximate approximations". *BIT*, 35:548–560, 1995.

[48] V. Karlin and V. Maz'ya. Time-marching algorithms for nonlocal evolution equations based upon "approximate approximations". *SIAM Journal on Scientific Computing*, 18(3):736–752, 1997.

[49] J. Keener and J. Sneyd. *Mathematical Physiology*. Springer, 2001.

[50] R. G. Khlebopros, V. A. Okhonin, and A. I. Fet. *Catastrophes in Nature and Society: Mathematical Modeling of Complex Systems*. World Scientific Publishing, 2007.

[51] A. I. Khuri, editor. *Response Surface Methodology and Related Topics*. World Scientific Publishing Co., 2006.

[52] D. Krewski, Y. Wang, S. Bartlett, and K. Krishnan. Uncertainty, variability, and sensitivity analysis in physiological pharmacokinetic models. *Journal of Biopharmaceutical Statistics*, 5(3):245–271, Nov 1995.

[53] D. Kröner. *Numerical Schemes for Conservation Laws*. Wiley & Sons, Chichester, 1997.

[54] F. Lanzara, V. Maz'ya, and G. Schmidt. Approximate approximations from scattered data. *Journal of Approximation Theory*, 145:141–170, 2007.

[55] A. Lasota and M. C. Mackey. *Chaos, Fractals, and Noise*. Springer, 1994. Stochastic Aspects of Dynamics.

[56] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2004.

[57] G. R. Liu and Y. T. Gu. *An Introduction to Meshfree Methods and Their Programming*. Springer, 2005.

[58] C. Lubich. *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*. Zürich Lectures in Advanced Mathematics. European Mathematical Society, 2008.

[59] D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.

[60] M. C. Mackey, A. Longtin, and A. Lasota. Noise-induced global asumptotic stability. *Journal of Statistical Physics*, 60(5/6):735–751, 1990.

[61] B. F. J. Manly. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall, 1997.

[62] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*, 254:178–196, 2008.

[63] V. Maz'ya. Approximate approximations. In J. Whiteman, editor, *The mathematics of finite elements and applications*, pages 77–104,. Wiley, Chichester, 1993.

[64] V. Maz'ya. A new approximation method and its applications to the calculation of volume potentials. boundary point method. *3. DFG-Kollquium des DFG-Forschungsschwerpunktes "Randelementmethoden"*, 30. Sept. - 5. Oct. 1991.

[65] V. Maz'ya and G. Schmidt. Approximate wavelets and the approximation of pseudodifferential operators. *Applied and Computational Harmonic Analysis*, 6:287–313, 1999.

[66] V. Maz'ya and G. Schmidt. *Approximate Approximations*. American Mathematical Society, 2007.

[67] W. J. Morokoff and R. E. Caflisch. Quasi-monte carlo integration. *Journal of Computational Physics*, 122:218–230, 1995.

[68] F. Müller and W. Varnhorn. An approximation method using approximate approximations. *Applicable Analysis*, 85:669–680, 2006.

[69] E. A. Nadaraya. On estimating regression. *Theory of Probability and its Applications*, 9(1):141–142, 1964.

[70] H. Niederreiter. Quasi-monte carlo methods and pseudo-random numbers. *Bulletin of the American Mathematical Society*, 84(6):957–1041, 1978.

[71] H. Niederreiter and D. Talay, editors. *Monte Carlo and Quasi-Monte Carlo Methods 2004*. 6th International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing & 2nd International Conference on Monte Carlo and Probabilistic Methods for Partial Differential Equations, Springer, 2004.

[72] E. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076, 1962.

[73] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer Series in Computational Mathematics. Springer Verlag, Berlin, 1994.

[74] H. Rabitz, M. Kramer, and D. Dacol. Sensitivity analysis in chemical kinetics. *Annual Reviews of Physical Chemistry*, 34:419–461, 1983.

[75] K. Rektorys. *The Method of Discretization in Time and Partial Differential Equations*. Dordrecht-Boston-London, D. Reidel, 1982.

[76] K. Rektorys. Numerical and theoretical treating of evolution problems by the method of discretization in time. In M. Zlámal and J. Vosmanský, editors, *Equadiff 6, Proceedings of the International Conference on Differential Equations and Their Applications*, pages 71–84, 1985.

[77] C. P. Robert and G. Casella. *Monte Carlo statistical methods*. Springer, New York, 2004.

[78] R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo method / Reuven Y. Rubinstein, Dirk P. Kroese*. John Wiley & Sons, Hoboken, N.J. :, 2nd edition, 2008.

[79] J. H. Schaibly and K. E. Shuler. Study of sensitivity of coupled reaction systems to uncertainties in rate coefficients. II Applications. *Journal of Chemical Physics*, 59(8):3879–3888, 1973.

[80] W. E. Schiesser. *The Numerical Method of Lines: Integration of Partial Differential Equations*. Academic Press, San Diego, 1991.

[81] G. Schmidt. On approximate approximations and their applications. In *The Maz'ya Anniversary Collection*, volume 1. Operator Theory: Advances and Applications, 1999.

[82] D. R. Shier and K. T. Wallenius. *Applied Mathematical Modeling: A Multidisciplinary Approach*. Chapman & Hall, 1999.

[83] C. J. Tomlin and J. D. Axelrod. Biology by numbers: mathematical modelling in developmental biology. *Nature Reviews Genetics*, 8(5):331–340, May 2007.

[84] L. N. Trefethen. *Finite Difference and Spectral Methods for Ordinary and Partial Differential Equations*. Unpublished text, available at http://www.comlab.ox.ac.uk/nick.trefethen/pdetext.html, 1996.

[85] L. N. Trefethen. *Spectral Methods in Matlab*. Society for Industrial & Applied Mathematics, Philadelphia, 2000.

[86] T. Turányi. Sensitivity analysis of complex kinetic systems. tools and applications. *Journal of Mathematical Chemistry*, 5:203–248, 1990.

[87] H. van de Waterbeemd and E. Gifford. ADMET in silico modelling: towards prediction paradise? *Nature Reviews Drug Discovery*, 2(3):192–204, 03 2003.

[88] L. Wasserman. *All of Statistics: A Concise Course in Statistical Inference*. Springer, 2004.

[89] L. Wasserman. *All of Nonparametric Statistics*. Springer, 2007.

[90] G. S. Watson. Smooth regression analysis. *Sankhya: The Indian Journal of Statistics*, 26:359–372, 1964.

[91] A. Y. Weiße, I. Horenko, and W. Huisinga. Adaptive approach for modelling variability in pharmacokinetics. In M. Berthold, R. Glen, and I. Fischer, editors, *CompLife 2006*, pages 194–204. Springer, Berlin, 2006.

[92] C. Zenger. Sparse grids. In W. Hackbusch, editor, *Parallel Algorithms for Partial Differential Equations*, volume 31 of *Notes on Numerical Fluid Mechanics*, pages 241–251. Vieweg, Braunschweig/Wiesbaden, 1991.

Abbreviations & Notation

Abbreviations

FAST	Fourier amplitude sensitivity test
MC	Monte Carlo
ODE	Ordinary differential equation
PDE	Partial differential equation
TRAIL	Trapezoidal Rule for Adaptive Integration of Liouville dynamics

Notation

$\mathcal{A}$	Infinitesimal generator of the semigroup of Frobenius-Perron operators
$\mathcal{A}_\mathcal{K}$	Infinitesimal generator of the semigroup of Koopman operators
$\lvert \cdot \rvert$	Vector norm, unless stated otherwise
$\Delta u_t^{(k)}$	Difference between two solutions of order $k+1$ and k at time t
$\Delta \hat{u}_t^{(k)}$	Difference between two spatially perturbed solutions of order $k+1$ and k at time t
$\delta_\varepsilon^{(k)}$	Spatial perturbation of the temporal error estimate for the kth-order solution
$\delta_t^{(k)}$	Spatial perturbation of the kth-order solution at time t
$\varepsilon_t(\tau)$	True temporal error at time $t+\tau$
$\hat{\varepsilon}_t^{(k)}$	Spatially perturbed temporal error estimate of a solution of order k at time $t+\tau$
$\hat{u}_t^{(k)}$	Spatially perturbed solution of order k at time t
Id	Identity operator
$\mathcal{K}_t$	Koopman operator with respect to the evolution Φ_t
$\langle \cdot, \cdot \rangle$	Scalar product
$\mathcal{M}_{h,\mathcal{D}} u(x)$	Approximate approximation of $u(x)$
$\mathcal{P}_t$	Frobenius-Perron operator corresponding to the evolution Φ_t

$\mathscr{L}_k^{\gamma}(x)$	Generalized Laguerre polynomial
$\mu(B)$	Measure of set B
$\|\|\cdot\|\|_{\mathcal{L}_p(\Omega)}$	$\mathcal{L}_p$-norm restricted to the domain $\Omega \subset \mathbb{R}^d$
$\mathbb{P}[B]$	Probability of set B
Φ_t	Evolution operator
$\varepsilon_t^{(k)}(\tau)$	Temporal error estimate of a solution of order k at time $t+\tau$
$F(x)$	Vector field of ODE with extended state space
$L_k(x)$	Laguerre polynomial of order k
$r(z)$	Rational approximation to e^z
$R_\tau^{(k)}$	Rational approximation of order k to the semigroup of Frobenius-Perron operators
TOL	Local tolerance
TOL_t	Local temporal tolerance
TOL_x	Local spatial tolerance
$u_t^{(k)}$	Solution of order k at time t
$u_t = u(t,\cdot)$	Probability density function of the random state variable X_t
X_t	Random state variable at time t
$\|\|\cdot\|\|$	$\mathcal{L}_p$-norm, unless stated otherwise

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe.

Berlin, den 28. April 2009 Andrea Y. Weiße