Technische Universität München
Department of Mathematics

State-Discrete Probabilistic Methods for Partial Differential Equations

Dominic Christopher Kohler

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:		Univ.-Prof. Dr. Silke Rolles
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Johannes Müller
	2.	Hon.-Prof. Dr. Dr. h. c. Albert Gilg
	3.	Prof. Youssef Marzouk, Ph.D. Massachusetts Institute of Technology (USA) (nur schriftliche Beurteilung)

Die Dissertation wurde am 10.10.2013 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 09.01.2014 angenommen.

Abstract

We suggest a modeling and simulation framework, which meets two practical requirements. First, it respects the uncertainty of imprecise models and related data. Second, it follows users' demand for an efficient computation of coarse-grained information like risk levels instead of a slow evaluation of precise statements. The approach is designed to allow for easy integration of measurement data and to be scalable to large systems that arise in technical applications.

As a mathematical basis we introduce a translation of distinct partial differential equations into objects which are discrete in time, space and state space. The translation is achieved by state space discretization as in set oriented numerics and the use of the locality concept from cellular automata theory. Depending on how it is performed in detail, we can either use the resulting objects in the context of cellular automata theory to extract information about the partial differential equations' dynamics, or as a novel method for density-based uncertainty propagation and Bayesian inference. We also provide a consistency result. The methods are tested and validated in contaminant fate and fluid dynamical problems from water grid applications.

Zusammenfassung

Wir schlagen einen Modellierungs- und Simulationsansatz vor, der zwei Anforderungen aus der Praxis gerecht wird. Zum einen berücksichtigt er die Ungenauigkeit von Modellen und die oftmals unsichere Datenlage. Zum anderen nimmt er Nutzerwünsche nach effizienter Berechnung grobkörniger Informationen im Sinne von Risikoklassen anstelle von präziser und dafür langsamer Evaluierung ernst. Mit dem Ansatz können Messdaten leicht integriert werden. Außerdem ist er auf große Systeme skalierbar, wie sie in technischen Anwendungen vorkommen.

Als mathematische Grundlage führen wir eine Übersetzung einer Klasse von partiellen Differentialgleichungen in zeit-, orts- und zustandsdiskrete Objekte ein. In der Übersetzung wird der Zustandsraum wie in der mengenorientierten Numerik diskretisiert und das Lokalitätskonzept von zellulären Automaten ausgenutzt. Je nach Durchführung können wir dann entweder im Rahmen der Theorie der zellulären Automaten die Dynamik der partiellen Differentialgleichungen analysieren, oder die diskreten Objekte als neuartige Methode zur dichtebasierten Propagation von Unsicherheiten und für Bayes'sche Inferenz einsetzen. Wir zeigen auch die Konsistenz der Übersetzung. Die Verfahren werden an Wasserverschmutzungs- und Fluiddynamikproblemen in Wassernetzanwendungen getestet und validiert.

Acknowledgements

First of all, I thank my supervisor Prof. Dr. Johannes Müller very much for the friendly support throughout the last years. He guided my research with very interesting and unconventional ideas, personal advice and various challenges. I am especially grateful for having always been open for enlightning and helpful discussions, that's not to be taken for granted.

Also, I am very thankful to my co-supervisor Dr. Utz Wever for all the kind support and the numerous discussions. His motivation and vision helped me a lot when I was struggling with difficulties from time to time.

I owe many thanks to Prof. Youssef Marzouk, Ph.D., for having hosted me as a long term visitor in his group. The fruitful discussions with him and his group broadened my view on my work enormously and led to interesting insights.

I had the opportunity to work with various people at Siemens AG to whom I'm very thankful for their contributions. My special thanks go to Prof. Dr. Dr. h. c. Albert Gilg for his multiple constructive feedback and for having given me the possibility to work in this inspiring environment. Dr. Meinhard Paffrath, Birgit Obst, Dr. Maria Davidich, Dr. Wolfram Klein, Dr. Annelie Sohr, Roland Rosen, Dr. Yayun Zhou, Oliver Weeger, Dr. Ali Koc and Ludger Meyer helped me in bridging the gap between university research and industrial applications.

Furthermore I thank Prof. Dr. Oliver Junge and Dr. Péter Koltai for instructive discussions about set oriented numerics, and Prof. Dr. Gerd Steinebach for sharing his fluid dynamics experience. Both inputs contributed substantially to the progress of my work.

Special thanks go to Alexander Reiss for having worked with me on the bacterial regrowth model for his Bachelor's thesis.

It was a great pleasure to work with my colleagues and friends at university. In particular I'd like to thank Dr. Phaedra Stavropoulou, Dr. Sina Degenfeld-Schonburg, Dr. Florian Augustin, Dr. Laura Mainini, Dr. Luca Tosatto, Eric Dow, Patrik Conrad, Prof. Dr. Florian Rupp and Prof. Dr. Rupert Lasser.

The deepest thank of all goes to my family and friends for their unconditional support.

I gratefully acknowledge funding from Siemens AG, TU München and the German Academic Exchange Service (DAAD).

Contents

1 Introduction

1.1 Background and Problem Statement

Modeling and simulation is essential for many technologies especially in industrial environments. Traditional applications include the design of technical products, where simulation is often cheaper, safer and more flexible than real experiments. A digital or virtual twin of a technical system can also be used for testing new components in a virtual environment or for training of staff [136, 151]. Recent technological advances in sensor and information technology enable the collection of large amounts of data about features of technical systems. The integration of this additional information and the according adaptation of modeling and simulation technology leads to further applications like monitoring, decision support systems and the control of technical systems. For example, the evaluation of different response scenarios to incidents or virtual fault diagnosis become possible, which is especially important when systems are not easily accessible. The processing and analysis of data with respect to models of technical systems is also at the heart of emerging technologies like the Internet of Things [48, 117, 152] as a control network of sensors and actuators in physical objects or Big Data [25, 116] with a focus on huge data sets. In the context of production processes it is the basis for future Smart Factories [24, 154].

In all of these fields it is of vital importance to tailor the modeling and simulation techniques to the specific requirements. We focus on two common requirements: on the one hand, some model parameters or features of reality are often not known exactly but given as probability distributions instead. The resulting *uncertainty* has to be considered for three reasons. First, in non-linear systems time-evolution and taking the mean do not commute, i.e., the mean of the output distribution may differ from the deterministic evolution with the mean of the input distributions by orders of magnitude. So results from a deterministic simulation may not have any significance. Second, it is important to determine error bounds for computational results, and third, it is often an important issue to quantify failure probabilities of technical systems. However, uncertainty quantification is computationally very expensive and therefore impractical for large systems.

On the other hand, one is often not interested in the precise value of some system properties, but rather in more coarse-grained *risk levels* or *threshold values*. Users often only need such qualitative instead of exact quantitative statements. In principle these statements can be extracted from precise calculations, but then there is a huge numerical overhead. A faster alternative is to operate directly on discrete states that correspond to risk levels. Also from a conceptual perspective there are arguments for state-discrete computations. One can argue that computers are inherently discrete and can only approximate continuous values anyway. Furthermore, often there are difficulties with existence and uniqueness of solutions to continuous-valued models, a problem that does not appear in discrete setups.

The thesis aims at combining these minimal user requirements with the uncertainty present in the models. By calculating directly on a discrete state space with a risk level

interpretation we propose an efficient modeling and simulation framework, which is *scalable to large systems* and which also allows for convenient *integration of measurement data*. The theoretical foundations as well as the possibilities in practice are considered.

As an exemplary practical problem we are interested in the industrial application of water grid monitoring, in particular in contamination of *drinking* or *process water grids*. A special interest lies on large systems like municipal grids with accidental or intentional contamination for example by construction sites or terroristic threats. On the one hand contaminant source concentrations and locations or reaction schemes might not be known with certainty [40, 177, 187], and on the other hand consumers are often only interested in health risk level statements rather than in precise contaminant concentrations. Also here the interplay with sensor measurements plays a crucial role [39, 167]. The important field of water quality has been recognized as part of the UN Millenium Development Goals [169] and as one of the Grand Challenges for Engineering by the US National Academy of Sciences [171]. Especially the idea of adding computational intelligence to water grid operation has received a lot of interest lately e.g. from the US Environmental Protection Agency [149, 150, 170] and from industry [178].

1.2 State of the Art

In the following we review state of the art methods that can be applied to the challenges outlined above. They can be assigned to three different research fields: uncertainty propagation, state-discrete modeling and inverse problems for the integration of measurement data.

1.2.1 Uncertainty Propagation

The numerical treatment of differential equation dynamics that is subject to uncertain data has attracted a lot of interest lately. The goal is to determine stochastic output information like probability distributions or moments of random variables. We can characterize random influences on a system by the time-correlation of the stochastic parameters. A stochastic process like white noise will influence the system, if there are no or only small time correlations. If, on the contrary, the time-correlation is very high, the uncertainty will reflect the lack of knowledge about static system properties like material constants or reaction parameters.

A number of prominent numerical approaches are based on the *Monte Carlo* idea [89, 144], such as Latin hypercube sampling [114], the quasi Monte Carlo method [54], importance sampling [52] and the multilevel Monte Carlo method [8]. In these methods the probability distribution of interest is calculated from many deterministic simulations which are conducted for different realizations of the stochastic influences. Other well known approaches are based on the *Itô calculus* [100, 140]. It is furthermore possible to use *transfer operators* [45, 90, 106] or the *Fokker-Planck equation* [105] to describe the time evolution of a system's probability density directly. For stochastic influences with non-zero time-correlation *(generalized) polynomial chaos* expansions have been introduced [62, 179]. The idea is to expand a function of a random variable into a series of orthogonal polynomials and to determine the coefficients either by stochastic collocation [53, 186] or a Galerkin pro-

jection [62,93]. One of several improvements is to decompose the random space [176], and only recently numerical implementations of this improvement have been investigated [3].

The application of different methods for industrial purposes is discussed, for example, in [1,2,5]. Although the approaches have proven to be successful for many tasks, they often encounter certain efficiency restrictions in higher dimensions of the random space. Current research tries to meet this challenge by developing new methods like the alternating least-squares scheme [41] or by exploiting a structure of weakly interacting subsystems in the large problem [161].

1.2.2 State-Discrete Modeling

In classical descriptions of the physical world sytem properties are assigned values from a continuum of states. Many sophisticated models have been developed as systems of (partial) differential and algebraic equations, where also time and space are continuous. An example are advection-reaction-equations for contaminant dynamics in water pipes, which are coupled by hydraulic and chemical algebraic conservation laws at junctions. However, in most common schemes for numerical analysis time or space are discretized, e.g. in Runge-Kutta or multistep methods [37] for ordinary differential equations and in the method of lines [147], the finite difference [164], finite element [74] or finite volume methods [81,110] for partial differential equations (PDE).

On the contrary, there also exist many completely discrete models, in which especially the state space is finite and discrete. *Cellular Automata* (CA) as dynamical systems with discrete time, space and state space are the discrete equivalent of PDE [32,92,174]. Important properties are their shift-invariant spatial grid structure and their local interaction [27,79]. Stochastic versions have also been introduced [38, 50], which are similar to particle systems [44, 113]. As automata CA can be handled, simulated and, up to a certain degree, analyzed in a simpler way than PDEs [82]. Well-known applications range from Conway's Game of Life [58] over excitable media [66] or biological pattern formation [38] to fluid dynamics through lattice-gas cellular automata and lattice Boltzmann methods [184]. Also, agent models on the basis of CA are very popular with applications in city [11] or society [59] modeling and crowd control [183].

If we drop the requirement of shift-invariance for the grid structure, we will get more general objects that are studied as *Boolean networks* [94] in the framwork of complex networks [137]. In general, there is a huge interest in topological properties of networks, and in *adaptive networks* with dynamics of and on the network topology [65, 69]. An application is, for example, the analysis of power grids [127, 155]. Dynamics on networks with discrete states at each node and probabilistic dependencies between them are studied as *dynamic Bayesian networks*, a generalization of *hidden Markov models* [33, 131]. They are in turn special cases of *probabilistic graphical models* as a general graph-based framework to compactly represent probability distributions over high-dimensional spaces [18, 102]. In case of directed graphs the general models are called *Bayesian networks*, whereas they are referred to as *Markov networks* [10, 107] in the undirected case. The major uses of probabilistic graphical models are information representation, statistical inference from data for a given model and the learning of models from data. Probabilistic graphical models are, for example, applied in speech recognition [12] and cellular networks [56].

Because of potential analytical or computational advantages it is of interest to study

approximations of classical continuous- with discrete-valued models. Time-continuous dynamical systems on continuous state space can be studied by symbolic dynamics [181] or approximated by time-discrete Markov chains on finite state space [84]. This technique of state space discretization has led to the powerful tools of *set oriented numerics* [34, 35]. It is especially useful to study ergodic theory, asymptotic dynamics, and optimal control [71, 104]. Recently, contributions to uncertainty quantification have been made [90]. Alternative discretizations use information about special behavior of the continuous-state system [31] or have been discussed in the context of *interval arithmetic* [126] and for probabilistic graphical models with continuous states [102]. Approaches for the specific transition from PDE to CA include *ultradiscretization* [139, 166] and a probabilistic method [7]. People have even argued in favor of a reformulation of physical laws in a discrete language [165, 185].

1.2.3 Integration of Data: Inverse Problems

In a forward problem, observations are predicted from a set of model parameters. At least in the deterministic case, usually the forward problem is well-posed in the sense of Hadamard [75]: there exists a solution (existence), there is at most one solution (uniqueness), and the solution depends continuously on the observations (stability). The corresponding *inverse problem* consists of inferring parameters from the observations. By such an integration of measurement data the model can be refined or updated. An example is the identification of contamination sources in drinking water grids from sensor data. Observations are often noisy, uncertain or not sufficient with respect to the complexity of the model, and hence inverse problems often lack the Hadamard properties. We say that they are ill-posed.

Classically ill-posed problems are solved with *regularization* methods [98, 172]. Formulating inverse problems in a statistical inference framework allows for deeper insights in these methods and additional solution strategies [15, 49]. The *Bayesian setting* [61, 91, 125] has gained a lot of interest lately, with applications ranging from regular parameter estimation to optimal experimental design [85] or machine learning [18]. An advantage of the Bayesian approach is that uncertain forward models, noisy data and prior information are integrated quite naturally. Also, instead of a single value for the model parameters, Bayesian inference leads to the more informative posterior density over their values. Thus, uncertainties in the inferred results can be assessed quantitatively: all desired statistical information like median, mean or higher moments etc. can be derived from it.

In order to calculate these statistical properties we require the evaluation of weighted integrals of the posterior density over the model properties. Deterministic quadrature [47] is often not possible, but there are several ways to approximate these integrals. The simplest is to use Bayes' theorem and *Monte Carlo* sampling [89, 144] from the prior. *Markov chain Monte Carlo* (MCMC) methods sample directly from the posterior distribution [64, 78, 121]. In these methods a Markov process is constructed, whose stationary distribution is the posterior. Popular versions are the Metropolis-Hastings algorithm [28] and Gibbs sampling [61]. If the data come sequentially and one wishes for real-time inference, sequential Monte Carlo methods [42] or Kalman filters [67] can be applied.

The computational bottleneck for all these methods is that the potentially complex forward model in the likelihood has to be evaluated over and over again. Improvements

use surrogate forward models based on Polynomial Chaos and Karhunen-Loève expansions [118, 119]. In other approaches to Bayesian inference *surrogate models* are also constructed by model reduction based on Gaussian processes [97] and on order reduction [112, 159]. However, for expensive, e.g. PDE-based forward models, it is still a big challenge to make Bayesian inference computationally tractable. Current research therefore focuses on methodological advances of classical methods as well as on the development of new techniques. An example is the approach based on *optimal maps*, which directly push forward the prior to the posterior distribution [46].

For state-discrete probabilistic graphical models there is a collection of Bayesian inference tools especially adapted to their special structure. It contains many sophisticated exact and approximate inference algorithms. The approximate algorithms can be divided into deterministic and stoachstic ones. The central idea for the design of fast inference algorithms is to use dynamic programming [36]: repeated computations of intermediate results can be cached for efficient evaluation of conditional probability distributions in complex networks. Well-known examples are *variable elimination* or a more advanced formulation thereof, *message passing in clique* or *junction trees* [102]. Continuous-valued Gaussian counterparts of such ideas have also been used for uncertainty quantification in flows in random porous media by constructing a probabilistic graphical as a surrogate model [175]. In principle the basic algorithms are applicable to dynamic Bayesian networks when unrolling them. However, special algorithms can be designed because of their temporal structure. One example of deterministic approximate inference is a version of the forward-backward algorithm, the *BK algorithm* named after Boyen and Koller [20–22]. It approximates the potentially entangled belief state in each time step by a factorized one. In the inference literature this approach is also known as assumed density filtering [102]. A similar approach is the *factored frontier algorithm*, which updates the belief state locally and is based on the frontier algorithm for inference [135].

There exists a large pool of software for inference in state-discrete probabilistic graphical models. We mention just a few packages and refer to [133, 134] for a more comprehensive overview. Very popular toolkits are BNT [132] or its successor PMTK3 [43] by Kevin Murphy and co-workers. Another software which is especially designed for dynamic Bayesian networks is GMTK [16] by Jeff Bilmes. Examples for enterprise packages are PNL from Intel [88] or Infer.NET from Microsoft [123].

1.3 Research Goals and Outline

The goal of the thesis is to develop a *modeling and simulation framework* for spatio-temporal systems, that suites two practical demands. First, it respects the *uncertainty* of imprecise models and related data. Second, it takes user requirements serious, according to which efficient computation of coarse-grained information about discrete *risk levels* is preferable to slow evaluation of precise statements. The approach is designed to allow for easy *integration of measurement data* and to be *scalable to large systems* that arise in technical applications. In addition to theoretical investigations the framework is tested in contaminant fate and fluid dynamics problems in water grids.

The thesis is divided into four projects which work out the main idea step by step. They constitute the main chapters 3-6 of the work and can be read independently from each

other. An exception is Chap. 5 which is based on the material of Chap. 4.

After an introduction of definitions and basic concepts that are used throughout the whole thesis in Chap. 2, in the first project in Chap. 3 we introduce *cellular non-deterministic automata* (CNDA). They are understood in the spirit of non-deterministic automata theory and different from well known stochastic cellular automata. Their purpose is to analyze in a time-, space- and state-discrete setup the dynamics of a deterministic PDE that does not depend explicitly on the independent variables. A discretization is chosen such that it emphasizes spatial shift-invariance and locality of the interaction and that the resulting CNDA covers all possible dynamics of the PDE for given initial values. A detailed study of CNDA with the tools of CA theory reveals the difficulty of preserving spatial correlations in such a local construction; an insight that will be central to all following chapters. We also develop three approaches to efficiently approximate the CNDA with so-called superautomata at the cost of further losing - in a controlled way - information. As an application we discuss the Fisher-KPP equation [68] and the reaction random walk [80] which are central in the theory of reaction-diffusion processes [157].

Chap. 4 reconsiders the problem of approximating PDE dynamics with a completely discrete object. This time we use an approximation that neglects unlikely trajectories instead of finding a cover for all possible solutions. The basic idea is to calculate local transition probabilities between discretized portions of phase space to approximate the evolution of the system's probability density in transfer operator theory like in set-oriented numerics. PDEs can therefore be considered in a probabilistic setting, and the resulting *cellular probabilistic automata* (CPA) can be used for density-based uncertainty propagation. CPA are similar in structure to CNDA, and the insights gained with superautomata for CNDA are used extensively. We develop the method using initial value uncertainties under deterministic dynamics and show that it is consistent.

After the theoretical introduction of the CPA method we discuss *applications and extensions* thereof in Chap. 5. First, we conduct a complexity analysis and outline approaches for efficient implementations. Then we apply the method to arsenate transportation and adsorption and to bacterial regrowth as exemplary problems of contaminant fate in drinking water pipes, and also to the isothermel Euler equation of fluid dynamics. The applications demonstrate the interpretation of the phase space domains as discrete risk levels. We compare our results to Monte Carlo computations. By integrating algebraic coupling equations the CPA framework is also scaled to water grids. Furthermore we introduce and test an extension of CPA that imposes conservation properties of the underlying PDE system.

In Chap. 6 we *integrate measurement data* in the CPA framework. For this purpose CPA are interpreted as dynamic Bayesian networks, for which many inference algorithms and a large amout of software are available. CPA hence allow the inference problem in PDEs to be solved with inference algorithms from probabilistic graphical models. We discuss requirements for inference tools in our specific setup and choose the BK algorithm. To demonstrate our ideas, the scheme is applied to the arsenate adsorption problem in a water pipe: from measurements of the concentration of dissolved arsenate at the outflow boundary condition we infer the strength of the arsenate source at the inflow boundary condition. Finally we give our conclusions of the thesis in Chap. 7.

2 Preliminaries

The thesis is mostly formulated in the language of cellular automata and probability theory. We review according basic concepts and settle a notation for the following in this chapter. In Sec. 2.1 we present elements of cellular automata theory, and in Sec. 2.2 some applicable concepts from dynamical systems theory. Then we give a brief introduction to probability theory in Sec. 2.4. It relies on underlying ideas of measure theory which we present first in Sec. 2.3. Then we focus on the Frobenius-Perron operator in Sec. 2.5 and stochastic processes in Sec. 2.6 to describe stochastic dynamics.

2.1 Cellular Automata

In the present section we introduce basic concepts and notation from cellular automata theory and special topological definitions for Chap. 3 on the basis of [92].

Here $(G, \cdot)$ denotes a finitely generated group with generators $\{\tau_1, ..., \tau_n\}$, and E a finite set. G can be interpreted as the set of vertices of an associated graph, the Cayley graph, and is therefore in our context called a *grid* with *grid sites* $g \in G$. We write E^H for the set of all functions from $H \subseteq G$ to E and $\overline{e}_H$ for the function which is constantly $e \in E$ on H. When defining a $\varphi \in E^{\mathbb{Z}}$ explicitly we use the notational convention that the first written element of E after the dots always is the one at site $j = 0$. A *shift operator* is a mapping $\sigma_g : E^G \to E^G, \sigma_g(\varphi)(h) = \varphi(g \cdot h)$ for $g, h \in G$.

Definition 2.1.

i) A *(deterministic) cellular automaton (CA)* is a tuple (G, U, E, f_0) with G and E as above and $U = \{0, \tau_i \,|\, i = 1, ..., n\}$ the *neighborhood* of the unit element $0 \in G$. $f_0 : E^U \to E$ is a *local function* that induces the *global function* $f : E^G \to E^G, \varphi \mapsto f(\varphi)$ with

$$f(\varphi)(g) = f_0(\sigma_g(\varphi)|_U)$$

for $g \in G$. The *(deterministic) trajectory starting with* $\varphi_0 \in E^G$ is given by the sequence $(\varphi_n)_{n\in\mathbb{N}}$, where $\varphi_n = f(\varphi_{n-1})$ for $n \in \mathbb{N}^+$.

ii) If there is exactly one marked element $0 \in E$ with $f_0(\overline{0}_U) = 0$, the CA is said to have a *resting state*.

iii) For finite $H \subset G$, $h \in E^H$ is said to be a *Garden of Eden pattern*, if there are no states $\varphi, \psi \in E^G$ with $\varphi|_H = h$ and $\varphi = f(\psi)$ [115].

Understanding the restriction $\varphi|_{g\cdot U}$ as a function with domain U, it holds that $\sigma_g(\varphi)|_U = \varphi(g \cdot U) = \varphi|_{g\cdot U}$ and therefore $f_0(\sigma_g(\varphi)|_U) = f_0(\varphi|_{g\cdot U})$.

Although many definitions and results may be extended to more general groups, we restrict ourselves mostly to the group $(\mathbb{Z}, +)$ in the following. It will turn out that this is sufficient to apply our ideas to the simulation of partial differential equations in one

spatial dimension. We call $|U| = m + n + 1$ the *length* of a set $U = \{-m, ..., n\}$ such as the neigborhood, where $m, n \in \mathbb{N}$. In case of E carrying an order $\leq$, $\varphi \in E^{\mathbb{Z}}$ is called *monotonic* if $\varphi(i) \leq \varphi(j)$ for all $i \leq j$ or if $\varphi(i) \geq \varphi(j)$ for all $i \leq j$, where $i, j \in \mathbb{Z}$.

Definition 2.2. Let $\varphi, \psi \in E^{\mathbb{Z}}$. The *Cantor metric* d_C is given by $d_C(\varphi, \psi) = 0$ if $\varphi = \psi$ and $d_C(\varphi, \psi) = \frac{1}{1+i}$ otherwise, where i is the least non-negative integer such that either $\varphi(i) \neq \psi(i)$ or $\varphi(-i) \neq \psi(-i)$. The induced topology is called the *Cantor topology*.

The intuitive understanding of the Cantor metric is that two configurations are close if they agree on a large set around the origin. In general open balls with radius ϵ around $x \in X$ in a metric space (X, d) are written as $B_\epsilon(x) = \{y \in X \mid d(x, y) < \epsilon\}$, and the closure of $Y \subseteq X$ is denoted as $\overline{Y}$.

It is a well known fact that $(E^{\mathbb{Z}}, d_C)$ is a complete Cantor space [180] and that the Cantor metric leads to a strictly topological characterization of CA [79].

Definition 2.3. A *Cantor space* is a non-empty metric space $(\mathfrak{C}, d)$ that is

i) compact,
ii) *perfect*, i.e. it does not contain any isolated points: $c \in \overline{\mathfrak{C} \backslash \{c\}}$ for all $c \in \mathfrak{C}$,
iii) *totally disconnected*, i.e. for any $c \in \mathfrak{C}$ and any open $U \subseteq \mathfrak{C}$ with $c \in U$ there is a *clopen* (closed and open) set V with $c \in V \subseteq U$.

Theorem 2.4. (*Curtis, Hedlund, Lyndon*)
The global functions f of all CA $(\mathbb{Z}, U, E, f_0)$ are exactly the continuous functions on $(E^{\mathbb{Z}}, d_C)$ that are shift-invariant, i.e. that $\sigma_i f = f \sigma_i$ for all $i \in \mathbb{Z}$.

We will work with the Cantor metric in the following, although it is not shift-invariant. Besicovitch and Weyl topological spaces are alternative approaches that overcome this problem [14, 128].

Lastly we define the Hausdorff metric on the central subset

$$\mathcal{K} := \{S \in \mathcal{P}(E^{\mathbb{Z}}) \backslash \{\emptyset\} \mid S \text{ is compact}\}$$

of the power set $\mathcal{P}(E^{\mathbb{Z}})$ of $E^{\mathbb{Z}}$. In the following we write $T(j) := \{\varphi(j) | \varphi \in T\}$, $T|_V := \{\varphi|_V | \varphi \in T\}$ and $\mathcal{K}|_V := \{S|_V | S \in \mathcal{K}\}$ for $T \in \mathcal{K}$, $j \in \mathbb{Z}$ and $V \subseteq \mathbb{Z}$.

Definition 2.5. The *distance between an element $\varphi \in E^{\mathbb{Z}}$ and a set $S \in \mathcal{K}$* is defined by

$$d(\varphi, S) = \min_{\psi \in S} \{d_C(\varphi, \psi)\}.$$

The *Hausdorff distance* d_H between $S, T \in \mathcal{K}$ is given by

$$d_H(S, T) = \max \left\{ \max_{\varphi \in S} \{d(\varphi, T)\}, \max_{\psi \in T} \{d(\psi, S)\} \right\}.$$

Theorem 2.6.

i) $(\mathcal{K}, d_H)$ is a compact and complete metric space.

ii) Let $(S_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{K}$ such that there is $S \in \mathcal{K}$ with $\lim_{n\to\infty} S_n = S$ in the Hausdorff metric. Then

$$S = \{\varphi \in E^{\mathbb{Z}} \,|\, \forall n \in \mathbb{N}\, \exists \varphi_n \in S_n : \lim_{n\to\infty} \varphi_n = \varphi\}.$$

Proof We refer to [95], in which the Hausdorff metric is extended from $\mathcal{K}$ to $\mathcal{K} \cup \{\emptyset\}$ by the definitions

$$d_H(\emptyset, \emptyset) = 0 \quad \text{and} \quad d_H(S, T) = 1 \text{ if exactly one of } S, T \in \mathcal{K} \cup \{\emptyset\} \text{ is } \emptyset.$$

Claim i) $\mathcal{K} \cup \{\emptyset\}$ is a compact metric space, because $E^{\mathbb{Z}}$ is compact [95]. We conclude that $\mathcal{K}$ is compact because it is a closed subset of $\mathcal{K} \cup \{\emptyset\}$: $\mathcal{K}^C = (\mathcal{K} \cup \{\emptyset\})\backslash\mathcal{K} = \{\emptyset\}$ is open, as e.g. $B_{\frac{1}{2}}(\emptyset) = \{T \in \mathcal{K} \cup \{\emptyset\} | d_H(T, \emptyset) < \frac{1}{2}\} = \{\emptyset\}$. $\mathcal{K}$ is also complete as it is compact.

Claim ii) It is well known in the literature that convergence in the Hausdorff metric implies that the so-called *Kuratowski limit inferior* or *topological lower limit* coincides with the limit point; see e.g. [95].

□

2.2 Dynamical Systems

Here we present some ideas of dynamical systems theory on the basis of [72] and apply them to cellular automata. We set $\mathbb{N}_0 = \{0, 1, 2, ...\}$ and $\mathbb{R}_0^+ = [0, \infty)$.

Definition 2.7. A *time-discrete* resp. *time-continuous (metric) dynamical system* is a tuple (T, X, Φ) consisting of a *time* $T = \mathbb{N}_0$ resp. $T = \mathbb{R}_0^+$, a metric space X and a *flow* $\Phi : T \times X \mapsto X$ with the following properties:

i) $\Phi(t, x)$ is continuous in $x \in X$ and, if $T = \mathbb{R}_0^+$, also in $t \in T$
ii) $\Phi(0, x) = x$ for all $x \in X$
iii) $\Phi(t, \Phi(s, x)) = \Phi(t + s, x)$ for all $s, t \in T$ and all $x \in X$.

A set $Y \subset X$ is said to be *positively* resp. *negatively invariant* under the flow, if $\Phi(t, Y) \subseteq Y$ resp. $\Phi(t, Y) \supseteq Y$ for all $t \in T$. It is *invariant* if it is positively and negatively invariant.

We also write $\Phi(t, x) = \Phi_t(x)$. If $\Phi_t : X \to X$ is even a homeomorphism, the time may be extended to $\mathbb{Z}$ resp. $\mathbb{R}$. Systems of ordinary differential equations (ODE) induce time-continuous dynamical systems. Time-discrete dynamical systems are e.g. given by iteration of a continuous function $f : X \to X$ by $\Phi(k, x) = f^k(x)$, and hence we also write (X, f) for these systems. Every time-continuous dynamical system induces such a time-discrete one with function Φ_P for fixed $P > 0$.

The observation that global functions are continuous allows to interpret CA as a special instance of time-discrete dynamical systems on compact metric spaces. Hence we focus on the latter, especially on the concepts of long-term behavior and embeddings.

Definition 2.8. Let (X, f) be a time-discrete dynamical system on compact X.

i) The *ω-limit set* of $x \in X$ is given by $\omega_f(x) = \{y \in X \,|\, \exists$ strictly monotonically increasing sequence $(t_k)_{k\in\mathbb{N}}$ in $\mathbb{N} : y = \lim_{k\to\infty} f^{t_k}(x)\}$.

ii) If there is an open $U_0 \subseteq X$ such that $f(\overline{U_0}) \subseteq U_0$, $A = \cap_{k\geq 0} f^k(U_0)$ is called an *attractor* and U_0 a *preattractor*. $\mathcal{B}(A) = \cup_{n\geq 0} f^{-n}(U_0)$ is the *basin of attraction*. If $\mathcal{B}(A) = X$, A is called the *global attractor*.

By choosing $(t_k)_{k\in\mathbb{N}}$ in $\mathbb{R}_0^+$ with $t_k \to \infty$ instead of a strictly monotonically increasing sequence in $\mathbb{N}$, ω-limit sets are analogously defined for time-continuous dynamical systems. For all further common dynamical system concepts like stationary states, periodic orbits, (in)stability etc. we refer to [72].

Definition 2.9. A time-discrete dynamical system (X, f) can be *embedded* in another dynamical system (Y, h), if there is a continuous injective map $\alpha : X \to Y$ such that $\alpha \circ f = h \circ \alpha$ on X.

Accordingly a time-discrete dynamical system (X, f) can be embedded in a CA $(\mathbb{Z}, U, E, h_0)$ with global function h if it can be embedded in $(E^{\mathbb{Z}}, h)$. In the literature such an embedding is sometimes called *weak*, if additional features are required for a *strong* version [129].

2.3 Measure Theory

We prepare the introduction to probability theory in the next section with some basic concepts from measure theory in this section. Both sections are based on [57, 99].

Definition 2.10. Let X be a set and $\mathcal{P}(X)$ its power set. A *σ-algebra* is a family $\mathcal{A} \subseteq \mathcal{P}(X)$ with the properties

i) $X \in \mathcal{A}$
ii) $A \in \mathcal{A} \Rightarrow X \backslash A \in \mathcal{A}$
iii) $A_i \in \mathcal{A}, 1 \leq i \leq n \leq \infty \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$

A trivial $\sigma-$algebra for a set X is $\mathcal{P}(X)$ itself. One can also show that for $A \in \mathcal{P}(X)$ there is a smallest $\sigma-$algebra $\sigma(A)$ with $A \subset \sigma(A)$. If $(X, \mathcal{T})$ is a topological space, $\mathcal{B}(X) := \sigma(\mathcal{T})$ is called the *Borel σ-algebra*. In this work $\mathbb{R}^n$ is always considered to be equipped with the σ-algebra $\mathcal{B}(\mathbb{R}^n)$ if considered in a measure theoretic context.

The tuple $(X, \mathcal{A})$ is called a *measurable space*, and $A \in \mathcal{A}$ are the *measurable sets*. Two measurable spaces can be structurally related by the important class of measurable maps.

Definition 2.11. Let $(X, \mathcal{A})$ and $(X', \mathcal{A}')$ be measurable spaces. A map $V : X \to X'$ is called *$(\mathcal{A} - \mathcal{A}'$-)measurable*, if $V^{-1}(A') \in \mathcal{A}$ for any $A' \in \mathcal{A}'$.

Definition 2.12. Let $(X, \mathcal{A})$ be a measurable space. A map $\mu : \mathcal{A} \to [0, \infty]$ with $\mu(\emptyset) = 0$ is

i) *σ-additive*, if for every family of disjoint $A_i \in \mathcal{A}$ with $1 \leq i \leq n \leq \infty$

$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i),$$

ii) *σ-finite*, if there exists a sequence $(A_i)_{i\in\mathbb{N}}$ in $\mathcal{A}$ such that $X = \bigcup_{i\in\mathbb{N}} A_i$ and $\mu(A_i) < \infty$ for all $i \in \mathbb{N}$.

A σ-additive map $\mu : \mathcal{A} \to [0, \infty]$ on measurable space $(X, \mathcal{A})$ with $\mu(\emptyset) = 0$ is called a *measure*, and $(X, \mathcal{A}, \mu)$ a *measure space*. On any measure space an integral may be defined. For this construction we refer to [9, 99].

Measurable $g, g' : X \to \mathbb{R}$ on a measurable space $(X, \mathcal{A})$ with measure μ are said to be in the same equivalence class $[g]$, if they agree *μ-almost everywhere* (μ-a.e.), i.e. if there is $N \subset X$ with $\mu(N) = 0$ such that $g(x) = g'(x)$ for $x \in X \backslash N$. By this equivalence relation and with $p \in [0, \infty)$

$$\mathcal{L}^p(X, \mathcal{A}, \mu) := \{[g] \mid \int |g|^p d\mu < \infty\}$$

becomes a Banach space with norm

$$\|[g]\|_p := \left(\int |g|^p d\mu \right)^{1/p}.$$

Having this in mind we write g for the elements $[g] \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ in the following.

For a measurable map $g : X \to [0, \infty)$ on measure space $(X, \mathcal{A}, \mu)$

$$\nu : \mathcal{A} \to [0, \infty], \quad A \mapsto \int g \chi_A d\mu \tag{2.1}$$

is a measure. The following important theorem reverses this statement.

Theorem 2.13. *(Radon-Nikodym)*

Let μ and ν be σ-finite measures on measurable space $(X, \mathcal{A})$ such that $\nu(A) = 0$ for all $A \in \mathcal{A}$ with $\mu(A) = 0$. Then there exists an integrable function $g : X \to [0, \infty)$ such that

$$\nu(A) = \int_A g d\mu \quad \text{for all } A \in \mathcal{A}.$$

We end this section with 3 important examples of measure spaces with σ-finite measures.

i) For any measurable space $(X, \mathcal{A})$ the *Dirac measure* in a point $x \in X$ is defined by

$$\delta_x : \mathcal{A} \to [0, \infty], \quad A \mapsto \chi_A(x).$$

ii) Consider a measurable space $(E, \mathcal{P}(E))$, where $E \neq \emptyset$ is at most countable. Then

$$\gamma : \mathcal{P}(E) \to [0, \infty], \quad A \mapsto \sum_{e \in E} \delta_e(A) = |A|$$

is the *counting measure*. Here $|A|$ denotes the cardinality of A.

iii) The measurable space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ can be equipped with the *Lebesgue-Borel measure* λ. It is the unique measure with the property

$$\lambda((a, b]) = \prod_{i=1}^{n} b_i - a_i$$

for all $a, b \in \mathbb{R}^n$ with $a_i < b_i, 1 \leq i \leq n$. We abbreviate $\mathcal{L}^p(\mathbb{R}^n) := \mathcal{L}^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$.

2.4 Probability Theory

The whole probability theory is based on a special case of measure spaces: A *probability space* is a measure space $(X, \mathcal{A}, \mu)$ with $\mu(X) = 1$. Here μ is called a *probability measure* and is σ-finite per definition. In a probability space $A \in \mathcal{A}$ is called an *event*. Probability spaces model random experiments, of which usually just certain aspects are observed in the form of random variables.

Definition 2.14. Let $(X, \mathcal{A}, \mu)$ be a probability space and $(X', \mathcal{A}')$ a measurable space. A *random variable* is a measurable mapping $V : X \to X'$. The *image measure*

$$\mu_V : \mathcal{A}' \to [0, 1], \quad A' \mapsto \mu(V^{-1}(A'))$$

is called the *distribution* of V, shortly $V \sim \mu_V$. If $(X', \mathcal{A}') = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ or $(X', \mathcal{A}') = (E, \mathcal{P}(E))$ for finite E, we speak of *real* or *discrete random variables*, respectively.

Note that μ_V is a probability measure as well. We also write $\mu(V \in A') = \mu_V(A')$ for $A' \in \mathcal{A}'$ and $\mu(V = v) = \mu(\{x \in X | V(x) \in \{v\}\})$ for $v \in X'$. Often there is $g \in \mathcal{L}^1(X', \mathcal{A}', \mu')$ such that the distribution μ_V of a random variable V may be written as

$$\mu_V(A') = \int_{A'} g d\mu', \quad \text{for all } A' \in \mathcal{A}'.$$

Then g is called the *density* of μ_V with respect to μ. We also say that V has density g. This motivates the study of the set of densities on a general measure space $(X, \mathcal{A}, \mu)$:

$$D(X, \mathcal{A}, \mu) := \{g \in \mathcal{L}^1(X, \mathcal{A}, \mu) \,|\, g \geq 0, \|g\|_1 = 1\}.$$

According to Eq. 2.1 every density defines a probability measure on the measure space $(X, \mathcal{A}, \mu)$. We give 2 examples.

i) Consider the measure space $(E, \mathcal{P}(E), \gamma)$, where $E \neq \emptyset$ is at most countable. Let further $(p_e)_{e \in E}$ be nonnegative numbers with $\sum_{e \in E} p_e = 1$. Then the *weight function*

$$g : E \to [0, 1], \quad g(e) = p_e$$

is in $D(E, \mathcal{P}(E), \gamma)$ and defines the probability measure $\mu : \mathcal{P}(E) \to [0, 1]$ by

$$\mu(A) = \int g \chi_A d\gamma = \sum_{e \in E} (g\chi_A)(e) = \sum_{e \in A} g(e) = \sum_{e \in A} p_e.$$

for $A \in \mathcal{P}(E)$. Obviously $\mu = \sum_{e \in E} p_e \delta_e$.

ii) Let $g \in D(\mathbb{R}^n) := D(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$ be given by

$$g(v) = \frac{1}{\lambda(\Omega)} \chi_\Omega(v)$$

for $\Omega \subset \mathbb{R}^n$, where χ_Ω is its characteristic function. A real random variable with the associated distribution is said to have *uniform distribution*.

To characterize the relationship of several events at a time, we need the concept of conditional probabilities and independent events.

Definition 2.15. Let $(X, \mathcal{A}, \mu)$ be a probability space and $n \in \mathbb{N}$.

i) For $B_1, ..., B_n \in \mathcal{A}$ with $\mu(B_1 \cap ... \cap B_n) > 0$, the *conditional probability distribution* of $A \in \mathcal{A}$ is given by

$$\mu(A|B_1, ..., B_n) = \frac{\mu(A \cap B_1 \cap ... \cap B_n)}{\mu(B_1 \cap ... \cap B_n)}.$$

ii) $A_1, ..., A_n \in \mathcal{A}$ are *independent* if for all $m \in \mathbb{N}$, $0 \leq m \leq n$, $\mu(A_{i_1} \cap ... \cap A_{i_m}) = \mu(A_{i_1})...\mu(A_{i_m})$.

It is easy to verify that $\mu(A|B_1, ..., B_n)$ is a probability distribution over the measurable space $(X, \mathcal{A})$. Now we transfer the ideas to random variables.

Definition 2.16. Let $(X, \mathcal{A}, \mu)$ be a probability space and $V_i : X \to X_i'$ random variables with values in the measurable spaces $(X_i', \mathcal{A}_i')$, where $i \in I = \{1, ..., n\}$ for $n \in \mathbb{N}$.

i) Consider $j \in I$. The *conditional distribution of V_j given $V_i \in A_i'$* for $A_i' \in \mathcal{A}_i'$, where $i \in \tilde{I} = I \setminus \{j\}$, is

$$\mu_{V_j|V_i, i\in\tilde{I}} : \mathcal{A}_j' \to [0, 1],\ \mu_{V_j|V_i, i\in\tilde{I}}(A_j'|\forall i \in \tilde{I} : V_i \in A_i') = \mu(V_j^{-1}(A_j')|\forall i \in \tilde{I} : V_i^{-1}(A_i')).$$

ii) $\{V_i\}_{i\in I}$ are *independent*, if $V_i^{-1}(A_i')$ are independent with respect to μ for all $A_i' \in \mathcal{A}_i'$. If in addition $\mu_{V_i} = \mu_{V_j}$ for all $i, j \in I$, we say that $\{V_i\}_{i\in I}$ are *independent and identically distributed (iid)*.

We also write $\mu(V_j \in A_j'|V_i \in A_i') = \mu_{V_j|V_i, i\in\tilde{I}}(A_j'|\forall i \in \tilde{I} : V_i \in A_i')$ and $\mu(V_j = v_j|V_i = v_i) = \mu_{V_j|V_i, i\in\tilde{I}}(\{v_j\}|\forall i \in \tilde{I} : V_i \in \{v_i\})$, where $v_i \in X_i'$ and $v_j \in X_j'$.

In practice one is often interested in statistical quantities of a real-valued random variable such as the expected values or variances.

Definition 2.17. Let $(X, \mathcal{A}, \mu)$ be a probability space. The *m-th moments* of a real-valued random variable $V : X \to \mathbb{R}^n$ are defined as

$$E[V_i^m] = \int_X V_i^m d\mu$$

for $i = 1, ..., n$. The first moments are called *expected values* or *means*, and $\text{Var}[V_i] = E[V_i^2] - (E[V_i])^2$ are the *variances*.

2.5 The Frobenius-Perron Operator

Now we focus on stochastic dynamics. In the present section we briefly introduce the Frobenius-Perron operator, a special transfer operator that describes how probability densities evolve under phase space dynamics. It has gained much attention e.g. in the investigation of chaotic systems [106] or in Ulam's method for the approximation of invariant measures in ergodicity theory [168].

Consider the measure space $(\mathbb{R}^{mn}, \mathcal{B}(\mathbb{R}^{mn}), \lambda)$, where $m, n \in \mathbb{N}$, $\mathcal{B}(\mathbb{R}^{mn})$ is the Borel σ-algebra and λ the Lebesgue measure. A measurable map $S : \mathbb{R}^{mn} \to \mathbb{R}^{mn}$ is called nonsingular if $\lambda(S^{-1}(A')) = 0$ for all $A' \in \mathcal{B}(\mathbb{R}^{mn})$ with $\lambda(A') = 0$. For any such map a unique operator can be defined on the basis of the Radon-Nikodym theorem [106].

Definition 2.18. Given a nonsingular map $S : \mathbb{R}^{mn} \to \mathbb{R}^{mn}$, for $g \in \mathcal{L}^1(\mathbb{R}^{mn})$ the *Frobenius-Perron operator* (FPO) $P_S : \mathcal{L}^1(\mathbb{R}^{mn}) \to \mathcal{L}^1(\mathbb{R}^{mn})$ is defined by

$$\int_{A'} P_S g(x)dx = \int_{S^{-1}(A')} g(x)dx \quad \forall A' \in \mathcal{B}(\mathbb{R}^{mn}).$$

The FPO preserves positivity and normalization and hence describes how densities are mapped under phase space evolution with S.

2.6 Stochastic Processes

Another important concept to describe stochastic dynamics are stochastic processes. We base our presentation on [99]. From now on we focus on random variables with values in measure space $(E, \mathcal{P}(E))$ with finite $E \neq \emptyset$, although the theory can be developed in more generality.

Definition 2.19. A (time- and state-discrete) *stochastic process* is a family of random variables $V = (V^t)_{t\in\mathbb{N}_0}$ on a probability space $(X, \mathcal{A}, \mu)$ with values in the measurable space $(E, \mathcal{P}(E))$. The elements of E are called *states*.

Definition 2.20. The stochastic process $(V^t)_{t\in\mathbb{N}_0}$ is called *Markovian*, if for all $t \in \mathbb{N}$ and all $v^0, ..., v^t \in E$

$$\mu(V^t = v^t | V^{t-1} = v^{t-1}, ..., V^0 = v^0) = \mu(V^t = v^t | V^{t-1} = v^{t-1}),$$

i.e. if the present state at time t only depends on the immediate predecessor at time $t-1$. If the Markov property does not depend on the index t, the process is called *stationary* or *homogeneous*.

For a stationary Markovian process it is possible to formally define a time-independent transition model.

Definition 2.21. Let $(E_1, \mathcal{P}(E_1))$ and $(E_2, \mathcal{P}(E_2))$ be measurable spaces for finite $E_1 \neq \emptyset, E_2 \neq \emptyset$. A map $\kappa : E_1 \times \mathcal{P}(E_2) \to [0, 1]$ is called a *transition distribution* or *Markov kernel* if

i) the function $e_1 \mapsto \kappa(e_1, \tilde{E}_2)$ is measurable for all $\tilde{E}_2 \in \mathcal{P}(E_2)$ and
ii) the function $\tilde{E}_2 \mapsto \kappa(e_1, \tilde{E}_2)$ is a probability distribution on $(E_2, \mathcal{P}(E_2))$ for all $e_1 \in E_1$.

We can describe the time-evolution of a stationary Markovian stochastic process just by an initial and a transition distribution. For random variables V^t, V^{t+1} with values in $(E, \mathcal{P}(E))$ for finite $E \neq \emptyset$, the transition distribution is equivalent to a conditional distribution $\mu_{V^{t+1}|V^t}$, or, even simpler, to a matrix.

3 Cellular Non-Deterministic Automata

In this chapter we introduce a non-probabilistic approach to study the dynamics of distinct partial differential equations (PDEs) in a completely discrete setup. The main idea is to translate them into objects that we call cellular non-deterministic automata (CNDA) and to analyze the latter. The translation is achieved by state space discretization as in set oriented numerics and the use of the locality concept from cellular automata (CA) theory.

In particular we consider a PDE which does not depend explicitly on the independent variables. We dismiss information stepwise; see Fig. 3.1a. First, we use a method of lines to go from a PDE on $\mathbb{R}$ to a countable system of ordinary differential equations (ODE); we replace $\mathbb{R}$ by $\mathbb{Z}$ by space discretization. Next, we discretize the coupled ODE in time and state by a variant of the set oriented method for dynamical systems. The outcome is interpreted as a CNDA in the spirit of non-deterministic automata theory [82]: the transition of a state at one site is not deterministically determined by a configuration in the neighborhood, we only know a set of possible image states. This is a consequence of the information reduction through state discretization. CNDA work on the power set of the completely discrete states to follow all possible trajectories at once.

We suggest to analyze CNDA with supersystems (SS), especially by embedding them in or covering them with deterministic cellular superautomata (SA); see Fig. 3.1b. This allows to approximate their dynamical behavior with the standard theory of CA at the cost of losing - in a controlled way - further information.

It depends very much on the system and the details of this procedure whether the SS reveals the essential features of the initial PDE. The method is chosen in such a way that no features of the approximating space and time discretized system of ODE are lost. However, if the information loss is too high, the outcome may be trivial in the sense that it does not rule out most dynamical patterns but accepts almost all patterns as possible structures.

The chapter is structured as follows: in Sec. 3.1 we abstractly introduce CNDA and discuss them in the CA context. Sec. 3.2 is concerned with the general analysis of CNDA with SS and SA. Then we show in detail how a CNDA may be constructed from a PDE as sketched above in Sec. 3.3. We use the Fisher-KPP equation [68] and the reaction random walk [80], which are both popular in the context of reaction-diffusion processes [157], as prototypical examples in order to obtain some practical insight into the developed theory. Finally we conclude our results in Sec. 3.4.

3.1 Introduction of Cellular Non-Deterministic Automata

In this section we first define cellular non-deterministic automata (CNDA) and compare them to conventional concepts in the cellular automata (CA) context. Then we investigate how our construction can be topologically characterized by a Curtis-Hedlund-Lyndon theorem.

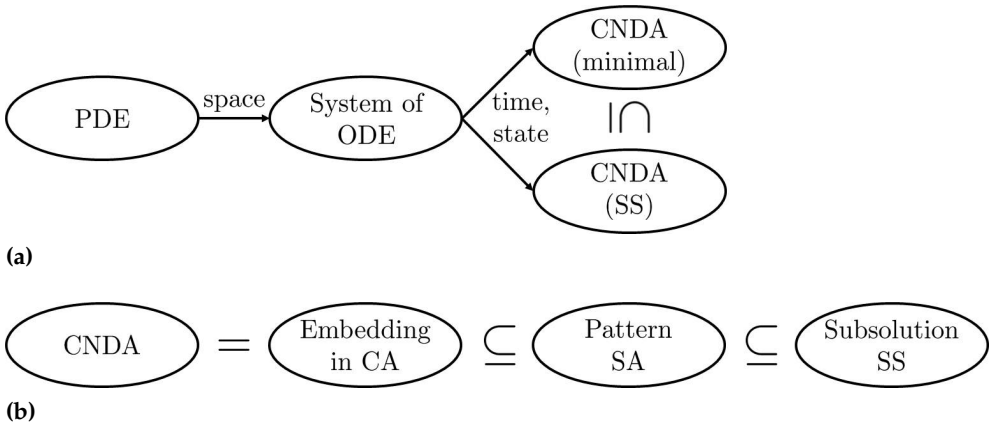


Figure 3.1: (a) Space, time and state are discretized to describe the dynamics of a PDE by the dynamics of a CNDA. In many cases it is hard to determine the smallest possible CNDA and one ends up with a bigger CNDA which already carries less information. (b) A CNDA is covered by its SS of decreasing precision; see Sec. 3.2 for details.

3.1.1 Definition of Cellular Non-Deterministic Automata

We only work on grid $\mathbb{Z}$, although certain generalizations are immediately possible. It will turn out that this is enough to account for analysis of PDE dynamics in one spatial dimension. CNDA are defined on the set $\mathcal{K}$ of all compact subsets of $E^{\mathbb{Z}}$, where we exclude the empty set for technical reasons. It can be endowed with the Hausdorff metric d_H to form a compact and complete metric space; see Sec. 2.1.

Definition 3.1. A *cellular non-deterministic automaton (CNDA)* is a tuple $(\mathbb{Z}, U, E, f_0)$ with *grid* $\mathbb{Z}$ and *neighborhood* $U = \{-m, ..., n\}$ of the origin $0 \in \mathbb{Z}$, where $m, n \in \mathbb{N}$. E is a finite set and $f_0 : E^U \to \mathcal{P}(E)\backslash\{\emptyset\}$ a *local function* that induces the *elementary global function*

$$f_1 : E^{\mathbb{Z}} \to \mathcal{K}, \qquad \varphi \mapsto f_1(\varphi) = \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \psi(i) \in f_0(\sigma_i(\varphi)|_U)\}$$

and the *global function*

$$f : \mathcal{K} \to \mathcal{K}, \qquad S \mapsto f(S) = \cup_{\varphi \in S} f_1(\varphi).$$

The *trajectory starting with* $S_0 \in \mathcal{K}$ is given by the sequence $(S_n)_{n\in\mathbb{N}}$, where $S_n = f(S_{n-1})$ for $n = 1, 2, ...$

The elements of E are called *elementary local*, the ones of $\mathcal{P}(E)$ *local*, the ones of $E^{\mathbb{Z}}$ *deterministic* and the ones of $\mathcal{K}$ *random states*. $\{e\} \in \mathcal{P}(E)$ for $e \in E$ are also referred to as elementary local states, and $\{\varphi\} \in \mathcal{K}$ for $\varphi \in E^{\mathbb{Z}}$ also as deterministic states.

By construction it is clear that a CNDA works on $\mathcal{P}(E^{\mathbb{Z}})\backslash\{\emptyset\}$, but not that it is possible to restrict the definition to the compact subsets of $E^{\mathbb{Z}}$. In the following we therefore check that a trajectory stays in $\mathcal{K}$ under evolution with respect to a CNDA. To show that a subset of $E^{\mathbb{Z}}$ is compact with respect to the Cantor metric it is always enough to prove that it is closed, as $E^{\mathbb{Z}}$ is compact.

Proposition 3.2. $f : \mathcal{P}(E^{\mathbb{Z}})\backslash\{\emptyset\} \to \mathcal{P}(E^{\mathbb{Z}})\backslash\{\emptyset\}$ is a closed map.

Proof We prove that $f(S)$ is closed, if $S \in \mathcal{P}(E^{\mathbb{Z}})\backslash\{\emptyset\}$ is. Let $(\psi_n)_{n\in\mathbb{N}}$ be a sequence in $f(S)$ s.t. $\psi_n \to \psi \in E^{\mathbb{Z}}$ as $n \to \infty$. Since $E^{\mathbb{Z}}$ is a metric space it is enough to show that $\psi \in f(S)$.

By using a suited subsequence if necessary we assume that $\psi|_{[-n,..,n]} = \psi_n|_{[-n,..,n]}$. For any $n \in \mathbb{N}$ there is $\varphi_n \in S$ s.t. $\psi_n \in f(\{\varphi_n\})$. As S is compact there is a converging subsequence of the φ_n,

$$\varphi_{n_k} \to \varphi \in S.$$

We claim that $\psi \in f(\{\varphi\})$. This is the case, if we have $\psi(j) \in f_0(\varphi|_{j+U})$ for each $j \in \mathbb{Z}$. Let us fix j. As $\varphi_{n_k} \to \varphi$ there is $k \in \mathbb{N}$ large enough such that $|j| < n_k$ and $\varphi_{n_k}|_{j+U} = \varphi|_{j+U}$. Thus, $f_0(\varphi_{n_k}|_{j+U)}) = f_0(\varphi|_{j+U})$. As $\psi(j) = \psi_{n_k}(j) \in f_0(\varphi_{n_k}|_{j+U})$, we find $\psi(j) \in f_0(\varphi|_{j+U})$.

So $\psi \in f(\{\varphi\}) \subseteq f(S)$, and f is a closed map. □

Note that especially the interesting deterministic states are in $\mathcal{K}$.

Lemma 3.3. Let $\varphi_1, ..., \varphi_n \in E^{\mathbb{Z}}$ for $n \in \mathbb{N}$. Then $\{\varphi_1, ..., \varphi_n\}$ is closed.

Proof Let $\varphi \in E^{\mathbb{Z}}$. It is enough to show that $\{\varphi\}^C = E^{\mathbb{Z}}\backslash\{\varphi\}$ is open, as finite unions of closed sets are closed.

$\psi \in \{\varphi\}^C$ implies that there exists $i \in \mathbb{Z}$ such that $\varphi(i) \neq \psi(i)$. We choose $0 < \epsilon < \frac{1}{1+|i|}$ and find that

$$B_\epsilon(\psi) = \{\chi \in E^{\mathbb{Z}} | d_C(\psi, \chi) < \epsilon\} = \{\chi \in E^{\mathbb{Z}} | \psi|_{\mathbb{Z}_{\frac{1}{\epsilon}-1}} = \chi|_{\mathbb{Z}_{\frac{1}{\epsilon}-1}}\}.$$

Therefore $\chi(i) = \psi(i) \neq \varphi(i)$ and $\chi \notin \{\varphi\}$ for $\chi \in B_\epsilon(\psi)$. So $B_\epsilon(\psi) \subseteq \{\varphi\}^C$. □

CA are a special case of CNDA: if the range of f_0 only contains elementary local states, the range of f_1 will just contain deterministic states, and f will coincide with f_1 on deterministic states. The trajectories under a general CNDA can be thought of as comprising many deterministic trajectories, i.e., trajectories of deterministic states.

We use the term 'non-deterministic' in the sense of automata theory [82]. CNDA should not be confused with probabilistic or stochastic CA [38, 50], which are sometimes also called non-deterministic in the CA literature. But unlike in the concept at hand there one of a set of deterministic local functions is chosen in each time step and at each site by means of probabilities. The trajectory does not follow all possible deterministic trajectories at once.

3.1.2 Cellular Non-Deterministic Automata and the CHL Theorem

In the present section we show that the global function of a CNDA acts continuously and shift-invariant on the compact metric space $\mathcal{K}$. This means that CNDA can be regarded as time-discrete dynamical systems. However, due to our special construction there are continuous and shift-invariant functions on $\mathcal{K}$ that are not the global function of a CNDA. We conclude that the Curtis-Hedlund-Lyndon theorem, Thm. 2.4, cannot be extended to CNDA. We end with a proposition on how to generalize CNDA to automata on $\mathcal{K}$ such that a Curtis-Hedlund-Lyndon theorem holds.

Proposition 3.4. The global function f of a CNDA $(\mathbb{Z}, U, E, f_0)$ is continuous.

Proof Let $S \in \mathcal{K}$. We show that for each $\epsilon > 0$ there is a $\delta > 0$ such that $d_H(f(S), f(T)) < \epsilon$ if $d_H(S,T) < \delta$ for $T \in \mathcal{K}$.

Let $\epsilon > 0$ and choose $k \in \mathbb{N}$ such that $\frac{1}{1+k} < \epsilon$. Let furthermore $m \in \mathbb{N}$ such that $U \subseteq [-m, ..., m]$, and choose $\delta > 0$ such that $\delta < \frac{1}{1+k+m}$. For $T \in \mathcal{K}$ with $d_H(S,T) < \delta$ we find that $\max_{\varphi \in S}\{\min_{\psi \in T}\{d_C(\varphi, \psi)\}\} < \delta$. So for all $\varphi \in S$ there exists a $\psi \in T$ such that $d_C(\varphi, \psi) < \delta$. Thus

$$\begin{aligned} &\forall i \in [-m-k, ..., m+k]: && \varphi(i) = \psi(i) \\ \Rightarrow &\forall i \in [-k, ..., k]: && \varphi|_{i+U} = \psi|_{i+U} \\ \Rightarrow &\forall i \in [-k, ..., k]: && f_0(\varphi|_{i+U}) = f_0(\psi|_{i+U}). \end{aligned}$$

As furthermore for every $\chi \in f(S)$ there is $\varphi \in S$ with $\chi \in f(\{\varphi\})$, for every $\chi \in f(S)$ there is $\omega \in f(\{\psi\}) \subseteq f(T)$ with $\chi(i) = \omega(i)$ for all $i \in \mathbb{Z}|_k$. So $\min_{\omega \in f(T)}\{d_C(\chi, \omega)\} < \epsilon$ for all $\chi \in f(S)$, which means that $\max_{\chi \in f(S)}\{\min_{\omega \in f(T)}\{d_C(\chi, \omega)\}\} < \epsilon$.

Analogously we find that $\max_{\omega \in f(T)}\left\{\min_{\chi \in f(S)}\{d_C(\omega, \chi)\}\right\} < \epsilon$ and conclude in summary that $d_H(f(S), f(T)) < \epsilon$. □

To talk about shift-invariance of a CNDA we need to lift the shift-operator from $E^{\mathbb{Z}}$ to $\mathcal{K}$. We define $\tilde{\sigma}_i : \mathcal{K} \to \mathcal{K}, S \mapsto \tilde{\sigma}_i(S) = \{\sigma_i(\varphi) | \varphi \in S\}$ for $i \in \mathbb{Z}$ and also write σ_i for $\tilde{\sigma}_i$ in the following. Since such a shift-operator can be regarded as the global function of a CNDA, in which the elementary global function is given by the shift-operator on $E^{\mathbb{Z}}$, Prop. 3.2 justifies that it maps to compact sets.

Proposition 3.5. The global function f of a CNDA $(\mathbb{Z}, U, E, f_0)$ is shift-invariant, i.e., $\forall i \in \mathbb{Z} : \sigma_i f = f \sigma_i$.

Proof Let $S \in \mathcal{K}$ and $i \in \mathbb{Z}$. Then

$$\begin{aligned} \sigma_i f(S) &= \{\sigma_i(\psi) \in E^{\mathbb{Z}} \,|\, \psi \in f(S)\} \\ &= \{\sigma_i(\psi) \in E^{\mathbb{Z}} \,|\, \exists \varphi \in S \, \forall j \in \mathbb{Z} : \, \psi(j) \in f_0(\sigma_j(\varphi)|_U)\} \\ &= \{\sigma_i(\psi) \in E^{\mathbb{Z}} \,|\, \exists \varphi \in S \, \forall j \in \mathbb{Z} : \, \psi(i+j) \in f_0(\sigma_{i+j}(\varphi)|_U)\} \\ &= \{\sigma_i(\psi) \in E^{\mathbb{Z}} \,|\, \exists \varphi \in S \, \forall j \in \mathbb{Z} : \, \sigma_i(\psi)(j) \in f_0(\sigma_j(\sigma_i(\varphi))|_U)\} \\ &= \cup_{\varphi \in S} f_1(\sigma_i(\varphi)) = f(\{\sigma_i(\varphi) \,|\, \varphi \in S\}) = f\sigma_i(S). \end{aligned}$$

□

Next we show that the Curtis-Hedlund-Lyndon theorem does not hold for CNDA. Heuristically the reason is that continuous and shift-invariant functions do not need to preserve spatial correlations. In general they allow, for example, for deterministic states in the image that are glued together by elementary local image states from different deterministic states in the preimage.

Theorem 3.6. If $|E| > 1$, there are continuous and shift-invariant functions $g : \mathcal{K} \to \mathcal{K}$ that cannot be interpreted as the global function of a CNDA.

Proof We give an example of a continuous and shift-invariant function on $\mathcal{K}$ and show that there is no local function inducing it as a global function of a CNDA. Consider

$$g : \mathcal{K} \to \mathcal{K}, \quad S \mapsto g(S) = \{\varphi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} \, \exists \psi \in S : \varphi(i) = \psi(i)\}.$$

Claim i) g is closed and hence well-defined.

As in Prop. 3.2 it is enough to show that the limit point of a converging sequence in $g(S)$ is in $g(S)$ as well.

Let $(\psi_n)_{n\in\mathbb{N}}$ be a sequence in $g(S)$ with $\psi_n \to \psi$ as $n \to \infty$. By going to a suited subsequence if necessary we may assume that there is $\psi_{|i|} \in g(S)$ such that $\psi(i) = \psi_{|i|}(i)$ for any $i \in \mathbb{Z}$. Per construction of g then there is also $\varphi_{|i|} \in S$ such that $\varphi_{|i|}(i) = \psi_{|i|}(i)$. Hence $\psi(i) = \varphi_{|i|}(i)$ and so $\psi \in g(S)$.

Claim ii) g is continuous.

We show that for all $\epsilon > 0$ there is $\delta > 0$ such that $d_H(S,T) < \delta$ implies $d_H(g(S), g(T)) < \epsilon$ for all $S, T \in \mathcal{K}$.

Let $\epsilon > 0$, choose $\delta = \epsilon$ and let $S, T \in \mathcal{K}$ such that $d_H(S,T) < \delta$. Then for all $\varphi \in S$ there is $\psi \in T$ such that for all $i \in \mathbb{Z}$ with $|i| < \frac{1}{\delta} - 1$ we have $\varphi(i) = \psi(i)$. Let now $\chi \in g(S)$ and $i \in \mathbb{Z}$ with $|i| < \frac{1}{\delta} - 1$. Then there is $\varphi_i \in S$ such that $\chi(i) = \varphi_i(i)$, and hence there is also $\psi_i \in T$ such that $\chi(i) = \psi_i(i)$.

So we can define $\omega \in E^{\mathbb{Z}}$ by

$$\omega(i) = \begin{cases} \psi_i(i) & \text{for } i \in \mathbb{Z} \text{ with } |i| < \frac{1}{\delta} - 1 \\ \tilde{\psi}(i) & \text{else} \end{cases}$$

for some arbitrary $\tilde{\psi} \in T$. Per construction $\omega \in g(T)$ and $\chi(i) = \omega(i)$ for all $i \in \mathbb{Z}$ with $|i| < \frac{1}{\delta} - 1$.

By interchanging S and T we analogoulsy find that for all $\omega \in g(T)$ there is $\chi \in g(S)$ such that $\omega(i) = \chi(i)$ for all $i \in \mathbb{Z}$ with $|i| < \frac{1}{\delta} - 1$ and conclude that in summary $d_H(g(S), g(T)) < \delta = \epsilon$.

Claim iii) g is shift-invariant.

This can be proven similarly to Prop. 3.5. We show that $\sigma_i g(S) = g\sigma_i(S)$ for $i \in \mathbb{Z}$ and $S \in \mathcal{K}$:

$$\begin{aligned} \sigma_i g(S) &= \{\sigma_i(\varphi) \,|\, \forall j \in \mathbb{Z} \exists \psi \in S : \varphi(j) = \psi(j)\} \\ &= \{\sigma_i(\varphi) \,|\, \forall j \in \mathbb{Z} \exists \psi \in S : \sigma_i(\varphi)(j) = \sigma_i(\psi)(j)\} \\ &= \{\varphi \,|\, \forall j \in \mathbb{Z} \exists \psi \in \sigma_i(S) : \varphi(j) = \psi(j)\} \\ &= g\sigma_i(S). \end{aligned}$$

Claim iv) There is no CNDA that has g as its global function.

Let $U = \{-m, ..., n\}$ for $m, n \in \mathbb{N}$, and let $f_0 : E^U \to \mathcal{P}(E)\backslash\{\emptyset\}$ induce $f_1 : E^{\mathbb{Z}} \to \mathcal{K}$ by $f_1(\varphi) = \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \psi(i) \in f_0(\sigma_i(\varphi)|_U)\}$. We show that there is $S \in \mathcal{K}$ such that $g(S) \neq \cup_{\varphi\in S} f_1(\varphi)$.

Case a) $f_0(u) = E' \in \mathcal{P}(E)\backslash\{\emptyset\}$ for all $u \in E^U$

Then $f_1(\varphi) = E'^{\mathbb{Z}}$ for all $\varphi \in E^{\mathbb{Z}}$. If $E' = E$, we choose two distinct elements of E, call them without restriction 0 and 1 and define $S = \{\overline{0}, ...0101...\}$. Then $\overline{1} \notin g(S)$, but $\overline{1} \in E^{\mathbb{Z}} = f_1(\overline{0}) \cup f_1(...0101...)$. Else $E' \neq E$, and there is $\varphi \in E^{\mathbb{Z}}\backslash E'^{\mathbb{Z}}$. Defining $S = \{\varphi\}$, then $\varphi \in \{\varphi\} = g(S)$, but $\varphi \notin E'^{\mathbb{Z}} = f_1(\varphi)$.

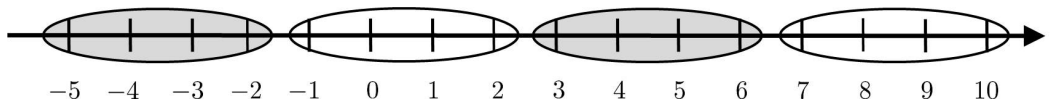


Figure 3.2: An illustration for the construction of the states in the proof of claim iv) in Thm. 3.6. With $U = \{-1, 0, 1, 2\}$, $u_1 \in E^U$ the white and u_2 the grey ellipse, the figure shows the construction of φ_1.

Case b) There is $u_1, u_2 \in E^U$ and a symbol in E which we denote by 0 such that $0 \in f_0(u_1)$, but $0 \notin f_0(u_2)$.

Consider $\varphi_1, \varphi_2 \in E^{\mathbb{Z}}$ which are constructed by gluing together u_1 and u_2 alternately, where we start with u_1 around $0 \in \mathbb{Z}$ for φ_1 and with u_2 around $0 \in \mathbb{Z}$ for φ_2; see Fig. 3.2. If $u_1(0) \neq 0$, we choose $S = \{\varphi_1\}$. Then there is $\psi \in f_1(\varphi_1)$ such that $\psi(0) = 0$, but $\varphi_1(0) \neq 0$ and hence $g(S) = \{\varphi_1\} \neq f_1(\varphi_1)$.

Else $u_1(0) = 0$, and we consider $S = \{\varphi_1, \varphi_2\}$ and let $\psi \in E^{\mathbb{Z}}$ be the state that is constructed by gluing together copies of u_1 such that $\psi(0) = 0$. Then $\psi \in g(S)$, but $\psi \notin f_1(\varphi_1)$ and $\psi \notin f_1(\varphi_2)$ because $\psi(k|U|) = 0$ for $k \in \mathbb{Z}$. Hence $g(S) \neq \cup_{\varphi \in S} f_1(\varphi)$. □

A different automaton definition would directly lead to the Curtis-Hedlund-Lyndon theorem. When applying the automaton all spatial correlation information is lost. We write $\mathcal{K}|_V := \{T|_V | T \in \mathcal{K}\}$ for $T \in \mathcal{K}$ and $V \subseteq \mathbb{Z}$.

Proposition 3.7. Let $U = \{-m, ..., n\}$ with $m, n \in \mathbb{N}$, E a finite set and $\tilde{f}_0 : \mathcal{K}|_U \to \mathcal{P}(E)\backslash\{\emptyset\}$ a local function. This induces a global function via

$$\tilde{f} : \mathcal{K} \to \mathcal{K}, \quad \tilde{f}(S) = \{\varphi \in E^{\mathbb{Z}} | \forall i \in \mathbb{Z} : \varphi(i) \in \tilde{f}_0(S|_{i+U})\}.$$

Then the continuous and shift-invariant function $\tilde{f} : \mathcal{K} \to \mathcal{K}$ is exactly the global function of this construction.

Proof Consider $\tilde{f}_0$ which induces $\tilde{f}$ as in the proposition.

Claim i) $\tilde{f}$ is closed and hence well-defined.

As in Prop. 3.2 it is enough to show that the limit point of a converging sequence in $\tilde{f}(S)$ is in $\tilde{f}(S)$ as well.

Let $(\psi_n)_{n \in \mathbb{N}}$ be a sequence in $\tilde{f}(S)$ with $\psi_n \to \psi$ as $n \to \infty$. By going to a suited subsequence if necessary we may assume that there is $\psi_{|i|} \in \tilde{f}(S)$ such that $\psi(i) = \psi_{|i|}(i)$ for any $i \in \mathbb{Z}$. Per construction of $\tilde{f}$ then $\psi(i) = \psi_{|i|}(i) \in \tilde{f}_0(S|_{i+U})$ for all $i \in \mathbb{Z}$ and hence $\psi \in \tilde{f}(S)$.

Claim ii) $\tilde{f}$ is continuous.

This statement can be proven in analogy to Prop. 3.4: We show that for each $\epsilon > 0$ there is a $\delta > 0$ such that $d_H(\tilde{f}(S), \tilde{f}(T)) < \epsilon$ if $d_H(S, T) < \delta$ for $S, T \in \mathcal{K}$.

Let $\epsilon > 0$ and choose $k \in \mathbb{N}$ such that $\frac{1}{1+k} < \epsilon$. Let furthermore $m \in \mathbb{N}$ such that $U \subseteq [-m, ..., m]$, and choose $\delta > 0$ such that $\delta < \frac{1}{1+k+m}$. For $S, T \in \mathcal{K}$ with $d_H(S, T) < \delta$

we find that

$$\begin{aligned} \forall i \in [-m-k, ..., m+k]: & \quad S(i) = T(i) \\ \Rightarrow \forall i \in [-k, ..., k]: & \quad S|_{i+U} = T|_{i+U} \\ \Rightarrow \forall i \in [-k, ..., k]: & \quad \tilde{f}_0(S|_{i+U}) = \tilde{f}_0(T|_{i+U}) \\ \Rightarrow \forall i \in [-k, ..., k]: & \quad \tilde{f}(S)(i) = \tilde{f}(T)(i), \end{aligned}$$

and therefore $d_H(\tilde{f}(S), \tilde{f}(T)) < \epsilon$.

Claim iii) $\tilde{f}$ is shift-invariant.

This can be proven similarly to Prop. 3.5. We show that $\sigma_i \tilde{f}(S) = \tilde{f}\sigma_i(S)$ for $i \in \mathbb{Z}$ and $S \in \mathcal{K}$:

$$\begin{aligned} \sigma_i \tilde{f}(S) &= \{\sigma_i(\varphi) \,|\, \forall j \in \mathbb{Z} : \varphi(j) \in \tilde{f}_0(S|_{j+U})\} \\ &= \{\sigma_i(\varphi) \,|\, \forall j \in \mathbb{Z} : \sigma_i(\varphi)(j) \in \tilde{f}_0(S|_{i+j+U})\} \\ &= \{\varphi \,|\, \forall j \in \mathbb{Z} : \varphi(j) \in \tilde{f}_0(\sigma_i(S)|_{j+U})\} \\ &= \tilde{f}\sigma_i(S), \end{aligned}$$

as $S|_{i+j+U} = \{\varphi|_{i+j+U} | \varphi \in S\} = \{\sigma_i(\varphi)|_{j+U} | \varphi \in S\} = \sigma_i(S)|_{j+U}$.

Let now in turn $\tilde{f} : \mathcal{K} \to \mathcal{K}$ be a continuous and shift-invariant function. It is enough to prove the existence of a local function $\tilde{f}_0 : \mathcal{K}|_U \to \mathcal{P}(E)\backslash\{\emptyset\}$ such that $\tilde{f}(S)(i) = \tilde{f}_0(S|_{i+U})$ for all $i \in \mathbb{Z}$ and all $S \in \mathcal{K}$ in order to prove the existence of an automaton behind $\tilde{f}$. This can be achieved by extending the idea of the classical Curtis-Hedlund-Lyndon theorem, Thm. 2.4, to power sets.

Claim iv) There is a finite set $U \subset \mathbb{Z}$ such that $\tilde{f}(S)(0)$ is fully determined by $S|_U$ for all $S \in \mathcal{K}$.

Choose $0 < \epsilon < 1$. Because $\tilde{f}$ is continuous there exists $\delta > 0$ such that $d_H(S_1, S_2) < \delta$ implies $d_H(\tilde{f}(S_1), \tilde{f}(S_2)) < \epsilon$ for $S_1, S_2 \in \mathcal{K}$. Consider $S_1, S_2 \in \mathcal{K}$ such that $S_1|_{[-i,...,i]} = S_2|_{[-i,...,i]}$ for $i \in \mathbb{Z}$ with $\frac{1}{\delta} - 2 < |i| \leq \frac{1}{\delta} - 1$, i.e., that they differ just at sites $j \in \mathbb{Z}$ with $|j| > \frac{1}{\delta} - 1$. But then already $d_H(S_1, S_2) < \delta$ and therefore $d_H(\tilde{f}(S_1), \tilde{f}(S_2)) < \epsilon$, which means that $\tilde{f}(S_1)(0) = \tilde{f}(S_2)(0)$. So we define $U = [-i, ..., i]$.

Claim v) There is a local function $\tilde{f}_0 : \mathcal{K}|_U \to \mathcal{P}(E)\backslash\{\emptyset\}$ such that $\tilde{f}(S)(i) = \tilde{f}_0(S|_{i+U})$ for all $i \in \mathbb{Z}$ and all $S \in \mathcal{K}$.

Let $\tilde{T} \in \mathcal{K}|_U$. Then there is $T \in \mathcal{K}$ such that $T|_U = \tilde{T}$, so that we can define the map by $\tilde{f}_0(\tilde{T}) = \tilde{f}(T)(0)$. Note that we can restrict any $S \in \mathcal{K}$ to U, apply $\tilde{f}_0$ and obtain $\tilde{f}_0(S|_U) = \tilde{f}(S)(0)$ per construction. Due to the shift-invariance of $\tilde{f}$ the requested property holds: for $i \in \mathbb{Z}$ and $S \in \mathcal{K}$

$$\tilde{f}(S)(i) = \sigma_{i^{-1}}\sigma_i \tilde{f}(S)(i) = \sigma_{i^{-1}}(\tilde{f}\sigma_i(S))(i) = \tilde{f}(\sigma_i(S))(0) = \tilde{f}_0(\sigma_i S|_U) = \tilde{f}_0(S|_{i+U}).$$

□

Note that for a CNDA $(\mathbb{Z}, U, E, f_0)$ with global function f we can define

$$\tilde{f}_0(\tilde{T}) = \cup_{u \in \tilde{T}} f_0(u)$$

for $\tilde{T} \in \{T|_U | T \in \mathcal{K}\}$, which induces a global function $\tilde{f}$ like introduced above. Then $f(S) \subseteq \tilde{f}(S)$ for all $S \in \mathcal{K}$. However, in the PDE application we exactly aim at the smaller $f(S)$ which preserves spatial correlations, as we want to preserve as much information as possible. This is because a physical system is supposed to be in one of the deterministic states in a random state, and so a deterministic state in the image can only have one deterministic state as predecessor. At the cost of losing the Curtis-Hedlund-Lyndon theorem we therefore stick to our CNDA definition. Nonetheless analysis of CNDA as in Sec. 3.2.4 will involve similar ideas of information loss. Also, in the construction of CPA in Chap. 4 we will allow for such loss of information to gain efficiency.

3.2 Analysis of Cellular Non-Deterministic Automata

On the one hand, we wish to analyze the long-term behavior of CNDA, but are aware that in practice this is hard for dynamical systems on power sets. On the other hand, the theory is very well developed for CA and provides classical results for ω-limit sets etc. So we suggest the concept of supersystems (SS) and especially of superautomata (SA) in Sec. 3.2.1 in order to be able to approximate a CNDA's long-time behavior by relating it to a CA's; see Fig. 3.1b.

Then we first show in Sec. 3.2.2 that a CNDA can even be embedded in a CA in principle, but that it is not possible to find a CA on $\Gamma' = \mathbb{Z}$ which serves this purpose reasonably. So second we introduce a set of SA on $\mathbb{Z}$ by the idea of de Bruijn states and pattern analysis in Sec. 3.2.3. This set is totally ordered and has a maximal element. And third we propose in Sec. 3.2.4 a subsolution SS which often enables convenient and fast analysis of CNDA and is closely linked to this maximal SA. This SS is often sufficiently exact for practical purposes.

3.2.1 Supersystems and Superautomata

We introduce SS and show why they are of interest to analyze the long-term-behavior of a CNDA. We are especially interested in deterministic SA.

Definition 3.8. Let $(\mathbb{Z}, U, E, f_0)$ be a CNDA with global function f.

i) A time-discrete dynamical system (X', f') on compact X' is called *supersystem (SS)*, if there are continuous maps

$$\beta' : X' \to \mathcal{K} \quad \text{and} \quad \gamma' : \mathcal{K} \to X'$$

such that for all $S \in \mathcal{K}$ and all $k \in \mathbb{N}$

$$f^k(S) \subseteq \beta' f'^k \gamma'(S).$$

ii) A SS (X', f') with $\beta' : X' \to \mathcal{K}$ and $\gamma' : \mathcal{K} \to X'$ is called *smaller* than a SS (X'', f'') with $\beta'' : X'' \to \mathcal{K}$ and $\gamma'' : \mathcal{K} \to X''$, if for all $S \in \mathcal{K}$ and all $k \in \mathbb{N}$

$$\beta' f'^k \gamma'(S) \subseteq \beta'' f''^k \gamma''(S).$$

It is called *strictly smaller* if there is $S \in \mathcal{K}$ and $k \in \mathbb{N}$ such that the subset-relation is strict. The property *(strictly) bigger* is defined with the subset- replaced by the superset-relation.

iii) A SS is called *minimal* in a set of SS, if there is no strictly smaller SS in this set, and *maximal* if there is no SS in it which is strictly bigger.
iv) If f' is the global function of a deterministic CA (Γ', U', E', f'_0) on $X' = E'^{\Gamma'}$, we call the SS a *(deterministic) superautomaton (SA)*.

We say that a SS *covers* the CNDA or a smaller SS, or that the trajectory under one of the latter is *contained* in the trajectory under the SS. Now we investigate an order relation on SS.

Definition 3.9. A *binary relation* $\leq$ on a set S is a collection of ordered pairs of elements of S. If the relation fulfills for all $r, s, t \in S$
i) $r \leq r$ (*reflexivity*),
ii) if $r \leq s$ and $s \leq r$, then $r = s$ (*antisymmetry*),
iii) if $r \leq s$ and $s \leq t$, then $r \leq t$ (*transitivity*),
then it is called a *partial order*. If it only fulfills i) and iii), then it is called a *preorder*.

Lemma 3.10. The smaller-relation on SS is a preorder.

Proof The preorder is induced by the partial order $\subseteq$ on $\mathcal{K}$. The relation on SS is reflexive, as $\subseteq$ is reflexive; $\beta' f'^k \gamma'(S) \subseteq \beta' f'^k \gamma'(S)$ for all $S \in \mathcal{K}$ and all $k \in \mathbb{N}$ implies that (X', f') is smaller than (X', f'). Similarly, it is transitive because if (X', f') is smaller than (X'', f'') and $(X'' f'')$ is smaller than (X''', f''') it is implied that $\beta' f'^k \gamma'(S) \subseteq \beta'' f''^k \gamma''(S)$ and $\beta'' f''^k \gamma''(S) \subseteq \beta''' f'''^k \gamma'''(S)$ for all $S \in \mathcal{K}$ and all $k \in \mathbb{N}$. Therefore $\beta' f'^k \gamma'(S) \subseteq \beta''' f'''^k \gamma'''(S)$ for all $S \in \mathcal{K}$ and all $k \in \mathbb{N}$, and so (X', f') is smaller than (X''', f'''). □

Clearly the smaller a SS, the exacter it describes the CNDA's time-behavior. If even an embedding α into a SS (X', f') is possible, it is minimal in the set of all SS because $\alpha^{-1} f'^k \alpha(S) = f^k(S)$ for all $S \in \mathcal{K}$ and all $k \in \mathbb{N}$ then.

Proposition 3.11. Let $(\mathbb{Z}, U, E, f_0)$ be a CNDA with global function f, (X', f') a SS with $\beta' : X' \to \mathcal{K}$ and $\gamma' : \mathcal{K} \to X'$ and (X'', f'') a SS with $\beta'' : X'' \to \mathcal{K}$ and $\gamma'' : \mathcal{K} \to X''$, where the first SS is smaller than the second. Then the ω-limit sets are ordered in the sense that for all $S \in \mathcal{K}$
i) $\forall T \in \omega_f(S)\, \exists T' \in \omega_{f'}(\gamma'(S)) : T \subseteq \beta'(T')$ and
ii) $\forall T' \in \omega_{f'}(\gamma'(S))\, \exists T \in \omega_f(S) : T \subseteq \beta'(T')$,
iii) $\forall T' \in \omega_{f'}(\gamma'(S))\, \exists T'' \in \omega_{f''}(\gamma''(S)) : \beta'(T') \subseteq \beta''(T'')$ and
iv) $\forall T'' \in \omega_{f''}(\gamma''(S))\, \exists T' \in \omega_{f'}(\gamma'(S)) : \beta'(T') \subseteq \beta''(T'')$.

Proof We only prove the first statement, the others can be shown analogously.

Let $T \in \omega_f(S)$. There is a strictly monotonically increasing sequence $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ such that $\lim_{k\to\infty} f^{n_k}(S) = T$. Then $(f'^{n_k}(\gamma'(S)))_{k\in\mathbb{N}}$ as a sequence on compact X' has a convergent subsequence $(f'^{n_l}\gamma'(S))_{l\in\mathbb{N}}$, and hence there is $T' \in \omega_{f'}(\gamma'(S))$ such that $\lim_{l\to\infty} f'^{n_l}(\gamma'(S)) = T'$. The proof is complete when we show that necessarily $T \subseteq \beta'(T')$.

Let $\varphi \in T$. According to Thm. 2.6 ii) there is a sequence (φ_{n_l}) with $\varphi_{n_l} \in f^{n_l}(S)$ such that $\lim_{n_l\to\infty} \varphi_{n_l} = \varphi$. But then $\varphi_{n_l} \in \beta' f'^{n_l} \gamma'(S)$ for all $l \in \mathbb{N}$, and as β' is continuous thus $\varphi \in \beta'(T')$ again by Thm. 2.6 ii). □

3.2.2 Embedding in Deterministic Cellular Automata

3.2.2.1 General Embedding

To show that a CNDA can be embedded in a deterministic CA we use a result of Müller and Spandl [129, 130]. It states that any time-discrete dynamical system $(\mathfrak{C}, f)$ given by a continuous function f on a Cantor space $\mathfrak{C}$ can be embedded in a deterministic CA on a non-Abelian Cayley graph.

Theorem 3.12. $(\mathcal{K}, d_H)$ is a Cantor space.

Proof It is well known that $\mathcal{K}$ is perfect [95], and according to Thm. 2.6 i) it is compact. Because $\mathcal{K}$ is clearly non-empty, it just remains to prove that it is totally disconnected.

Let $S \in \mathcal{K}$ and $U \subset \mathcal{K}$ an open set with $S \in U$. We show that there is a clopen set V with $S \in V \subseteq U$. As U is open, there is $k \in \mathbb{N}$ such that $B_{\frac{1}{k}}(S) \subseteq U$. Because the distance d_H only takes on the discrete values 0 and $\frac{1}{k+1}$ for $k = 0, 1, ...$, a property inherited from the underlying Cantor metric,

$$B_{\frac{1}{k}}(S) = \{T \in \mathcal{K} | d_H(S,T) < \frac{1}{k}\} = \{T \in \mathcal{K} | d_H(S,T) \leq \frac{1}{k + \frac{1}{2}}\}.$$

So the open ball $B_{\frac{1}{k}}(S)$ conincides with the closed ball $\overline{B_{\frac{1}{k+\frac{1}{2}}}(S)}$ and is hence a clopen set V in U which contains S. □

Note that the empty set is isolated in $\mathcal{K} \cup \{\emptyset\}$ and that therefore it is excluded in our definition of $\mathcal{K}$.

Corollary 3.13. Let $(\mathbb{Z}, U, E, f_0)$ be a CNDA with global function f. Then $(\mathcal{K}, f)$ is a dynamical system that can be embedded in a CA.

As can be seen in the construction of the CA it is a shortcoming of this embedding that there is almost no chance to determine its actual global function in practice. Furthermore, the grid underlying the CA belongs to an aritificial group that does not resemble the original structure, $\mathbb{Z}$, at all. In consequence this embedding seems not to be applicable in practice. We ask next if there is a reasonable embedding in a CA which at least operates on grid $\mathbb{Z}$.

3.2.2.2 Embedding in a CA on $\mathbb{Z}$?

There exists a homeomorphism $\alpha : \mathcal{K} \to E'^{\mathbb{Z}}$ for arbitrary finite E', as both sets are Cantor spaces [180]. So there is an embedding of the CNDA $(\mathbb{Z}, U, E, f_0)$ with global function f in a dynamical system $(E'^{\mathbb{Z}}, f' := \alpha f \alpha^{-1})$. However, f' is not necessarily the global function of a CA $(\mathbb{Z}, V, E', f'_0)$. According to Thm. 2.4 f' would at least need to be shift-invariant for this purpose, but from a reasonable embedding α we expect a little bit more:

Definition 3.14. Let E' be a finite set. A map $\alpha : \mathcal{K} \to E'^{\mathbb{Z}}$ is called *shift-invariant* if $\alpha \circ \sigma_i = \sigma_i \circ \alpha$ for all $i \in \mathbb{Z}$.

If the embedding α has this property, the shift-invariance of f' follows immediately with Lm. 3.5: for all $i \in \mathbb{Z}$ we have

$$\sigma_i f' \alpha = \sigma_i \alpha f = \alpha \sigma_i f = \alpha f \sigma_i = f' \alpha \sigma_i = f' \sigma_i \alpha.$$

The following theorem now answers the question, whether there is an embedding of a CNDA in a CA which is reasonable in this sense: there is none.

Theorem 3.15. Let $|E| > 1$. In general there is no continuous, injective and shift-invariant map from $\mathcal{K}$ to any $E'^{\mathbb{Z}}$, where E' is a finite set.

Proof We show that because of the continuity and the shift-invariance any potential map acts locally and then rule out injectivity by constructing a counterexample. The proof of the locality is very similar to the proof of Prop. 3.7.

Let E' be a finite set and $\alpha : \mathcal{K} \to E'^{\mathbb{Z}}$ continuous and shift-invariant.

Claim i) There is a finite set $U \subset \mathbb{Z}$ such that $\alpha(S)(0)$ is fully determined by $S|_U$ for all $S \in \mathcal{K}$.

Choose $0 < \epsilon < 1$. Because α is continuous there exists $\delta > 0$ such that $d_H(S_1, S_2) < \delta$ implies $d_C(\alpha(S_1), \alpha(S_2)) < \epsilon$ for $S_1, S_2 \in \mathcal{K}$. Consider $S_1, S_2 \in \mathcal{K}$ such that $S_1|_{[-i,...,i]} = S_2|_{[-i,...,i]}$ for $i \in \mathbb{Z}$ with $\frac{1}{\delta} - 2 < |i| \leq \frac{1}{\delta} - 1$, i.e., that they differ just at sites $j \in \mathbb{Z}$ with $|j| > \frac{1}{\delta} - 1$. But then already $d_H(S_1, S_2) < \delta$ and therefore $d_C(\alpha(S_1), \alpha(S_2)) < \epsilon$, which means that $\alpha(S_1)(0) = \alpha(S_2)(0)$. So we define $U = [-i, ..., i]$.

Claim ii) There is a map $\alpha_0 : \mathcal{K}|_U \to E'$ such that $\alpha(S)(i) = \alpha_0(S|_{i+U})$ for all $i \in \mathbb{Z}$ and all $S \in \mathcal{K}$.

Let $\tilde{T} \in \mathcal{K}|_U$. Then there is $T \in \mathcal{K}$ such that $T|_U = \tilde{T}$, so that we can define the map by $\alpha_0(\tilde{T}) = \alpha(T)(0)$. Note that we can restrict any $S \in \mathcal{K}$ to U, apply α_0 and obtain $\alpha_0(S|_U) = \alpha(S)(0)$ per construction. Due to the shift-invariance of α the requested property holds: for $i \in \mathbb{Z}$ and $S \in \mathcal{K}$

$$\alpha(S)(i) = \sigma_{i^{-1}} \sigma_i \alpha(S)(i) = \sigma_{i^{-1}}(\alpha \sigma_i(S))(i) = \alpha(\sigma_i(S))(0) = \alpha_0(\sigma_i S|_U) = \alpha_0(S|_{i+U}).$$

Claim iii) α is not injective.

As $|E| > 1$, we find two different symbols which we call $0, 1 \in E$ without restriction and define the deterministic states

$$\begin{aligned} \varphi_1 &= (...01\ 0...0\ 10...), & \varphi_2 &= (...01\ 0...0\ 00...), \\ \psi_1 &= (...00\ \underbrace{0...0}_{|U_0|+1}\ 00...), & \psi_2 &= (...00\ \underbrace{0...0}_{|U_0|+1}\ 10...). \end{aligned}$$

According to Lm. 3.3 $S_1 = \{\varphi_1, \psi_1\} \in \mathcal{K}$ and $S_2 = \{\varphi_2, \psi_2\} \in \mathcal{K}$. Then for all $i \in \mathbb{Z}$

$$\alpha(S_1)(i) = \alpha_0(S_1|_{i+U}) = \alpha_0(S_2|_{i+U}) = \alpha(S_2)(i)$$

and therefore $\alpha(S_1) = \alpha(S_2)$. As $S_1 \neq S_2$, α is not injective. □

3.2.3 Pattern Based Analysis

We suggest pattern SA on $\mathbb{Z}$ that allow for control of a CNDA's limit set and for easier analysis than with the embedding of the last section. Before pattern SA are introduced we first introduce the underlying de Bruijn state calculus which we developed on the basis of pattern ideas in CA theory [76, 173], in the theory of de Bruijn graphs [162] and in pair approximation [96].

3.2.3.1 De Bruijn State Calculus

As usual in this section E denotes a finite set and $V = \{-m, ..., n\}$ for $m, n = 0, 1, ...$

Definition 3.16. A *de Bruijn state* is an element of the *de Bruijn space* $X_{dB} := (\mathcal{P}(E^V))^{\mathbb{Z}}$.

Note that the Cantor metric can be introduced on X_{dB} because $\mathcal{P}(E^V)$ is finite.

Definition 3.17. For all $\Phi, \Psi \in X_{dB}$

$$\Phi \subseteq_{dB} \Psi \Leftrightarrow \forall i \in \mathbb{Z} : \Phi(i) \subseteq \Psi(i).$$

Lemma 3.18. $\subseteq_{dB}$ is a partial order on X_{dB}.

Proof The partial order on X_{dB} is induced by the partial order $\subseteq$ on $\mathcal{P}(E^V)$. Let $\Phi, \Psi, \Omega \in X_{dB}$. Then the relation $\subseteq_{dB}$ is reflexive, as $\subseteq$ is reflexive; $\Phi(i) \subseteq \Phi(i)$ for all $i \in \mathbb{Z}$ implies $\Phi \subseteq_{dB} \Phi$. It is antisymmetric, as per definition $\Phi \subseteq_{dB} \Psi$ and $\Psi \subseteq_{dB} \Phi$ implies that $\Phi(i) \subseteq \Psi(i)$ and $\Psi(i) \subseteq \Phi(i)$ for all $i \in \mathbb{Z}$. Because $\subseteq$ is antisymmetric, it follows that $\Phi(i) = \Psi(i)$ for all $i \in \mathbb{Z}$ and therefore $\Phi = \Psi$. Similarly, it is transitive because $\Phi \subseteq_{dB} \Psi$ and $\Psi \subseteq_{dB} \Omega$ imply $\Phi(i) \subseteq \Psi(i)$ and $\Psi(i) \subseteq \Omega(i)$ for all $i \in \mathbb{Z}$. Therefore $\Phi(i) \subseteq \Omega(i)$ for all $i \in \mathbb{Z}$ and so $\Phi \subseteq_{dB} \Omega$. □

We also write $\subseteq$ for $\subseteq_{dB}$ in the following, as the meaning is always clear from the context. Next we define functions relating X_{dB} and $\mathcal{K}$.

Definition 3.19. We set $\beta : X_{dB} \to \mathcal{K} \cup \{\emptyset\}$, $\Phi \mapsto \beta(\Phi)$ with

$$\beta(\Phi) = \{\varphi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \varphi|_{i+V} \in \Phi(i)\}$$

and $\gamma : \mathcal{K} \cup \{\emptyset\} \to X_{dB}$, $S \mapsto \gamma(S)$ with

$$(\gamma(S))(i) = \{v \in E^V \,|\, \exists \varphi \in S : v = \varphi|_{i+V}\}.$$

Fig. 3.3 illustrates how β and γ work. As usual it has to be checked first that β really maps to compact sets.

Lemma 3.20. Let $\Phi \in X_{db}$. Then $\beta(\Phi)$ is closed.

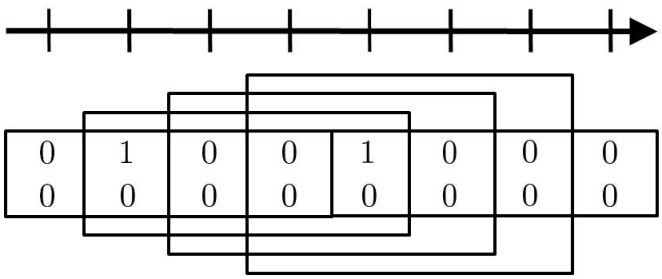


Figure 3.3: Reconsider S_1 of the example of the third claim in the proof of Thm. 3.15 with $U_0 = \{0\}$. For $V = \{0, 1, 2, 3\}$ the de Bruijn state $\gamma(S_1)$ is depicted by the rectangles. $(\gamma(S_1))(0) = \{(0100), (0000)\}$ is the smallest rectangle at the very left, $(\gamma(S_1))(1) = \{(1001), (0000)\}$ the next tallest etc. By gluing them back together such that they coincide on the overlap we find that $\beta\gamma(S_1) = S_1$.

Proof We show that the limit of any convergent sequence in $\beta(\Phi)$ is also in $\beta(\Phi)$.

Let $(\varphi_n)_{n\in\mathbb{N}}$ be a sequence in $\beta(\Phi)$ with $\varphi_n \to \varphi \in \mathcal{K}$ for $n \to \infty$. Let $i \in \mathbb{Z}$. There exists $N \in \mathbb{N}$ such that $\varphi_N|_{[-i-|V|,\ldots,i+|V|]} = \varphi|_{[-i-|V|,\ldots,i+|V|]}$. Therefore $\varphi|_{i+V} = \varphi_N|_{i+V} \in \Phi(i)$ and $\varphi \in \beta(\Phi) = \{\psi \in E^{\mathbb{Z}} | \forall i \in \mathbb{Z} : \psi|_{i+V} \in \Phi(i)\}$. □

A state $\Phi \in X_{db}$ with $\text{card}(\beta(\Phi)) = 1$ is called a *deterministic de Bruijn state*. For any deterministic state $\varphi \in E^{\mathbb{Z}}$ there is a deterministic de Bruijn state $\Phi \in X_{dB}$ given by $\Phi(i) = \{\varphi|_{i+V}\}$ for all $i \in \mathbb{Z}$. It holds that $\beta(\Phi) = \{\varphi\}$, $\gamma(\{\varphi\}) = \Phi$, $\gamma\beta(\Phi) = \Phi$ and $\beta\gamma(\{\varphi\}) = \{\varphi\}$ then, such that β and γ are each other's inverse on these states.

There are many de Bruijn states that are mapped to the empty set by β: let, e.g., $E = \{0, 1\}$ and $\Phi \in X_{dB}$ given by $\Phi(i) = \{\overline{0}_V\}$ and $\Phi(i + 1) = \{\overline{1}_V\}$ for even $i \in \mathbb{Z}$. Then $\beta(\Phi) = \emptyset$, as an incompatibility at any neighboring sites in the grid forces the image of the de Bruijn state under β to be the empty set.

As a consequence, β is not continuous on X_{dB}. We approximate e.g. $\Phi \in X_{db}$ with $\beta(\Phi) \neq \emptyset$ arbitrarily close by a sequence $(\Phi_n)_{n\in\mathbb{N}}$ in X_{db} with $\beta(\Phi_n) = \emptyset$ for all $n \in \mathbb{N}$ by inserting in Φ an incompatibility at neighboring sites $n, n + 1 \in \mathbb{N}$. Then $d_C(\Phi, \Phi_n) = \frac{1}{1+n}$, but $d_H(\beta(\Phi), \beta(\Phi_n)) = d_H(\beta(\Phi), \emptyset) = 1$.

We characterize a subset of X_{dB} on which β has nicer properties. In particular this subset is not mapped to the empty set by β.

Definition 3.21. Let $V^0 := \{0\}$, $V^1 := V$ and recursively for $p \in \mathbb{N}, 1 < p$

$$V^p := \{l \in \mathbb{Z} \,|\, \exists k \in V : l \in k + V^{p-1}\}.$$

Consider $\Phi \in X_{dB}$ with $\Phi(i) \neq \emptyset$ for all $i \in \mathbb{Z}$, $j \in \mathbb{Z}$, and $p = 1, 2, \ldots$. Then $v \in E^{V^p}$ with

$$\forall k \in V^{p-1} : v|_{k+V} \in \Phi(j + k).$$

is called a *pattern* or *template of* Φ *at location* j *of extension* p. If for all patterns v of Φ there is $\varphi \in \beta(\Phi)$ s.t. $\varphi|_{j+V^p} = v$, we say that Φ *has the extension property*. The subset of elements in X_{dB} that have the extension property is denoted by X_{dBe}.

So for $\Phi \in X_{dBe}$ and given $j \in \mathbb{N}$ we find for all $i \in [-j, j]$ and all $v_i \in \Phi(i)$ a $\varphi \in E^{\mathbb{Z}}$ such that $(\gamma(\{\varphi\}))(i) = v_i$. By gluing together such finite patterns we can extend the construction to infinity. Fig. 3.4a gives an example of how patterns can be understood in terms of the underlying sites.

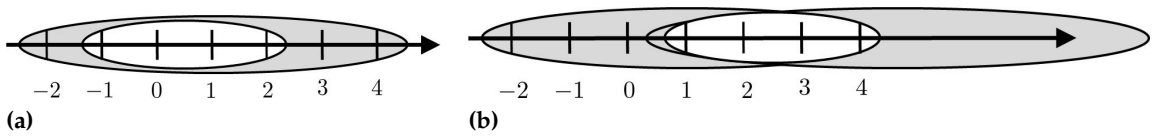


Figure 3.4: (a) Exemplary pattern sketches for $m = 1, n = 2$. The inner/outer ellipse depicts the sites $i \in \mathbb{Z}$ belonging to patterns at location $j = 0$ of extension $p = 1/p = 2$. (b) Example of how a pattern can be extended to the right; see the proof of Lm. 3.22. The left dark ellipse depicts v, the right dark ellipse is the part of ψ that is used for extension and the white ellipse is the interval on which they coincide, $(0 + (2 - 1)2) + \{-1, 0, 1, 2\}$.

Lemma 3.22. $\gamma(S) \in X_{dBe}$ for $S \in \mathcal{K}$.

Proof We show that $\gamma(S)$ possesses the extension property for $S \in \mathcal{K}$. Let $v \in E^{V^p}$ be a pattern of $\gamma(S)$ at location $j \in \mathbb{Z}$ of extension p. This is, the length of v is $p(n + m) + 1$. The right hand side of v ends at site $j + pn$; the last neighborhood of length $n + m + 1$ in the interval in $\mathbb{Z}$ that is covered by $j + V^p$ is centered around $j + (p - 1)n$. There is $w \in \gamma(S)(j + (p - 1)n)$ such that $w(k) = v((p - 1)n + k)$ for $k \in V$. Then there is $\psi \in S$ s.t. ψ also coincides with v on $(j + (p - 1)n) + V$; see also Fig. 3.4b. We may thus extend the pattern v to the right using ψ,

$$\varphi(k) = \begin{cases} v(l - j) & \text{for} \quad l \in j + V^p \\ \psi(k) & \text{for} \quad k > j + pn \end{cases} .$$

Similarly, we can extend the pattern to the left. In the end $\varphi|_{j+V^p} = v$ for $\varphi \in \beta\gamma(S)$, as φ can be thought of as glued together by the patterns in the extending states and v. □

Thus from now on we use β and γ only as maps between the restricted sets X_{dBe} and $\mathcal{K}$.

Lemma 3.23. Consider $\beta : X_{dBe} \to \mathcal{K}$ and $\gamma : \mathcal{K} \to X_{dBe}$. Then
i) β is continuous and shift-invariant.
ii) γ is continuous and shift-invariant.

Proof Claim i) Let $\epsilon > 0$. Choose $\delta = \epsilon$ and assume $d_C(\Phi_1, \Phi_2) < \delta$ for $\Phi_1, \Phi_2 \in X_{dBe}$. For $i \in \mathbb{Z}$ with $|i| < \frac{1}{\delta} - 1$ then $\Phi(i) = \Psi(i)$, so that for all $\varphi \in \beta(\Phi_1)$ there is $\psi \in \beta(\Phi_2)$ with $\varphi(i) = \psi(i)$ and vice versa. Note that for such a ψ to exist it is necessary that Φ_2 is extendable. Therefore $d_H(\beta(\Phi_1), \beta(\Phi_2)) < \epsilon$ and hence β is continuous.

Furthermore β is shift-invariant since for $i \in \mathbb{Z}$ and $\Phi \in X_{dBe}$

$$\begin{aligned} \sigma_i\beta(\Phi) &= \{\sigma_i(\varphi) \in E^{\mathbb{Z}} \,|\, \forall j \in \mathbb{Z} : \varphi|_{j+V} \in \Phi(j)\} \\ &= \{\varphi \in E^{\mathbb{Z}} \,|\, \forall j \in \mathbb{Z} : \varphi|_{j+V} \in \Phi(i + j)\} = \beta\sigma_i(\Phi). \end{aligned}$$

Claim ii) Let $\epsilon > 0$. Choose $\delta = \frac{1}{\frac{1}{\epsilon} + |V|}$ and assume $d_H(S_1, S_2) < \delta$ for $S_1, S_2 \in \mathcal{K}$. Consider $v \in \gamma(S_1)(i)$ for $|i| < \frac{1}{\epsilon} - 1$. By definition there exists $\varphi \in S_1$ such that $v = \varphi|_{i+V}$, and because of $d_H(S_1, S_2) < \delta$ there is $\psi \in S_2$ such that $\varphi(j) = \psi(j)$ for $j \in \mathbb{Z}$ with $|j| < \frac{1}{\delta} - 1 = \frac{1}{\epsilon} - 1 + |V|$. So $v \in \gamma(S_2)(i)$ and $\gamma(S_1)(i) \subseteq \gamma(S_2)(i)$, and analogous considerations also yield $\gamma(S_2)(i) \subseteq \gamma(S_1)(i)$. Therefore $d_C(\gamma(S_1), \gamma(S_2)) < \epsilon$ and hence γ is continuous.

Furthermore γ is shift-invariant since for $i, j \in \mathbb{Z}$ and $S \in \mathcal{K}$

$$\begin{aligned}(\sigma_i\gamma(S))(j) = \gamma(S)(i+j) &= \{v \in E^V \mid \exists\varphi \in S : \; v = \varphi|_{i+j+V}\} \\ &= \{v \in E^V \mid \exists\varphi \in \sigma_i(S) : \; v = \varphi|_{j+V}\} = (\gamma\sigma_i(S))(j).\end{aligned}$$

□

According to Thm. 3.15 γ cannot be injective, but is surjective on X_{dBe} per construction. β on contrast is injective by construction, but not surjective: consider again the example in the third claim of the proof of Thm. 3.15 and identify $V = U_0$. Then e.g. S_1 does not have any preimage in X_{dBe}, because $\beta(\Phi) = S_1 \cup S_2$ if the patterns of S_1 are collected in $\Phi \in X_{dBe}$.

Although β and γ served as each other's inverse on deterministic states, the considerations so far indicate that this cannot be true in general. However, we find the following:

Lemma 3.24. $\Phi = \gamma\beta(\Phi)$ for $\Phi \in X_{dBe}$.

Proof We show both inclusions for $\Phi \in X_{dBe}$. Consider first $i \in \mathbb{Z}$ and $v \in \Phi(i)$. Since Φ has the extension property and v is a pattern of extension 1, there is $\varphi \in \beta(\Phi)$ such that $\varphi|_{i+V} = v$. By definition of γ then also $v \in (\gamma\beta(\Phi))(i)$.

Consider now $i \in \mathbb{Z}$ and $v \in (\gamma\beta(\Phi))(i)$. By definition of γ then there is $\varphi \in \beta(\Phi)$ such that $v = \varphi|_{i+V}$. By definition of β then already $v \in \Phi(i)$.

□

Lemma 3.25. $S \subseteq \beta\gamma(S)$ for $S \in \mathcal{K}$.

Proof The claim is a direct consequence of the definition of the maps γ and β. Application of γ extracts the patterns of extension 1 from the given states, and application of β glues them back together. There may be more states in the end, but at least the original states are reconstructed.

□

In general the other inclusion $\beta\gamma(S) \subseteq S$ is not true for $S \in \mathcal{K}$: reconsider again the example of the third claim in the proof of Thm. 3.15 with $V = U_0 = \{0\}$. Then

$$\gamma(S_1)(i) = \begin{cases} \{(0),(1)\} & \text{for } i = 1, i = 4 \\ \{(0)\} & \text{else} \end{cases},$$

for $i \in \mathbb{Z}$ and e.g. $\varphi_2 = (...010000...) \in \beta\gamma(S_1)$, but $\varphi_2 \notin S_1$.

Proposition 3.26. $\gamma(S)$ is the smallest state in X_{dBe} of which the image under β covers a given $S \in \mathcal{K}$.

Proof Let $S \in \mathcal{K}$. According to Lm. 3.22 and 3.25 $\gamma(S) \in X_{dBe}$, and $\beta\gamma(S)$ covers S. So it remains to show that $\gamma(S)$ is the smallest state with that property, i.e., that for all $\Phi \in X_{dBe}$ with $S \subseteq \beta(\Phi)$ we find $\gamma(S) \subseteq \Phi$. But as for every $i \in \mathbb{Z}$ every element in $\gamma(S)(i)$ corresponds to a state $\varphi \in S$, the image under β of any state in X_{dBe} that is strictly smaller than $\gamma(S)$ cannot cover the complete set S. □

We close our basic considerations by comparing properties of patterns of different length. Let $V_1, V_2 \subseteq \mathbb{Z}$, $0 \in V_1, V_2$ define two de Bruijn spaces X_{dB_i} and according maps $\beta_i : X_{dBe_i} \to \mathcal{K}, \gamma_i : \mathcal{K} \to X_{dBe_i}$ for $i = 1, 2$.

Lemma 3.27. $V_1 \subseteq V_2$ implies that $\beta_2\gamma_2(S) \subseteq \beta_1\gamma_1(S)$ for all $S \in \mathcal{K}$.

Proof Let $S \in \mathcal{K}, i \in \mathbb{Z}$ and consider $\varphi \in \beta_2\gamma_2(S)$. By definition $\varphi|_{i+V_2} \in \gamma_2(S)(i)$, and therefore there exists $\psi \in S$ such that $\varphi|_{i+V_2} = \psi|_{i+V_2}$. As $V_1 \subseteq V_2$ then also $\varphi|_{i+V_1} = \psi|_{i+V_1}$ and $\varphi|_{i+V_1} \in \gamma_1(S)(i)$. As this is true for all $i \in \mathbb{Z}$ it can be concluded that $\varphi \in \beta_1\gamma_1(S)$. □

3.2.3.2 Pattern Superautomata

In this section we introduce SA on X_{dBe}. We suggest a system, show first that it is a CA and then that it is really a SS. Afterwards we investigate an order of these SA for different pattern lengths.

$U \subseteq \mathbb{Z}$ always denotes the neighborhood of a CNDA and $V \subseteq \mathbb{Z}$ a set to define the de Bruijn states $X_{dB} = (\mathcal{P}(E^V))^{\mathbb{Z}}$. We need the definition $X_{dBe}|_U := \{\Psi|_U \mid \Psi \in X_{dBe}\} = \{v \in (\mathcal{P}(E^V))^U \mid \exists \Psi \in X_{dBe} : v = \Psi|_U\}$ in the following.

Theorem 3.28. Let $(\mathbb{Z}, U, E, f_0)$ be a CNDA with global function f and $V = \{-m, ..., n\}$ for $m, n \in \mathbb{N}$. There is a local function $\tilde{g}_0 : X_{dBe}|_U \to \mathcal{P}(E^V)$ such that for $i \in \mathbb{Z}$ the map

$$g : X_{dBe} \to X_{dBe}, \quad \Phi \mapsto g(\Phi) = \gamma f \beta(\Phi)$$

acts as

$$g(\Phi)(i) = \tilde{g}_0(\Phi|_{i+V}).$$

Proof Let $\Phi \in X_{dBe}, i \in \mathbb{Z}$ and denote $S = f(\beta(\Phi))$. $\gamma(S)(i)$ can be computed if we know S only locally, this is, if we only know $\{\varphi|_{i+V} \mid \varphi \in S\}$. In order to determine $\{\varphi|_{i+V} \mid \varphi \in S\}$ it is sufficient to know $\{\varphi|_{i+j+U} \mid j \in V, \varphi \in \beta(\Phi)\}$. However, as Φ has the extension property any local pattern can be extended to a global state. Therefore we only need to know $\Phi(k)$ for $k \in i + U$, and hence there is a map $\tilde{g}_0$ with the required properties. □

Corollary 3.29. Let $(\mathbb{Z}, U, E, f_0)$ denote a CNDA with global function f and $V = \{-m, ..., n\}$ for $m, n \in \mathbb{N}$. Then there is a CA $(\mathbb{Z}, U, \mathcal{P}(E^V), g_0)$ with global function $g : X_{dB} \to X_{dB}$ such that $g(\Phi) = \gamma f \beta(\Phi)$ for $\Phi \in X_{dBe}$.

Proof We just extend the local function $\tilde{g}_0$ of Thm. 3.28 from $X_{dBe}|_U$ to $(\mathcal{P}(E^V))^U$, as a CA is defined on the whole space. This can be achieved via

$$g_0 : (\mathcal{P}(E^V))^U \to \mathcal{P}(E^V), \qquad v \mapsto \begin{cases} \tilde{g}_0(v) & \text{if} \quad v \in X_{dBe}|_U \\ \emptyset & \text{else} \end{cases}.$$

Then per construction the according global function g agrees with $\gamma f \beta$ on X_{dBe}. □

Fig. 3.5 shows how the CA works. The reason for the introduction of X_{dBe} is the requirement of locality for β.

It remains to show that for arbitrary V such a CA is truly a SA, and that the set of these SA can be totally ordered.

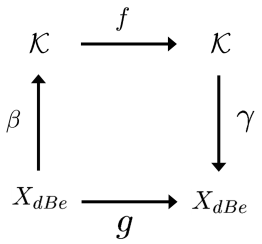


Figure 3.5: The global function g of the CA works on X_{dBe} like the global function f of the CNDA after having translated the underlying spaces.

Lemma 3.30. $f^k(S) \subseteq \beta g^k \gamma(S)$ for all $k = 0, 1, ...$ and all $S \in \mathcal{K}$.

Proof The claim follows by induction, and the initial step for $k = 0$ is given by Lm. 3.25. Let $k \in \mathbb{N}$ and $S \in \mathcal{K}$. Assuming that the claim holds for $k - 1$, we can calculate

$$f^k(S) = ff^{k-1}(S) \subseteq f\beta g^{k-1}\gamma(S) \subseteq \beta\gamma f\beta g^{k-1}\gamma(S) = \beta g^k \gamma(S).$$

□

Lemma 3.31. Let $V_1, V_2 \subseteq \mathbb{Z}$ define two de Bruijn spaces X_{dB_i} and according maps $\beta_i : X_{dBe_i} \to \mathcal{K}, \gamma_i : \mathcal{K} \to X_{dBe_i}$ and $g_i = \beta_i f \gamma_i$ for $i = 1, 2$. Then $V_1 \subseteq V_2$ implies that $\beta_2 g_2^k \gamma_2(S) \subseteq \beta_1 g_1^k \gamma_1(S)$ for all $k = 0, 1, ...$ and all $S \in \mathcal{K}$.

Proof The claim follows by induction, and the initial step for $k = 0$ is given by Lm. 3.27. Let $V_1 \subseteq V_2$ and $S \in \mathcal{K}$. Assuming that the claim holds for $k - 1$, we can calculate

$$\beta_2 g_2^k \gamma_2(S) = \beta_2\gamma_2 f\beta_2 g_2^{k-1}\gamma_2(S) \subseteq \beta_2\gamma_2 f\beta_1 g_1^{k-1}\gamma_1(S) \subseteq \beta_1\gamma_1 f\beta_1 g_1^{k-1}\gamma_1(S).$$

□

As there is $S \in \mathcal{K}$ with $\beta\gamma(S) \not\subseteq S$, we cannot expect equality in Lm. 3.30 for a whole trajectory in general. Indeed there are CNDA such that no pattern SA is capable of covering it precisely, even if one restricts oneself to trajectories starting with deterministic states. This is a consequence of correlations beyond the patterns' locality.

Proposition 3.32. There is a CNDA $(\mathbb{Z}, U, E, f_0)$ and $\varphi \in E^{\mathbb{Z}}$ such that for all V there exists $k \in \mathbb{Z}$ with $\beta g^k \gamma(\{\varphi\}) \not\subseteq f^k(\{\varphi\})$.

Proof Consider $U = \{-1, 0\}, E = \{0, 1, 2, 3, 4, 5, 6\}$ and f_0 given by

$\varphi(-1)$	$\varphi(0)$	$f_0((\varphi(-1), \varphi(0)))$
0	0	$\{0\}$
0	1	$\{0, 2\}$
0	2	$\{2\}$
0	3	$\{0\}$
0	5	$\{0\}$
0	6	$\{0\}$
1	0	$\{3\}$

$\varphi(-1)$	$\varphi(0)$	$f_0((\varphi(-1), \varphi(0)))$
2	0	$\{0\}$
2	3	$\{4\}$
2	4	$\{0\}$
3	0	$\{5\}$
4	5	$\{6\}$
5	0	$\{0\}$
6	0	$\{6\}$,

where the lacking combinations may be defined at will. We choose $\{\varphi\} = \{(...001000...)\}$ and calculate for all V with $|V| = l \in \mathbb{N}$

$$f(\{\varphi\}) = f\beta\gamma(\{\varphi\}) = \beta\underbrace{\gamma f\beta}_{g}\gamma(\{\varphi\}) = \{(...000300...), (...002300...)\},$$
$$f^2(\{\varphi\}) = f\beta g\gamma(\{\varphi\}) = \{(...000050...), (...002450...)\}.$$

Applying again $\beta\gamma$, for $l = 1$, i.e., $V = \{0\}$, we are done as $\eta_1 = (...000450...) \in \beta g^2\gamma(\{\varphi\})$, but $\eta_1 \notin f^2(\{\varphi\})$. For $l > 1$ this does not change anything. Furthermore

$$f^3(\{\varphi\}) = f\beta g^2\gamma(\{\varphi\}) = \{\overline{0}, (..002060...)\},$$

and by induction it is clear that in general application of f^{l-1} to $f^2(\{\varphi\})$ leads to $l-1$ zeros between 2 and 6 in the second state, and that

$$f^{l-1+2}(\{\varphi\}) = f\beta g^l\gamma(\{\varphi\}) = \{\overline{0}, (...002\underbrace{0...0}_{l-1}60...)\}.$$

So we find that e.g. $\eta_2 = (...0020...) \in \beta(g^{l+1}\gamma(\{\varphi\}))$, but $\eta_2 \notin f^{l+1}(\{\varphi\})$. □

This example also shows that there is no minimal pattern SA in general, as there are cases where for all pattern SA a bigger choice of V leads to a strictly smaller SA. In turn the SA for $V = \{0\}$ is maximal.

We note that similar considerations will become important again in the analysis of cellular probabilistic automata in Sec. 4.3.2.

3.2.4 Subsolution Supersystems

In this section we suggest the SS of minimal and maximal subsolutions to analyze a CNDA $(\mathbb{Z}, U, E, f_0)$ and relate it to the pattern SA. It is assumed in the whole section that E is a totally ordered set with order $\leq$, which is necessary for the method. Note that this order induces a partial order on E^U: $u \leq u'$ if $u(j) \leq u'(j)$ for all $j \in U$, where $u, u' \in E^U$.

3.2.4.1 Minimal and Maximal Subsolutions

Definition 3.33. For a CNDA $(\mathbb{Z}, U, E, f_0)$ with global function f, let

$$\underline{f}_0 : E^U \to E, \qquad u \mapsto \min f_0(u),$$
$$\overline{f}_0 : E^U \to E, \qquad u \mapsto \max f_0(u).$$

We call the deterministic CA $(\mathbb{Z}, U, E, \underline{f}_0)$ *minimal* and $(\mathbb{Z}, U, E, \overline{f}_0)$ *maximal CA* and denote their global functions by $\underline{f}$ and $\overline{f}$, respectively. We set further $F : \mathcal{K} \to \mathcal{K}$ by

$$S \to F(S) = \{\varphi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z}\, \exists \psi_1, \psi_2 \in S : \underline{f}(\psi_1)(i) \leq \varphi(i) \leq \overline{f}(\psi_2)(i)\}.$$

As usual we first check the definition and then show how F can be used to analyze the CNDA's long-term-behavior.

Lemma 3.34. F is closed and continuous.

Proof
Claim i) F is closed.

We show that $F(S)$ is closed, if $S \in \mathcal{K}$ is closed. Let $S \in \mathcal{K}$ and $(\varphi_n)_{n\in\mathbb{N}}$ be a sequence in $F(S)$ s.t. $\varphi_n \to \varphi \in E^{\mathbb{Z}}$ as $n \to \infty$. As $E^{\mathbb{Z}}$ is a metric space it is enough to show that $\varphi \in F(S)$.

For all $i \in \mathbb{Z}$ there is $N \in \mathbb{N}$ such that $\varphi(i) = \varphi_N(i)$. Per definition there are $\psi_1, \psi_2 \in E^{\mathbb{Z}}$ such that $\underline{f}(\psi_1)(i) \leq \varphi_N(i) \leq \overline{f}(\psi_2)(i)$. So for all $i \in \mathbb{Z}$ there are $\psi_1, \psi_2 \in S$ such that $\underline{f}(\psi_1)(i) \leq \varphi(i) \leq \overline{f}(\psi_2)(i)$ and therefore $\varphi \in F(S)$.

Claim ii) F is continuous.

We show that for each $\epsilon > 0$ there is a $\delta > 0$ such that $d_H(F(S), F(T)) < \epsilon$ if $d_H(S, T) < \delta$ for $S, T \in \mathcal{K}$.

Let $\epsilon > 0$ and choose $\delta = \epsilon$ and $k \in \mathbb{N}$ such that $\frac{1}{1+k} < \delta$. For $S, T \in \mathcal{K}$ with $d_H(S, T) < \delta$ we find that $S(i) = T(i)$ and hence $F(S)(i) = F(T)(i)$ for $i \in [-k, ..., k]$, and therefore $d_H(F(S), F(T)) < \epsilon$. □

Proposition 3.35. $(\mathcal{K}, F)$ is a SS of the CNDA $(\mathbb{Z}, U, E, f_0)$ with global function f.

Proof As both dynamical systems are defined on $\mathcal{K}$, it is enough to show that for all $S \in \mathcal{K}$ it holds that $f(S) \subseteq F(S)$.

Let $\varphi \in f(S)$. Then there exists $\psi \in S$ sucht that $\varphi(i) \in f_0(\psi|_{i+U})$ for all $i \in \mathbb{Z}$. Therefore for all $i \in \mathbb{Z}$ we can choose $\psi_1 = \psi_2 = \psi \in S$ such that $\underline{f}_0(\psi|_{i+U}) \leq \varphi(i) \leq \overline{f}_0(\psi|_{i+u})$, and so $\varphi \in F(S)$. □

We refer to $(\mathcal{K}, F)$ as the subsolution SS. Recall that we gain knowledge about $\omega_f(S)$ with Prop. 3.11 if we know how to determine $\omega_F(S)$ for $S \in \mathcal{K}$. It turns out that the latter is easy in the case of monotonic CNDA.

Definition 3.36. A deterministic CA $(\mathbb{Z}, U, E, f_0)$ with global function f is *monotonic*, if
i) $\varphi(i) \leq f(\varphi)(i)$ for all $\varphi \in E^{\mathbb{Z}}$ and all $i \in \mathbb{Z}$,
ii) $\varphi|_U \leq \psi|_U \Rightarrow f_0(\varphi|_U) \leq f_0(\psi|_U)$ for all $\varphi, \psi \in E^{\mathbb{Z}}$.
A CNDA is called *monotonic* if the according minimal and maximal CA are monotonic.

For a CA $(\mathbb{Z}, U, E, f_0)$ the two monotonicity conditions are independent, as the following two examples show.

Example for i) $\wedge\neg$ ii): consider $U = \{0\}, E = \{1, 2, 3, 4\}$ and f_0 given by

$\varphi(0)$	$f_0(\varphi(0))$
1	4
2	3
3	3
4	4

Then i) is true. Take e.g. $\varphi, \psi \in E^{\mathbb{Z}}$ defined by $\varphi(i) = 1, \psi(i) = 2$ for all $i \in \mathbb{Z}$. Then $f(\varphi)(i) = 4 > 3 = f(\psi)(i)$, and therefore ii) is not fulfilled.

Example for ii) $\wedge\neg$ i): consider $U = \{0\}, E = \{0, 1, 2\}$ and f_0 given by

$\varphi(0)$	$f_0(\varphi(0))$
0	0
1	0
2	1.

Then ii) is fulfilled, but take $\varphi \in E^{\mathbb{Z}}$ with $\varphi(i) = 1$ for all $i \in \mathbb{Z}$. Then $f(\varphi)(i) = 0 < \varphi(i)$ for all $i \in \mathbb{Z}$.

Lemma 3.37. The ω-limit set of $\eta \in E^{\mathbb{Z}}$ under a monotonic deterministic CA $(\mathbb{Z}, U, E, f_0)$ with global function f contains only one element, i.e., $\lim_{n\to\infty} f^n(\eta)$ exists.

Proof Let $\varphi, \psi \in \omega_f(\eta)$. Then there exist strictly monotonically increasing sequences $(n_i)_{i\in\mathbb{N}}$ and $(m_j)_{j\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$\lim_{i\to\infty} f^{n_i}(\eta) = \varphi, \qquad \lim_{j\to\infty} f^{m_j}(\eta) = \psi$$

in the Cantor topology. Choose subsequences $(\tilde{n}_k)_{k\in\mathbb{N}}$ and $(\tilde{m}_k)_{k\in\mathbb{N}}$ such that $\tilde{n}_k \leq \tilde{m}_k$ for all $k \in \mathbb{N}$. Then we have because of monotonicity property i) for all $l \in \mathbb{Z}$

$$f^{\tilde{n}_k}(\eta)(l) \leq f^{\tilde{m}_k}(\eta)(l).$$

Since this holds for all $k \in \mathbb{N}$, the relation is also true in the limit $k \to \infty$: $\varphi(l) \leq \psi(l)$. Interchanging the roles of φ and ψ additionally yields $\varphi(l) \geq \psi(l)$ and therefore $\varphi = \psi$. □

The property transfers to monotonic CNDA, such that ω_F-limit sets of deterministic states contain only one element and can be determined by the corresponding $\omega_{\underline{f}}$- and $\omega_{\overline{f}}$-limit sets.

Lemma 3.38. Let the CNDA $(\mathbb{Z}, U, E, f_0)$ be monotonic and $\{\varphi\} \in \mathcal{K}$. Then for all $k \in \mathbb{N}$

$$F^k(\{\varphi\}) = \{\psi \in E^Z \,|\, \forall i \in \mathbb{Z} : \underline{\mathrm{f}}^k(\varphi)(i) \leq \psi(i) \leq \overline{f}^k(\varphi)(i)\}.$$

Proof We prove the claim by induction. Per definition it is clear that

$$F(\{\varphi\}) = \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \underline{f}(\varphi)(i) \leq \psi(i) \leq \overline{f}(\varphi)(i)\}.$$

Let $k \in \mathbb{N}$. We assume that

$$F^{k-1}(\{\varphi\}) = \{\psi \in E^Z \,|\, \forall i \in \mathbb{Z} : \underline{\mathrm{f}}^{k-1}(\varphi)(i) \leq \psi(i) \leq \overline{f}^{k-1}(\varphi)(i)\},$$

let $\psi \in F^{k-1}(\{\varphi\})$ and fix $i \in \mathbb{Z}$. Because $\underline{\mathrm{f}}^{k-1}(\varphi)(j) \leq \psi(j)$ for all $j \in i + U$, from monotonicity property ii) it follows that $\underline{\mathrm{f}}^k(\varphi)(i) \leq \underline{\mathrm{f}}(\psi)(i)$. Similar reasoning holds for the upper bound, and so

$$\begin{aligned} F^k(\{\varphi\}) &= \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} \exists \psi_1, \psi_2 \in F^{k-1}(\{\varphi\}) : \underline{\mathrm{f}}(\psi_1)(i) \leq \psi(i) \leq \overline{f}(\psi_2)(i)\} \\ &= \{\psi \in E^Z \,|\, \forall i \in \mathbb{Z} : \underline{\mathrm{f}}^k(\varphi)(i) \leq \psi(i) \leq \overline{f}^k(\varphi)(i)\}. \end{aligned}$$

□

Proposition 3.39. Let the CNDA $(\mathbb{Z}, U, E, f_0)$ be monotonic and $\{\varphi\} \in \mathcal{K}$. Then

$$\omega_F(\{\varphi\}) = \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \underline{\varphi}(i) \leq \psi(i) \leq \overline{\varphi}(i)\},$$

where $\omega_{\underline{f}}(\varphi) = \{\underline{\varphi}\}$ and $\omega_{\overline{f}}(\varphi) = \{\overline{\varphi}\}$.

Proof For a monotonic CNDA by definition the according mininimal and maximal CA are monotonic, and hence Lm. 3.37 guarantees that the ω-limit sets are well-defined. We denote $T := \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \underline{\varphi}(i) \leq \psi(i) \leq \overline{\varphi}(i)\}$ and show that $F^{m_j}(\{\varphi\})$ converges to T in the Hausdorff topology for all sequences (m_j). As $\mathcal{K}$ is complete, necessarily $T \in \mathcal{K}$.

As E is finite, for all $i \in \mathbb{Z}$ there exists $n_i \in \mathbb{N}$ such that $\underline{f}^n(\varphi)(i) = \underline{\varphi}(i)$ and $\overline{f}^n(\varphi)(i) = \overline{\varphi}(i)$ for all $n \geq n_i$.

Let $(m_j)_{j\in\mathbb{N}}$ be a strictly monotonically increasing sequence in $\mathbb{N}$ and $\epsilon > 0$. Choose $k \in \mathbb{N}$ such that $0 < \frac{1}{1+k} < \epsilon$ and choose $N_0 = \max_{i\in\{-k,\dots,k\}}\{n_i\}$. With Lm. 3.38 for all $j \geq N_0$, for all $\chi \in F^{m_j}(\{\varphi\})$ there is $\psi \in T$ with $\chi(i) = \psi(i)$ for $i \in \{-k, ..., k\}$ and vice versa, and therefore $d_H(F^{m_j}(\{\varphi\}), T) < \frac{1}{1+k} < \epsilon$. □

This proposition can be extended to random states. If $\underline{f}$ and $\overline{f}$ are monotonic, ω_F-limit-sets only contain one element and can be determined by the $\omega_{\underline{f}}$- and $\omega_{\overline{f}}$-limit sets of appropriate states.

Lemma 3.40. Let the CNDA $(\mathbb{Z}, U, E, f_0)$ be monotonic, $S \in \mathcal{K}$, $\varphi_1(i) = \min_{\varphi\in S}\{\underline{f}(\varphi)(i)\}$ and $\varphi_2(i) = \max_{\varphi\in S}\{\overline{f}(\varphi)(i)\}$. Then

$$\omega_F(S) = \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \underline{\varphi}(i) \leq \psi(i) \leq \overline{\varphi}(i)\},$$

where $\omega_{\underline{f}}(\varphi_1) = \{\underline{\varphi}\}$ and $\omega_{\overline{f}}(\varphi_2) = \{\overline{\varphi}\}$.

Proof Let $S \in \mathcal{K}$. Recalling that F is a function that acts on states only single-site-wise it becomes clear that $F(S)$ does not possess any correlations. This guides us to the definition of $\varphi_1, \varphi_2 \in E^{\mathbb{Z}}$ and encourages us to use the same proof idea as in the former proposition. Here one uses that in analogy to Lm. 3.38 for all $k \in \mathbb{N}$

$$F^k(S) = \{\psi \in E^{\mathbb{Z}} | \forall i \in \mathbb{Z} : \underline{f}^{k-1}(\varphi_1)(i) \leq \psi(i) \leq \overline{f}^{k-1}(\varphi_2)(i)\}$$

and that here for all $i \in \mathbb{Z}$ there exists $1 \leq n_i \in \mathbb{N}$ such that $\underline{f}^{n-1}(\varphi_1)(i) = \underline{\varphi}(i)$ and $\overline{f}^{n-1}(\varphi_2)(i) = \overline{\varphi}(i)$ for all $n \geq n_i$.

By the same argument as above we see that $F^{m_j}(S)$ converges to the suggested set for all strictly monotonically increasing sequences $\{m_j\}_{j\in\mathbb{N}}$ in the Hausdorff topology. □

3.2.4.2 The Subsolution Supersystem and Pattern Superautomata

It will be shown that for trajectories starting with deterministic states the subsolution SS of a CNDA $(\mathbb{Z}, U, E, f_0)$ with global function f can also be interpreted as originating from a SA, which often coincides with the maximal pattern SA.

In this section we choose $V = \{0\}$ and according de Bruijn states $X_{dB(e)}$ and maps β and γ. Here g is the global function of the corresponding maximal pattern SA, and F is defined as in Def. 3.33.

Lemma 3.41. Let $G : X_{dBe} \to X_{dBe}$ be given by $\Phi \to G(\Phi) = \gamma F \beta(\Phi)$. For all $S \in \mathcal{K}$ and all $\Phi \in X_{dBe}$ it holds that

i) $\beta\gamma F(S) = F(S)$,

ii) $\beta G(\Phi) = F\beta(\Phi)$, i.e., the dynamical system (X_{dBe}, G) can be embedded in $(\mathcal{K}, F)$.

Proof Claim i) According to Lm. 3.25 $\beta\gamma(S) \supseteq S$ for all $S \in \mathcal{K}$, so we just have to prove $\beta\gamma F(S) \subseteq F(S)$.

Let $\varphi \in \beta\gamma F(S)$. Then we know that $\varphi(i) \in \gamma F(S)(i)$ for all $i \in \mathbb{Z}$. According to the defintion of γ this implies that there is $\psi \in F(S)$ such that $\varphi(i) = \psi(i)$. As $\psi \in F(S)$, we know that there are $\psi_1, \psi_2 \in S$ such that $\underline{f}(\psi_1)(i) \leq \psi(i) = \varphi(i) \leq \overline{f}(\psi_2)(i)$, and thus $\varphi \in F(S)$.

Claim ii) follows from *i)*: $\beta G = \beta\gamma F\beta = F\beta$. □

Proposition 3.42. Using β and γ, G can be extended to a SA on X_{dB} of the CNDA $(\mathbb{Z}, U, E, f_0)$ which is bigger than all pattern SA, and $\beta(\omega_G(\gamma(\{\varphi\}))) = \omega_F(\{\varphi\})$ for all $\varphi \in E^{\mathbb{Z}}$.

Proof First, we show that G can be extended to a CA. Let $\Phi \in X_{dBe}$ and $i \in \mathbb{Z}$. We know that

$$\begin{aligned} G(\Phi)(i) &= \{\varphi|_{i+V} \,|\, \varphi \in F\beta(\Phi)\} \\ &= \{v \in E^V \,|\, \forall j \in V \exists \psi_1, \psi_2 \in \beta(\Phi) : \underline{f}(\psi_1)(i+j) \leq u(j) \leq \overline{f}(\psi_2)(i+j)\} \end{aligned}$$

As $\Phi \in X_{dbe}$, we only need to know $\Phi|_{i+U}$ in order to determine this set. The extension to a CA on X_{dB} can be achieved in analogy to the argumentation for Cor. 3.29.

It is clear that this CA is a SA bigger than all pattern SA, if we show that it covers the maximal pattern SA. With Prop. 3.35 it holds that $g(\Phi) = \gamma f \beta(\Phi) \subseteq \gamma F \beta(\Phi) = G(\Phi)$ for all $\Phi \in (\mathcal{P}(E))^{\mathbb{Z}}$. Therefore $\beta g^k \gamma(S) = \beta \subseteq \beta G^k \gamma(S)$ for all $k \in \mathbb{N}$ and all $S \in \mathcal{K}$.

It remains to show that $\omega_F(\{\varphi\}) \subseteq \beta(\omega_G(\gamma(\{\varphi\}))$ for all $\varphi \in E^{\mathbb{Z}}$, the other inclusion follows analogously.

Like in Prop. 3.11 we see that for $T \in \omega_F(\{\varphi\})$ there is a sequence $(n_l)_{l\in\mathbb{N}}$ in $\mathbb{N}$ and $\Phi \in \omega_G(\gamma(\{\varphi\}))$ such that $T = \lim_{l\to\infty} F^{n_l}(\{\varphi\})$ and $\Phi = \lim_{l\to\infty} G^{n_l}\gamma(\{\varphi\})$. Since $\beta\gamma(\{\varphi\}) = \{\varphi\}$ and with Lm. 3.41 i) for all $l \in \mathbb{N}$

$$F^{n_l}(\{\varphi\}) = (\beta\gamma F)^{n_l}(\{\varphi\}) = \beta(\gamma F\beta)^{n_l}\gamma(\{\varphi\}) = \beta G^{n_l}\gamma(\{\varphi\}),$$

and $\lim_{l\to\infty} \beta G^{n_l}\gamma(\{\varphi\}) = \beta(\Phi)$ because β is continuous. As the limit point is unique therefore $T = \beta(\Phi) \in \beta(\omega_G(\gamma(\{\varphi\}))$. □

In practice the local function often maps intervals to intervals in some sense. This property already ensures that the subsolution SA coincides with the maximal pattern SA for monotonic CNDA on trajectories starting with a deterministic state.

Definition 3.43. Let $W = \{-m, ..., n\}$ for $m, n \in \mathbb{Z}$. Then $I \subseteq E^W$ is an *interval on* E^W if for all $i_1, i_2 \in I$ with $i_1 < i_2$ there is no $w \in E^W \backslash I$ such that $i_1 < w < i_2$. A CNDA $(\mathbb{Z}, E, U, f_0)$ *maps intervals to intervals*, if $\cup_{i\in I} f_0(i)$ is an interval on $E^{\{0\}} = E$ for all intervals I on E^U.

Lemma 3.44. If the CNDA is monotonic and maps intervals to intervals, it holds that for all $\varphi \in E^{\mathbb{Z}}$ and all $k \in \mathbb{N}$

i) $F^k(\{\varphi\}) = (\beta\gamma f)^k(\{\varphi\})$,

ii) $\beta g^k \gamma(\{\varphi\}) = \beta G^k \gamma(\{\varphi\})$.

Proof Claim i) Let $\varphi \in E^{\mathbb{Z}}$. Using Lm. 3.38 it is enough to show that for all $k \in \mathbb{N}$

$$(\beta\gamma f)^k(\{\varphi\}) = \{\psi \in E^Z \,|\, \forall i \in \mathbb{Z} : \underline{f}^k(\varphi)(i) \leq \psi(i) \leq \overline{f}^k(\varphi)(i)\}.$$

We use induction and start with $k = 1$: since $\gamma f(\{\varphi\})(i) = f_0(\varphi|_{i+U})$ for all $i \in \mathbb{Z}$ and since $f_0(\varphi|_{i+U})$ is an interval

$$\begin{aligned}\beta\gamma f(\{\varphi\}) &= \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \psi(i) \in f_0(\varphi|_{i+U})\} \\ &= \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \underline{f}(\varphi)(i) \leq \psi(i) \leq \overline{f}(\varphi)(i)\}.\end{aligned}$$

Let $k \in \mathbb{N}$, $k > 1$. Then

$$\begin{aligned}\beta\gamma f(\beta\gamma f)^{k-1}(\{\varphi\}) &= \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} \exists \tau \in f(\beta\gamma f)^{k-1}(\{\varphi\}) : \psi(i) = \tau(i)\} \\ &= \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} \exists \eta \in (\beta\gamma f)^{k-1}(\{\varphi\}) : \psi(i) \in f_0(\eta|_{i+U})\}.\end{aligned}$$

Fix $i \in \mathbb{Z}$ and $\eta \in (\beta\gamma f)^{k-1}(\{\varphi\})$. Assuming that the claim is true for $k-1$, for all $j \in i + U$ then $\underline{f}^{k-1}(\varphi)(j) \leq \eta(j)$, and because of monotonicity hence $\underline{f}^k(\varphi)(i) \leq \underline{f}_0(\eta|_{i+U})$. An analogous statement holds for the upper bound. As $\{\eta|_{i+U} \,|\, \eta \in (\beta\gamma f)^{k-1}(\{\varphi\})\}$ is an interval on E^U and as the CNDA maps intervals to intervals, all together

$$\beta\gamma f(\beta\gamma f)^{k-1}(\{\varphi\}) = \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \underline{f}^k(\varphi) \leq \psi(i) \leq \overline{f}^k(\varphi)\}.$$

Claim ii) is a consequence of i) and of Lm. 3.41: for $k \in \mathbb{N}$ and $\varphi \in E^{\mathbb{Z}}$

$$\begin{aligned}\beta g^k \gamma(\{\varphi\}) &= \beta(\gamma f \beta)^k \gamma(\{\varphi\}) = (\beta\gamma f)^k \beta\gamma(\{\varphi\}) = (\beta\gamma f)^k(\{\varphi\}) = F^k(\{\varphi\}) \\ &= (\beta\gamma F)^k \beta\gamma(\{\varphi\}) = \beta G^k \gamma(\{\varphi\}).\end{aligned}$$

□

3.3 From PDEs to Cellular Non-Deterministic Automata

We suggest a generic procedure to construct a CNDA that describes the dynamics of a given PDE; see Fig. 3.1a. This procedure is applied in two examples, the Fisher-KPP equation [68] and the reaction random walk [80]. Both are central in the theory of reaction-diffusion equations [157].

3.3.1 Generic Procedure

Two steps are necessary in order to describe the dynamics of a PDE by a CNDA. First, time, space and state space have to be discretized, and second, all transitions between the discrete states with respect to a suitable time discretization have to be collected. Here we use a method of lines to discretize space, and then the dynamics to flexibly discretize time and state space at the same time. We note that in CPA in Chap. 4 we will only be interested in simpler state space discretizations and therefore discretize space together with time in one step with a finite difference scheme. Afterwards there the state space is discretized independently of the dynamics.

3.3.1.1 Discretization

We start by defining a partition of $\Omega \subseteq V \subseteq \mathbb{R}^n$ for a time-continuous dynamical system $(\mathbb{R}, V, \Phi_t)$ as in symbolic dynamics [181].

Definition 3.45. A *partition* or *coding* E of Ω is a finite collection of disjoint sets $\{\Omega_e\}_{e\in E}$ whose union is Ω. We call $e \in E$ the *symbol* of *coding domain* Ω_e, and the *coding map* is the function $T : \Omega \mapsto E$, where $T(v) = e$ if $v \in \Omega_e$.

Let now $\Omega \subseteq V$ be compact and positively invariant with respect to Φ_t. For many ODE $\partial_t v = h(v)$ there is the possibility to construct a partition of Ω that is compatible with the flow in the sense that $T(v_1) = T(v_2)$ implies $T(\Phi_p(v_1)) = T(\Phi_p(v_2))$ for a given $p \in \mathbb{R}$. Then there is a deterministic function $\tilde{f} : E \to E, e \mapsto \tilde{f}(e) = T\Phi_P T^{-1}(e)$ for $P = kp$, $k \in \mathbb{N}$, which can be iterated to resemble the dynamics of the continuous system. Such a partition may often be constructed by integrating the dynamics of a hypersurface or the boundary $\partial\Omega$ of Ω:

$$\begin{aligned}\Omega_{e_0} &= \{\Phi_t(\partial\Omega) \,|\, t \in [0, p)\},\\ \Omega_{e_{k+1}} &= \{\Phi_t(\Phi_{kp}\partial\Omega) \,|\, t \in [0, p)\}, \quad k = 1, ..., K-1,\\ \Omega_{e_K} &= \Omega \setminus \cup_{k=0}^{K-1}\Omega_{e_k}.\end{aligned}$$

Ω_{e_k} may also be split into a finite number of subsets along the trajectories of a finite number of points in $\partial\Omega$ to get a better resolution.

If one chooses a partition without that special property, a coding domain is mapped to several coding domains under h. The result is a non-deterministic $\tilde{f}$ on the symbolic level; see Sec. 3.3.3.2 for a pronounced example.

Now we move from considerations about a single ODE to a translationally invariant system of ODE

$$\partial_t v_i = h(v_{i-1}, v_i, v_{i+1}), \qquad i \in \mathbb{Z},$$

on dynamically positively invariant state space $\hat{V} = \Omega^{\mathbb{Z}}$. Our central example is the application of the method of lines on PDEs of the form

$$\partial_t v = \tilde{h}(v, \partial_x v, \partial_{xx} v).$$

Note that according to Tychonoff's theorem [180] $\hat{V}$ is compact if Ω is. We assume in the following that the solution operator $\hat{\Phi} : \mathbb{R}_+ \times \hat{V} \to \hat{V}$ of this dynamical system is well-defined.

Any coding map $T : \Omega \to E$ of Ω induces a map

$$\hat{T} : \hat{V} \to E^{\mathbb{Z}}, v \mapsto \hat{T}(v) \text{ with } (\hat{T}(v))_i = T(v_i) \text{ for } i \in \mathbb{Z}$$

that we can use for state space discretization of our system of ODE. We observe that a partition of Ω which is compatible with h at $i \in \mathbb{Z}$ for given v_{i-1} and v_{i+1} is not compatible any more if the latter are changed. Therefore it is in general impossible to construct a partition which is compatible with h. If the contribution of v_{i-1} and v_{i+1} to h is separable from and smaller than the one of v_i, the best we can do is to choose a partition with respect to the dominant process and consider the other ones as perturbations.

However, in general we get non-deterministic symbolic dynamics and different transitions $\tilde{f}$ for each configuration of states in neighboring sites. To account for this phenomenon we pass from $\tilde{f}$ to the local function f_0 of a CNDA.

3.3.1.2 Determination of Possible Transitions

We do not want our symbolic description to miss any trajectory of the system of coupled ODEs and hence have to collect all possible image symbols for any initial neighborhood configuration.

Definition 3.46. Let $T : \Omega \to E$ be a coding map, select $P > 0$ and define $U := \{-1, 0, 1\}$. A CNDA $(\mathbb{Z}, U, E, f_0)$ is given as follows: $\tilde{e}_0 \in f_0((e_{-1}, e_0, e_1))$ if there is a state $v \in \hat{V}$ such that $T(v_j) = e_j$ for $j \in U$ and $T((\hat{\Phi}_P(v))_0) = \tilde{e}_0$.

The CNDA $(\mathbb{Z}, U, E, f_0)$ with global function f covers the dynamic behavior of the time-discretized ODE system everywhere except for on a set of measure 0.

Proposition 3.47. Let $v \in \hat{V}$. If $\omega_{\Phi_P}(v) \cap \partial\Omega_e = \emptyset$ for all $e \in E$, then for all $w \in \omega_{\Phi_P}(v)$ there is $S \in \omega_f(\{\hat{T}(v)\})$ such that $\hat{T}(w) \in S$.

Proof The lemma follows in analogy to Prop. 3.11. Let $v \in \hat{V}$ and $w \in \omega_{\Phi_P}(u)$. Then there is a strictly monotonously increasing sequence (n_k) in $\mathbb{N}$ such that $\lim_{k\to\infty} \hat{\Phi}_P^{n_k}(v) = w$. As $\mathcal{K}$ is compact, there is a converging subsequence $f^{n_l}\hat{T}(v)$ of $f^{n_k}\hat{T}(v)$, and we set $S = \lim_{l\to\infty} f^{n_l}\hat{T}(u)$. By construction it holds that $\hat{T}\Phi_P^{n_l}(v) \in f^{n_l}\hat{T}(v)$ for all $l \in \mathbb{N}$, and as $\omega_{\Phi_P}(v) \cap \partial\Omega_e = \emptyset$ for all $e \in E$, there is $N \in \mathbb{N}$ such that $\hat{T}\Phi_P^{n_l}(v) = \hat{T}(w)$ for $l > N$. Therefore with Thm. 2.6 ii)

$$\{\hat{T}(w)\} = \{\hat{T}(\lim_{l\to\infty} \Phi_P^{n_l}(v))\} = \{\lim_{l\to\infty} \hat{T}\Phi_P^{n_l}(v)\} \subseteq \lim_{l\to\infty} f^{n_l}\hat{T}(v) = S.$$

□

In conclusion two steps are required in our CNDA construction. First, time, space and state space have to be discretized, and then all possible transitions between the discretized portions of phase space have to be determined. We point out that in practice one has to rely on approximating algorithms for the second task. They cannot guarantee that all transitions are found. On the other hand, depending on the algorithm one may often find spurious transitions which leads to a CNDA that is a SS to the minimal CNDA introduced in Def. 3.46; see Fig. 3.1a.

3.3.2 Application I: Fisher-KPP Equation

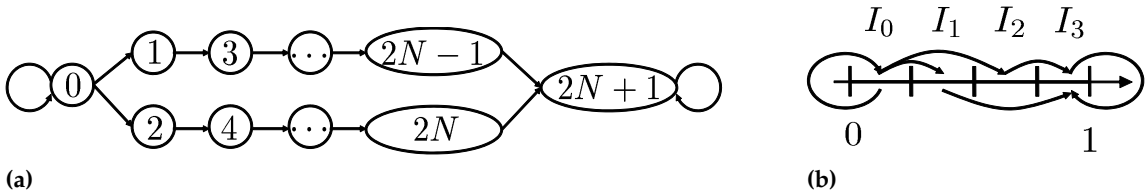
Consider the Fisher-KPP equation [68]

$$\partial_t v - \partial_{xx} v = v(1 - v)$$

for $v \in \Omega = [0, 1]$ and apply the method of lines to it. This leads to a system of ODE on $\hat{V} = \Omega^{\mathbb{Z}}$,

$$\partial_t v_i = v_i(1 - v_i) + \varepsilon(v_{i-1} - 2v_i + v_{i+1}) =: h(v_{i-1}, v_i, v_{i+1}),$$

where $\epsilon = (\Delta x)^{-2}$. We note that Ω is compact and show that $\hat{V}$ is positively invariant. Let $v_{i,R} = 1, v_{i,L} = 0, v_{i-1} = r, v_{i+1} = s$ for $r, s \in \Omega$ and n the normal vector of Ω. So $n(v_{i,R}) = 1$ and $n(v_{i,L}) = -1$, and hence:



(a) (b)

Figure 3.6: (a) $\tilde{f}$ as a transition graph: node i is mapped to node j if there is a directed edge from i to j. (b) The underlying partition of $\Omega = [0, 1]$ for four symbols. The arrows indicate $\tilde{f}$.

$$\langle h(v_{i-1}, v_{i,R}, v_{i+1}), n(v_{i,R})\rangle = \epsilon(r + s - 2) \leq 0,$$
$$\langle h(v_{i-1}, v_{i,L}, v_{i+1}), n(v_{i,L})\rangle = -\epsilon(r + s) \leq 0.$$

Therefore Ω is, independently of v_{i-1} and v_{i+1}, positively invariant under h, and so is $\hat{V}$ under the system of ODE.

3.3.2.1 State Space Discretization

We choose ε small by considering a low spatial resolution and thus construct the partition of Ω according to the dominant process $\partial_t w = w(1 - w)$. There are two stationary points in this system, the unstable $w_u = 0$ and the stable $w_s = 1$. We take $w_0 = 1/2$ and iterate forward and backward in time steps of p,

$$w_{k+1} = \frac{w_k e^p}{1 + w_k(e^p - 1)}, \qquad k \in \mathbb{Z}$$

We aim at a finite number of symbols and use 0, w_k ($k = -N, .., N$) and 1 as boundaries of coding domains. This is, we define $2N + 2$ intervals

$$I_0 = [0, w_{-N}), \quad I_k = [w_{-N+k-1}, w_{-N+k}) \quad \text{for } k = 1, .., 2N, \qquad I_{2N+1} = [w_N, 1],$$

name them by their index and obtain $E = \{0, ..., 2N + 1\}$. The resulting $\tilde{f}$ is depicted in Fig. 3.6a for $P = 2p$. It is already non-deterministic at the unstable w_u. We use $P = 2p$, because then small perturbations from the diffusion term are not able to break the system's strict monotonicity. The partition with this dynamics is kind of robust.

In order to demonstrate this construction, we choose from now on the parameters

$$p = 1.0, \quad P = 2.0, \quad N = 1, \quad \epsilon = 0.085,$$

and find the partition $E = \{0, ..., 3\}$ with $w_{-1} = 2.6894, w_0 = 1/2$ and $w_1 = 0.73106$; see Fig. 3.6b.

3.3.2.2 Determination of Possible Transitions

We utilize the monotonicity of the system in order to construct a CNDA. We do not have any restrictions for the state at sites $i = -2$ and $i = 2$, although these sites will also influence the dynamics of the state at site $i = 0$. This lack of information forces us to work with sub- and supersolutions. So we cannot exclude that our CNDA contains spurious trajectories.

Proposition 3.48. Consider $\partial_t v_i = h(v_{i-1}, v_i, v_{i+1})$, $v \in \hat{V}$ and $e_j \in E$ such that $T(v_j) = e_j$ for $j \in U$. Let $\Omega_{e_{-1}} = [w_L^-, w_L^+)$, $\Omega_{e_0} = [w_M^-, w_M^+)$, and $\Omega_{e_1} = [w_R^-, w_R^+)$ and

$$\partial_t \overline{u}_{-1} = h(1, \overline{u}_{-1}, \overline{u}_0), \quad \partial_t \overline{u}_0 = h(\overline{u}_{-1}, \overline{u}_0, \overline{u}_1), \quad \partial_t \overline{u}_1 = h(\overline{u}_0, \overline{u}_1, 1),$$

with $\overline{u}_{-1}(0) = w_L^+$, $\overline{u}_0(0) = w_M^+$, and $\overline{u}_1(0) = w_R^+$. Then, $(\hat{\Phi}_t(v))_0 \leq \overline{u}_0(t)$.

Similarly, we define

$$\partial_t \underline{u}_{-1} = h(0, \underline{u}_{-1}, \underline{u}_0), \quad \partial_t \underline{u}_0 = h(\underline{u}_{-1}, \underline{u}_0, \underline{u}_1), \quad \partial_t \underline{u}_1 = h(\underline{u}_0; \underline{u}_1, 0),$$

with $\underline{u}_{-1}(0) = w_L^-$, $\underline{u}_0(0) = w_M^-$, and $\underline{u}_1(0) = w_R^-$. Then, $(\hat{\Phi}_t(v))_0 \geq \underline{u}_0(t)$.

Proof The claim is an immediate consequence of the monotonicity of h in v_{i-1} and in v_{i+1} for all $i \in \mathbb{Z}$. □

Using this proposition and numerical integration we are able to find a maximal interval $[\underline{u}_0(P), \overline{u}_0(P)]$ that contains $(\hat{\Phi}_P(v))_0$ for any initial conditions fulfilling $T(v_j) = e_j$ for $j \in U$. Picking all coding intervals with a non-trivial intersection with $[\underline{u}_0(P), \overline{u}_0(P)]$, we can determine the local function of a CNDA:

e_{-1}	e_0	e_1	$f_0((e_{-1}, e_0, e_1))$
0	0	0	$\{0,1,2,3\}$
0	0	1	$\{0,1,2,3\}$
0	0	2	$\{0,1,2,3\}$
0	0	3	$\{1,2,3\}$
1	0	0	$\{0,1,2,3\}$
1	0	1	$\{1,2,3\}$
1	0	2	$\{1,2,3\}$
1	0	3	$\{1,2,3\}$
2	0	0	$\{0,1,2,3\}$
2	0	1	$\{1,2,3\}$
2	0	2	$\{1,2,3\}$
2	0	3	$\{1,2,3\}$
3	0	0	$\{1,2,3\}$
3	0	1	$\{1,2,3\}$
3	0	2	$\{1,2,3\}$
3	0	3	$\{1,2,3\}$

e_{-1}	e_0	e_1	$f_0((e_{-1}, e_0, e_1))$
0	1	0	$\{2,3\}$
0	1	1	$\{2,3\}$
0	1	2	$\{2,3\}$
0	1	3	$\{2,3\}$
1	1	0	$\{2,3\}$
1	1	1	$\{2,3\}$
1	1	2	$\{3\}$
1	1	3	$\{3\}$
2	1	0	$\{2,3\}$
2	1	1	$\{3\}$
2	1	2	$\{3\}$
2	1	3	$\{3\}$
3	1	0	$\{2,3\}$
3	1	1	$\{3\}$
3	1	2	$\{3\}$
3	1	3	$\{3\}$

and

$$f_0((*, 2, *)) = \{3\}, \qquad f_0((*, 3, *)) = \{3\},$$

where $*$ can be replaced by any symbol.

Note that although the assumption in Prop. 3.47 is violated by w_u and w_s, we still expect the CNDA to resemble the dynamics of the system of ODE. A look at the proposition's proof reveals that our stationary points on $\partial\Omega$ cannot lead to the difficulties that may appear for general non-empty intersections between ω-limit sets and boundaries of coding domains.

3.3.2.3 Subsolution Analysis of the CNDA

We note that the CNDA is monotonic and maps intervals to intervals; see Def. 3.36 and Def. 3.43. We use the subsolution method for analysis, which, according to Prop. 3.42 and Lm. 3.44, coincides with the maximal pattern SA when starting in a deterministic state.

We start with the minimal (deterministic) CA $(\mathbb{Z}, U, E, \underline{\mathrm{f}}_0)$ given by

e_{-1}	e_0	e_1	$\underline{\mathrm{f}}_0((e_{-1}, e_0, e_1))$
0	0	0,1,2	0
0,1,2	0	0	0
0	0	3	1
3	0	0	1
1,2,3	0	1,2,3	1
0	1	*	2
*	1	0	2

e_{-1}	e_0	e_1	$\underline{\mathrm{f}}_0((e_{-1}, e_0, e_1))$
1	1	1	2
1	1	2,3	3
2,3	1	1	3
2,3	1	2,3	3
*	2	*	3
*	3	*	3

and state some of its dynamic properties.

i) The minimal CA has resting state 0.

ii) $\overline{0}$ and $\overline{3}$ are stationary states.

iii) $010, 011, 110, 111, 012, 210, 112, 211$ and 212 are Garden of Eden patterns. This is, because a 3 is needed at a neighboring site to create a 1, and as 3 is mapped to itself a 1 can just appear next to a 3. As they do not have predecessors, states containing Garden of Eden patterns cannot appear in the global attractor.

iv) The global attractor of the CA just contains monotonic states. This is, because $\varphi = ...e_0e_1...e_{k+2}e_{k+3}...$, where $k \in \mathbb{N}$, $0 \leq e_0 < e_1 \leq 3$ and $3 \geq e_{k+2} > e_{k+3} \geq 0$ cannot arise under the CA's dynamics:
As $e_0, e_{k+3} \in \{0, 1, 2\}$, and as sites in elementary local state 1 and 2 have been in strictly smaller elementary states in the time step before, φ stems from a state ψ with $\psi(0) = \psi(k + 3) = 0$. ψ in turn eventually goes back to a $\chi \in E^{\mathbb{Z}}$ with $\chi(j) \in \{0, 1\}$ for $j \in \{0, ..., k + 3\}$, for which there exists at least one $i \in \{1, ..., k + 2\}$ with $\chi(i) = 1$. But states of this type contain a Garden of Eden pattern.

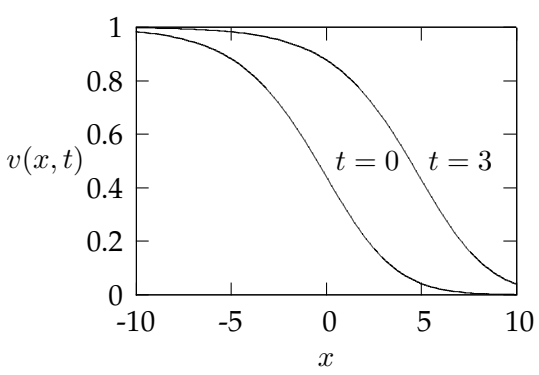
v) Non-constant states in the global attractor are the travelling wave states, i.e., states in

$$R := \{\varphi \in E^{\mathbb{Z}} \mid \exists j \in \mathbb{Z} \forall i < j < k : \text{either } \varphi(i) = 0,\ \varphi(k) = 3 \text{ or } \varphi(i) = 3,\ \varphi(k) = 0\};$$

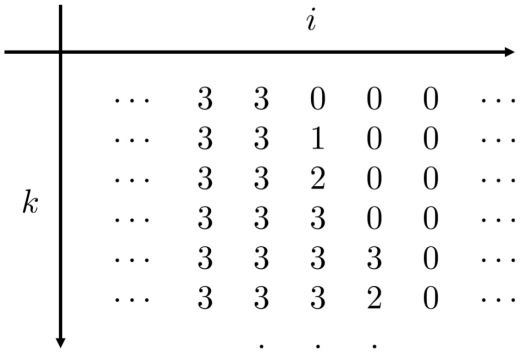
see Fig. 3.7b.
Any potential non-constant state in the global attractor is monotonic and hence arrives in $R \cup \{\overline{0}, \overline{3}\}$ after at most 2 iterations. As $R \cup \{\overline{0}, \overline{3}\}$ is positively invariant, monotonic states in $E^{\mathbb{Z}} \backslash (R \cup \{\overline{0}, \overline{3}\})$ cannot be reached again and hence cannot be contained in the global attractor. In turn R must be part of the global attractor as it is also negatively invariant.

Corollary 3.49. The global attractor of the minimal CA is $A = R \cup \{\overline{0}, \overline{3}\}$. Furthermore $\omega_{\underline{\mathrm{f}}}(\overline{0}) = \{\overline{0}\}$, and $\omega_{\underline{\mathrm{f}}}(\varphi) = \{\overline{3}\}$ for any $\varphi \in E^{\mathbb{Z}}, \varphi \neq \overline{0}$.

We turn now to the trivial maximal CA $(\mathbb{Z}, U, E, \overline{f}_0)$ given by



(a)



(b)

Figure 3.7: (a) Running front solution of the Fisher-KPP equation for speed $c = \frac{5}{\sqrt{6}}$ at time $t = 0$ and $t = 3$. (b) Running front solution of the minimal CA with speed $c_{min} = 0.57$. As usual i is the space and k the time variable.

e_{-1}	e_0	e_1	$\overline{f}_0((e_{-1}, e_0, e_1))$
*	*	*	$\overline{3}$.

Lemma 3.50. The global attractor of the maximal CA is $A = \{\overline{3}\}$. Furthermore $\omega_{\overline{f}}(\varphi) = \{\overline{3}\}$ for any $\varphi \in E^{\mathbb{Z}}$.

According to Prop. 3.11, 3.35 and 3.39 then

$$\omega_f(\{\overline{0}\}) \subseteq \omega_F(\{\overline{0}\}) = \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \overline{0}(i) \leq \psi(i) \leq \overline{3}(i)\} = E^{\mathbb{Z}}$$

and

$$\omega_f(\{\varphi\}) \subseteq \omega_F(\{\varphi\}) = \{\psi \in E^{\mathbb{Z}} \,|\, \forall i \in \mathbb{Z} : \overline{3}(i) \leq \psi(i) \leq \overline{3}(i)\} = \{\overline{3}\}$$

for $\varphi \in E^{\mathbb{Z}}, \varphi \neq \overline{0}$. As ω-limit sets on compact $\mathcal{K}$ are non-empty we even have equality in the last line.

So the simulation of the Fisher-KPP equation with a CNDA produces the expected results in the long term. It is also possible to extract information about attractors relatively easy with our means. This is remarkable as such an analysis is usually very difficult and technically involved [124]. It is interesting to note that the minimal deterministic CA even shows some of the transient behavior of the PDE, the travelling waves; see Fig. 3.7. For our parameters their speed can be calculated to be

$$c_{min} = \frac{\Delta x}{\Delta t} = \frac{\epsilon^{-0.5}}{3P} \approx 0.57.$$

The travelling wave speed of the maximal CA is infinitely large, and we can give a bound on the Fisher-KPP equation's travelling wave speed: $0.57 \leq c$. It is well known [68] that there are only PDE travelling wave solutions for $2 \leq c$, which shows that we are only approximating the true transient PDE behavior.

3.3.3 Application II: Reaction Random Walk

In this example we suggest two state space discretizations and resulting CNDA for a PDE on a 2-dimensional phase space. We only pave the way towards CNDA construction but do not go into the analysis, as we just want to point out the different constructions.

Consider the reaction random walk equation [80]

$$\partial_t \begin{pmatrix} v^+ \\ v^- \end{pmatrix} + c\partial_x \begin{pmatrix} v^+ \\ -v^- \end{pmatrix} = \begin{pmatrix} -\mu & \mu \\ \mu & -\mu \end{pmatrix} \begin{pmatrix} v^+ \\ v^- \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \end{pmatrix} (v^+ + v^-) - \gamma \begin{pmatrix} v^+ \\ v^- \end{pmatrix} (v^+ + v^-)$$

on $\Omega = [0, \frac{2b}{\gamma}]^2 \subseteq \mathbb{R}^2$ for $\mu, b, \gamma \geq 0$. With the notation $v_i = \begin{pmatrix} v_i^+ \\ v_i^- \end{pmatrix}$, application of the method of lines leads to a system of ODE on $\hat{V} = \Omega^{\mathbb{Z}}$

$$\partial_t v_i = h(v_{i-1}^+, v_i, v_{i+1}^-) = h_1(v_i) + h_2(v_{i-1}^+, v_i, v_{i+1}^-), \quad i \in \mathbb{Z}, \tag{3.1}$$

where

$$h_1(v_i) = \begin{pmatrix} (b-\mu)v_i^+ + (b+\mu)v_i^- - \gamma v_i^+(v_i^+ + v_i^-) \\ (b+\mu)v_i^+ + (b-\mu)v_i^- - \gamma v_i^-(v_i^+ + v_i^-) \end{pmatrix}, h_2(v_{i-1}^+, v_i, v_{i+1}^-) = \epsilon \begin{pmatrix} -(v_i^+ - v_{i-1}^+) \\ (v_{i+1}^- - v_i^-) \end{pmatrix}.$$

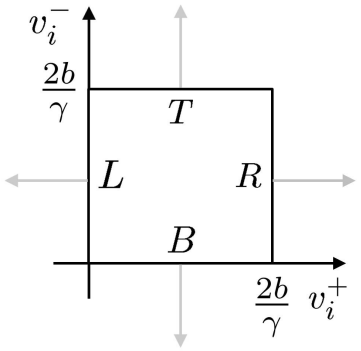


Figure 3.8: The black square Ω in single-ODE phase space together with the grey normal vectors at the boundaries R, T, L, B.

Note that we use the backward difference quotient for v^+, the particles that move in the positive direction, and the forward one for v^-, the particles that move in the negative direction. Ω is compact and $\hat{V}$ invariant, as is seen in the following calculation. Let

$$v_{i,R} = \begin{pmatrix} \frac{2b}{\gamma} \\ t_R \end{pmatrix}, \qquad v_{i,T} = \begin{pmatrix} t_T \\ \frac{2b}{\gamma} \end{pmatrix}, \qquad v_{i,L} = \begin{pmatrix} 0 \\ t_L \end{pmatrix}, \qquad v_{i,B} = \begin{pmatrix} t_B \\ 0 \end{pmatrix}$$

denote the right, top, left and bottom boundary of Ω, and $v_{i-1}^+ = r, v_{i+1}^- = s$, where $r, s, t_R, t_T, t_L, t_B \in [0, \frac{2b}{\gamma}]$. $n : \mathbb{R}^2 \to \mathbb{R}^2$ is the normal vector field of Ω, hence

$$n(v_{i,R}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad n(v_{i,T}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad n(v_{i,L}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad n(v_{i,B}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix};$$

see Fig. 3.8. With $\begin{Bmatrix} * \\ * \end{Bmatrix}$ as short-hand notation for two similar calculations therefore

$$\langle h(v^+_{i-1}, v_{i,\begin{Bmatrix} R \\ T \end{Bmatrix}}, v^-_{i+1}), n(v_{i,\begin{Bmatrix} R \\ T \end{Bmatrix}}) \rangle = \mu \left(t_{\begin{Bmatrix} R \\ T \end{Bmatrix}} - \frac{2b}{\gamma} \right) - \frac{2b^2}{\gamma} - bt_{\begin{Bmatrix} R \\ T \end{Bmatrix}} - \epsilon \left(\frac{2b}{\gamma} - \begin{Bmatrix} r \\ s \end{Bmatrix} \right) < 0,$$
$$\langle h(v^+_{i-1}, v_{i,\begin{Bmatrix} L \\ B \end{Bmatrix}}, v^-_{i+1}), n(v_{i,\begin{Bmatrix} L \\ B \end{Bmatrix}}) \rangle = -(\mu + b) t_{\begin{Bmatrix} L \\ B \end{Bmatrix}} - \epsilon \begin{Bmatrix} r \\ s \end{Bmatrix} \leq 0.$$

3.3.3.1 State Space Discretization I and CNDA I

Parallelizing the ideas of Example I we start with a partition according to the dynamics of $\partial_t w = h_1(w)$. There are 2 stationary points w_k, at which the Jacobian matrix has eigenvalues $\lambda_{k,l}$ and corresponding eigenvectors $w_{k,l}$, where $k \in \{u, s\}, l \in \{1, 2\}$ and

$$w_u = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \lambda_{u,1} = 2b, \quad w_{u,1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda_{u,2} = -2\mu, \quad w_{u,2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$
$$w_s = \frac{b}{\gamma} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda_{s,1} = -2b, \quad w_{s,1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \lambda_{s,2} = -2(b + \mu), \quad w_{s,2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

According to the criterion of Bendixson-Dulac [72] there are neither periodic orbits nor homoclinic orbits nor heteroclinic cycles in $(\mathbb{R}^+)^2$,

$$\operatorname{div}\left(\frac{1}{w^+ w^-} h_1(w)\right) = -(b + \mu) \left(\frac{1}{(w^+)^2} + \frac{1}{(w^-)^2} \right) - \gamma \left(\frac{1}{w^+} + \frac{1}{w^-} \right) < 0,$$

and the criterion of Poincaré-Bendixson [72] tells that the stationary points are hence the only limiting objects.

Following the dynamics of the circle with radius $\frac{b}{\gamma}$ around w_2 in multiples of p yields a partition $T_1 : \Omega \to E = \{0, 1, ..., N\}$, where E is ordered. An example is given in Fig. 3.9a for the parameters

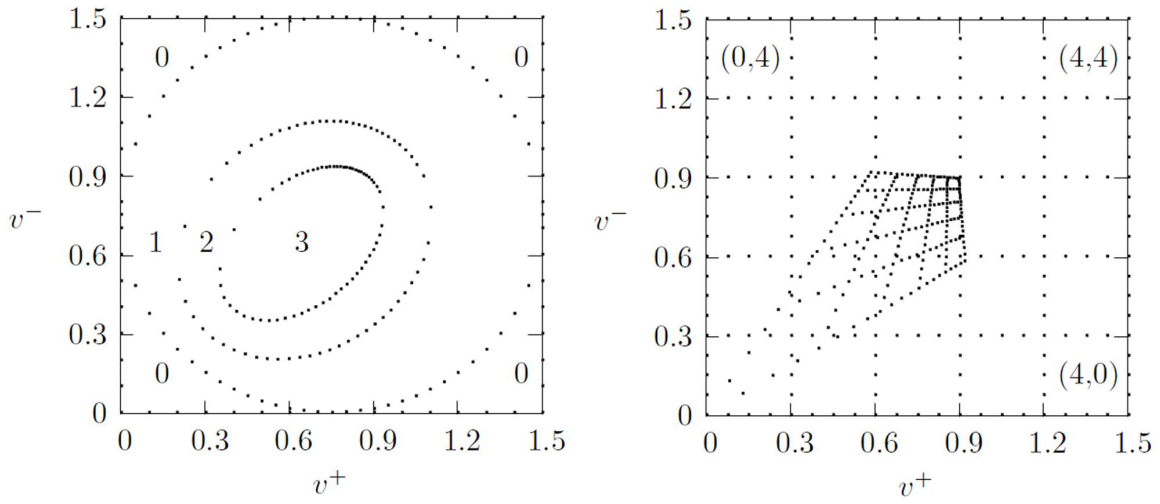
$$b = 0.75, \quad \gamma = 1, \quad \mu = 0.25, \quad p = \frac{1}{3}, \quad N = 3.$$

If perturbations are small enough, the resulting CNDA for $P = kp, k \in \mathbb{N}, 1 < k$ is monotonic and maps intervals to intervals. It can be analyzed with the subsolution method.

3.3.3.2 State Space Discretization II and CNDA II

Instead of one preserving $h_1(w_i)$ we now consider a partition T_2 which is invariant with respect to the transportation equation $\partial_t w_i = h_2(w^+_{i-1}, w_i, w^-_{i+1})$. If time and space discretization p and Δx are chosen such that $\frac{cp}{\Delta x} = 1$, we will end up with an equation

$$\begin{pmatrix} w^+_i((k+1)p) \\ w^-_i((k+1)p) \end{pmatrix} = \begin{pmatrix} w^+_{i-1}(kp) \\ w^-_{i+1}(kp) \end{pmatrix}$$



(a) **(b)**

Figure 3.9: (a) Partition I, where the domains between two consecutive images of the circle are labeled from $E = \{0, 1, 2, 3\}$ (b) Partition II with 25 squared boxes that are labeled as indicated. The partition is not invariant under h_1-dynamics, as can be seen from the image of the domain boundaries (parameters see the text).

for $k \in \mathbb{N}$. The suitable domains will be boxes. We take 25 squares as an example for

$$b = 0.75, \quad \gamma = 1, \quad \mu = 0.25, \quad \epsilon = 1, \quad P = 3/4$$

and label them with $E = \{(i, j) \,|\, i, j \in \{0, ..., 4\}\}$ as in Fig. 3.9b. Then we find e.g.

$$\tilde{f}((0, 0), *, (2, 4)) = (0, 4).$$

The partition is not invariant with respect to h_1-dynamics. The behavior of Eq. 3.1 can be numerically approximated to find a CNDA $(\mathbb{Z}, \{-1, 0, 1\}, E, f_0)$ with e.g.

$$f_0(((2, 2), (0, 0), (0, 0))) = \{(1, 0), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2)\},$$
$$f_0(((2, 2), (4, 4), (4, 4))) = \{(2, 2), (2, 3)\}.$$

Because we do not have a preserved order on E any more, we have to use pattern SA to analyze the CNDA.

Discretization II is computationally much easier to handle than discretization I. If we map test points in order to investigate the dynamics, standard algorithms can be used to identify the domains that they are mapped to. This is more difficult for domains with other shapes. However, the error in comparison to the PDE dynamics is bigger in the CNDA from discretization II. There the number of image domains for a given preimage is usually higher than with discretization I. Note that in the application at hand our choice of parameters ensures that at least not too much information is lost: interaction with a neighbor at site $i - 1$ in the stable $(2, 2)$ moves the local state at site i away from $(0, 0)$. In general the idea of discretization I may also be interesting for usage in uncertainty propagation with cellular probabilistic automata; see Chap. 5.

3.4 Conclusion and Outlook

In this chapter we have introduced cellular non-deterministic automata (CNDA) as time-discrete dynamical systems on the power set of deterministic cellular automata's (CA) state space. We have studied the construction in a CA context and have especially investigated a topological characterization. It has been shown that CNDA can be embedded in deterministic CA on a different grid, but that one needs other structures to approximate their long-term behavior in practice. We have developed pattern superautomata and the subsolution supersystem for this purpose. It has turned out that a special challenge in all approaches is the conservation of spatial correlations.

By choosing a suitable discretization of space, time and state space the dynamics of a PDE can be reproduced with a CNDA to some extent. This works well for the illustrative example of the Fisher-KPP equation, but with growing number of symbols one runs into the problem of exponential state space inflation. However, we point out that a potential simulation in the CNDA setup will still exhibit the computational advantages of CA like spatial parallelization.

Although it is possible to approximate certain PDE dynamics with the proposed methods, often the loss of information is too high such that the simplified model loses information content. An improvement that we follow in the next chapter is to weight the different transitions by occurrence probabilities and thus reveal more of the underlying PDE dynamics. This leads to cellular probabilistic automata (CPA) that are capable of working probabilistically on the power set of a deterministic CA's state space. Hence the new objects can be used for density based uncertainty propagation in PDE models. However, the conservation of spatial correlations will persist to be a central problem in the CPA approach, and we will use the insights of this chapter to resolve some problems. Especially the pattern idea will be used extensively.

Apart from that it might be an interesting goal for future work to extend the CNDA approach for PDE to infinitely but countably many symbols. This may help to better resolve the PDE's state space near unstable stationary states. Uniform structures might provide theoretical access to this idea [26].

4 Cellular Probabilistic Automata

In this chapter we suggest a novel numerical scheme for uncertainty propagation in distinct spatio-temporal processes. The main idea is to translate them into objects that we call cellular probabilistic automata (CPA) and to evolve the latter. The translation is achieved by state space discretization as in set oriented numerics and the use of the locality concept from cellular automata (CA) theory. We develop the method for the example of the propagation of initial value uncertainties under deterministic partial differential equation (PDE) dynamics. In the next chapter we will then pave the way towards an extension to more general stochastic influences on the system and test the method in several applications.

In particular we introduce a discretization of a PDE which does not depend explicitly on the independent variables. It is similar to the construction of cellular non-deterministic automata (CNDA) in Chap. 3. This time we first apply a finite difference scheme to a PDE; the spatial and temporal continuum is replaced by discrete sites and discrete time steps. Second, the state space of the resulting system is discretized. The discretized state space allows for an interpretation of risk levels or threshold values. As this procedure emphasizes the interaction between neighboring sites, a property that strongly resembles the locality and shift-invariance in cellular automata, the resulting completely discrete system is termed a cellular probabilistic automaton (CPA). Such an automaton is much simpler than the PDE and becomes accessible to very efficient simulation techniques.

While with CNDA we simply collected all possible transitions between discretized portions of phase space in a site's neighborhood, CPA basically consist of information about transition probabilities between them. The transition probabilities are interpreted as approximating the evolution of the system's probability density in transfer operator theory [106]. Hence CPA can be used for uncertainty propagation [122]. The translation from PDEs into CPA may be rather time-consuming, but the evolution of uncertainties with CPA is fast. The accuracy of the approximation depends on two parameters: one measures the state space resolution at every site, and the other the degree of locality, i.e., the extent to which correlations between neighboring sites are preserved. To formalize the degree of locality we develop the non-deterministic de Bruijn calculus from Sec. 3.2.3 further to a probabilistic de Bruijn calculus.

The chapter can be read independently of the latter chapter and is structured as follows. In Sec. 4.1 we formulate the problem of initial value uncertainty propagation under deterministic dynamics and deterministic boundary conditions. Here we also present the idea of density based uncertainty propagation through phase space discretization. By exploiting locality and shift-invariance of our problem this leads to the definition and discussion of CPA in Sec. 4.2. Here we introduce the probabilistic de Bruijn calculus. In Sec. 4.3 we present a consistency result for our construction. Finally, we give our conclusions in Sec. 4.4.

4.1 Density Based Uncertainty Propagation

In this section we first formulate the problem and then describe what we mean by density based uncertainty propagation. Finally, the CPA idea is derived in this context.

4.1.1 Problem Formulation

We are interested in the time evolution of uncertain initial data in a specific deterministic dynamical system. First, we specify the deterministic dynamical system that we will work with, and second we formulate the problem.

Consider the measure space $(\mathbb{R}^{mn}, \mathcal{B}(\mathbb{R}^{mn}), \lambda)$, where $m, n \in \mathbb{N}$, $\mathcal{B}(\mathbb{R}^{mn})$ is the Borel σ-algebra and λ the Lebesgue measure. Let $P_S : \mathcal{L}^1(\mathbb{R}^{mn}) \to \mathcal{L}^1(\mathbb{R}^{mn})$ denote the Frobenius-Perron operator (FPO) that describes how densities are mapped under phase space evolution with a nonsingular map $S : \mathbb{R}^{mn} \to \mathbb{R}^{mn}$; see Sec. 2.5. We focus on a particular type of phase space evolution.

Definition 4.1. Consider a deterministic dynamical system $(T, \mathbb{R}^{mn}, \Phi)$ specified as follows:

i) $(T, +)$ is an additive semigroup of time,
ii) $\mathbb{R}^{mn} = \times_I \mathbb{R}^n$ is the state space, where $I = \{1, ..., m\}$ and $1 \leq m$,
iii) the flow $\Phi : T \times \mathbb{R}^{mn} \to \mathbb{R}^{mn}$ is nonsingular for all $t \in T$,
iv) there is a neighborhood $U = \{-1, 0\}$, such that Φ has the locality property, i.e., that there is $h : T \times (\mathbb{R}^n)^2 \to \mathbb{R}^n$ with

$$\Phi(t, v)_i = h(t, v_{i-1}, v_i)$$

for all $t \in T, v = (v_1, ..., v_m) \in \times_I \mathbb{R}^n$ and $i \in \{2, ..., m\}$, and
v) that the system acts as the identity at the left, i.e., $\Phi(t, v)_1 = v_1$ for all $t \in T$ and all $v \in \times_I \mathbb{R}^n$.

We will write $\Phi^t(v) := \Phi(t, v)$ in the following. Assume that there is a compact $\Omega \subsetneq \mathbb{R}^n$ such that Ω^m is positively invariant under the flow, and fix $\tau \in T, \tau \neq 0$.

Our main application is the analysis of a PDE

$$\partial_t v = \tilde{h}(\partial_{xx} v, \partial_x v, v), \quad v(x, t) \in \Omega$$

on a one-dimensional compact spatial domain $x \in [a, b]$ for $a, b \in \mathbb{R}$. Note that unlike for CNDA we are now only interested in a bounded space domain to avoid technical difficulties. Under certain assumptions a dynamical system like the above is obtained by applying a finite difference method with space discretization $\Delta x = \frac{b-a}{m-1}$, where $m \in \mathbb{N}, m \geq 2$, and time step τ. Then U is naturally induced by the choice of the finite difference stencil. For notational simplicity we restrict to $U = \{-1, 0\}$ in this chapter. It is suitable to account for left first order difference quotients in advection reaction equations. In applications in Chap. 5 we also consider more complex PDEs and neighborhoods and comment on the generalizations when needed. In contrast to CNDA we use a finite difference scheme instead of a method of lines here to discretize space and time in one step. The reason is that we will not use dynamics to partition the phase space, but rather a very simple geometrical discretization. Because of the PDE context we call I the set of sites. By considering

only trajectories with $v^0(1) = k \in \mathbb{R}^n$, the system can be interpreted as to obey boundary conditions.

The time evolution of uncertain initial data in the deterministic dynamical system is described by real random variables $V^0, V^1, ... : X \to \Omega^m$ on probability space $(X, \mathcal{A}, \mu)$, where $V^{n+1} = \Phi^\tau V^n$. We focus on deterministic boundary conditions: $V^0(x)_1 = k \in \mathbb{R}^n$ for all $x \in X$. If V^n has density $g^n \in D(\mathbb{R}^{mn})$, the density of V^{n+1} is given by application of the associated FPO: $g^{n+1} = P_{\Phi^\tau}(g^n)$. The goal is to develop an algorithm that approximates the density evolution. It will be achieved by translating the system into a CPA in two steps. First, the FPO is discretized via a state discretization procedure, and then locality and shift-invariance are used to further transform it into a CPA.

4.1.2 State Space Discretization

In this section, first, we introduce the concept of state space discretization similar to Sec. 3.3.1.1 for CNDA. Second, we investigate according densities, and third, we construct a discretized version of the FPO. In principle these ideas are well known in the literature [34,35]. Here they are adapted to the special structure of the dynamical system.

Definition 4.2. A *partition* or *coding* E of Ω is a finite collection of disjoint sets $\{\Omega_e\}_{e\in E}$ whose union is Ω. We call $e \in E$ the *symbol* of *coding domain* Ω_e, and the *coding map* is the function $T : \Omega \mapsto E$, where $T(v) = e$ if $v \in \Omega_e$. A partition is called *uniform* if there is a resolution $\Delta\Omega \in \mathbb{R}$ such that Ω_e is an n-dimensional hypercube with side length $\Delta\Omega$ for all $e \in E$.

To avoid technical complications in the following proofs we consider only uniform partitions while developing the theory. They are also the ones that are relevant in practical algorithms.

A partition E of Ω with coding map T and $|E| = N$ naturally induces a partition E^I of Ω^m with coding map

$$\hat{T} : \Omega^m \to E^I, v \mapsto \hat{T}(v) \text{ with } (\hat{T}(v))_i = T(v_i) \text{ for } i \in I.$$

Note that $|E^I| = N^m$. For $\varphi \in E^J$, where $J \subseteq I$, we write

$$\Omega_\varphi = \{v \in \Omega^m \,|\, \forall j \in J : \hat{T}(v)(j) = \varphi(j)\}.$$

Now we study densities that are compatible with state space discretization. For this purpose we introduce the measure space $(E^I, \mathcal{P}(E^I), \gamma)$, where $\mathcal{P}(E^I)$ is the power set of E^I and γ is the counting measure. The densities $D(E^I)$ consist of the weight functions

$$g : E^I \to [0, \infty], \quad g(\varphi) = p_\varphi,$$

where $(p_\varphi)_{\varphi\in E^I}$ are nonnegative numbers with $\sum_{\varphi\in E^I} p_\varphi = 1$.

Definition 4.3.

i) $\mathcal{L}^1_{\hat{T}}(\mathbb{R}^{mn}) = \text{span}(B)$ is the finite-dimensional $\mathcal{L}^1(\mathbb{R}^{mn})$-subspace of piecewise constant functions with basis $B = \{\chi_{\Omega_\varphi}/\lambda(\Omega_\varphi)\}_{\varphi\in E^I}$. The set of piecewise constant densities is given by $D_{\hat{T}}(\mathbb{R}^{mn}) := \mathcal{L}^1_{\hat{T}}(\mathbb{R}^{mn}) \cap D(\mathbb{R}^{mn})$.

ii) The coordinate representation $\kappa_B : \mathcal{L}^1_{\hat{T}}(\mathbb{R}^{mn}) \to \mathbb{R}^{E^I}, \quad g \mapsto \kappa_B(g)$ with respect to the basis B is given by $\kappa_B(g)(\varphi) = c_\varphi$ for $\varphi \in E^I$ and $g = \sum_{\psi \in E^I} \frac{c_\psi}{\lambda(\Omega_\psi)} \chi_{\Omega_\psi}$. Obviously $\kappa_B(D_{\hat{T}}(\mathbb{R}^{mn})) = D(E^I)$.

iii) Let $\rho \in E$ such that $\rho = T(k)$. The densities that are compatible with the boundary conditions are given by

$$D_{BC}(E^I) := \{g \in D(E^I) \,|\, g(\varphi) = 0 \text{ if } \varphi(1) \neq \rho\}.$$

By averaging in the coding domains, every function in $\mathcal{L}^1(\mathbb{R}^{mn})$ can be mapped to a piecewise constant function.

Definition 4.4. A restriction operator to the subspace of piecewise constant functions is given by

$$R : \mathcal{L}^1(\mathbb{R}^{mn}) \to \mathcal{L}^1_{\hat{T}}(\mathbb{R}^{mn}), \quad R(g) = \sum_{\varphi \in E^I} \frac{c_\varphi}{\lambda(\Omega_\varphi)} \chi_{\Omega_\varphi},$$

where

$$c_\varphi = \int_{\Omega_\varphi} g(w) dw.$$

R is idempotent, i.e., $R \circ R = R$, and furthermore $R(D_{\hat{T}}(\mathbb{R}^{mn})) \subseteq D_{\hat{T}}(\mathbb{R}^{mn})$. In the following we will use the restriction operator to construct a discretized version of the FPO on the density level: RP_{Φ^τ}. This procedure is well known in ergodicity theory when invariant measures are approximated. There it is called Ulam's method [168].

The matrix representation of the linear $RP_{\Phi^\tau}|_{\mathcal{L}^1_{\hat{T}}(\mathbb{R}^{mn})}$ is given by $P_B = \kappa_B R P_{\Phi^\tau} \kappa_B^{-1} \in \mathbb{R}^{E^I \times E^I}$ with entries

$$P_{B,\varphi,\psi} = \int_{\Omega_\psi} P_{\Phi^\tau} \frac{\chi_{\Omega_\varphi}}{\lambda(\Omega_\varphi)} d\lambda = \int_{\Phi^{-\tau}(\Omega_\psi)} \frac{\chi_{\Omega_\varphi}}{\lambda(\Omega_\varphi)} d\lambda = \frac{\lambda(\Omega_\varphi \cap \Phi^{-\tau}(\Omega_\psi))}{\lambda(\Omega_\varphi)}.$$

$P_{B,\varphi,\psi}$ is the probability of finding a realization of a random variable with uniform density in Ω_φ in Ω_ψ, when Φ^τ is applied. Hence we may interpret $P_{B,\varphi,\psi}$ as the transition rate from Ω_φ to Ω_ψ of a finite state Markov chain on $\{\Omega_\varphi\}_{\varphi \in E^I}$. This chain approximates the behavior of the dynamical system for uncertain initial values.

In the following we regard $P_B : D_{BC}(E^I) \to D_{BC}(E^I)$ as a function which maps densities that are compatible with the boundary conditions by matrix multiplication.

4.1.3 Using Locality - Towards Cellular Probabilistic Automata

E^I grows exponentially in m. For a growing number of sites it becomes numerically expensive to obtain global transition rates and to handle global states and densities.

However, our dynamical system has a special structure: We use the locality property to approximate the set of global transition probabilities by several identical sets of local ones. This is possible for two reasons. The first is because we find identical dynamics at all sites away from the boundaries, and the second is because the transition probabilities at one particular site mainly depend on the state of its neighborhood rather than on the whole global configuration.

For the formal definition of these local transition probabilities we need to introduce the shift by $l \in \mathbb{Z}$ on finite grid $J \subset \mathbb{Z}$. It is given by

$$\sigma_l : F^J \to F^{-l+J}, \quad \varphi \mapsto \sigma_l(\varphi), \quad \sigma_l(\varphi)(-l+j) = \varphi(j),$$

where F is an arbitrary set, e.g. $F = E$ or $F = D(E^V)$. Moreover, for arbitrary $V = \{-p, ..., q\}, W = \{-t, ..., u\}$ with $p, q \in \mathbb{N}_0$, $t, u \in \mathbb{Z}, -t \leq u$, and $l \in \mathbb{Z}$ we use the conventions $l + V = \{-p + l, ..., q + l\}$ and $V + W = \{-p - t, ..., q + u\}$.

Definition 4.5. Let $V = \{-p, ..., q\}$ with $p, q \in \mathbb{N}_0$ and $1 + p + q \leq m$. A *local function* $f_0 : E^{U+V} \to D(E^V)$ is then given by

$$f_0(\varphi)(\psi) = \frac{\lambda(\Omega_{\sigma_{-i}(\varphi)} \cap \Phi^{-\tau}(\Omega_{\sigma_{-i}(\psi)}))}{\lambda(\Omega_{\sigma_{-i}(\varphi)})},$$

where $i = 2 + p$, $\varphi \in E^{U+V}$ and $\psi \in E^V$

Note that because of the locality property the definition is independent of the chosen site $i \in \{2 + p, ..., m - q\}$. The set V controls the degree of locality, i.e., the number of sites that give rise to a local transition. It will turn out that by enlarging it we can diminish the error of the locality approximation.

In the following section we develop a method of how to combine several such local transitions to approximate a global one. This will finish the construction of a CPA from the FPO.

4.2 Introduction of Cellular Probabilistic Automata

CPA are defined by extending the definition of deterministic CA according to [38, 50]: In CPA the local transition function specifies a time- and space-independent probability distribution of next states for each possible neighborhood configuration. As we do not want to follow one realization but rather the whole ensemble, unlike in the literature we define CPA to work on densities. This enables their utilization for uncertainty propagation. The difference to CNDA is that there the transitions are just collected, but not weighted by probabilities like in CPA.

In the last section we showed how the discretized FPO P_B on state space $D_{BC}(E^I)$ can be used to approximate the FPO P_{Φ^τ} on $D(\mathbb{R}^{mn})$. CPA further approximate the discretized FPO on a product space of local densities; see Fig. 4.1 for a sketch. This is again the de Bruijn pattern idea that we also used to analyze CNDA in Sec. 3.2.3, but now in a probabilistic setting. Uncertainty propagation with CPA therefore requires two definitions. The first one is about how to translate between global densities and the product space of local densities, and the second one is about how to evolve local densities in time with the help of the local function.

Since the definitions can be best understood for $V = \{0\}$, in Sec. 4.2.1 we first introduce CPA in this special case to demonstrate the basic construction. Afterwards we develop the de Bruijn calculus as a connection between local and global objects for more general V on finite grids in Sec. 4.2.2. This connection leads to the generalization of CPA to general V in Sec. 4.2.3.

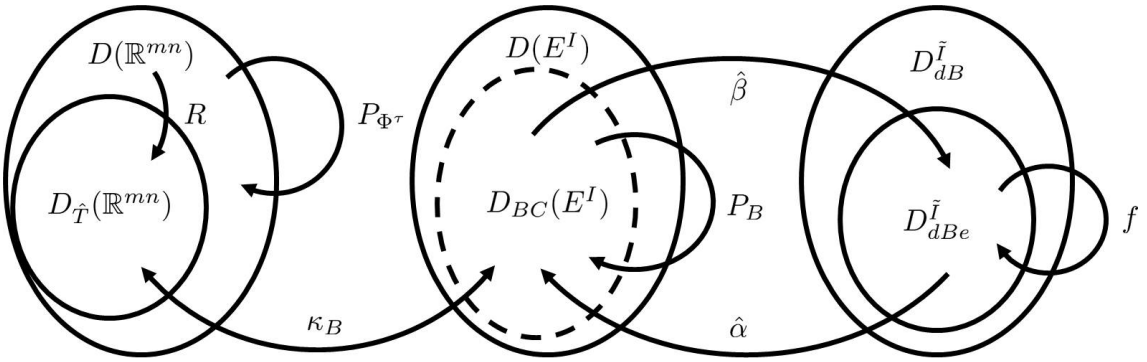


Figure 4.1: The relations between the FPO P_{Φ^τ} and its approximations. By state discretization we obtain the discretized FPO P_B which still works globally, and by exploiting locality we approximate P_B further by the CPA with global function f. The state space on which the CPA operates is a collection of local densities; see the text.

4.2.1 Cellular Probabilistic Automata: a Special Case

A crucial step is to translate between global densities $D(E^{V+W})$ and a (subset of a) collection of local densities $(D(E^V))^W$, where $V = \{-p, ..., q\}$ and $W = \{-t, ..., u\}$ for $p, q, t, u \in \mathbb{N}_0$. We introduce a projection operator $\beta_W : D(E^{V+W}) \to (D(E^V))^W$ and an embedding $\alpha_W : (D(E^V))^W \to D(E^{V+W})$. They correspond to γ and β in the de Bruijn calculus on power sets in Def. 3.19, respectively. β_W localizes the information to densities on states of length $|V|$ and thus erases far-reaching correlations. α_W in turn constructs global densities out of information about local densities. As we will see in the next section, this process is by no means unique and requires some technical refinement of the space of local densities. However, for $V = \{0\}$ there are canonical definitions for α_W and β_W: multiplication of local probabilities for independent events and calculation of marginal distributions.

Definition 4.6. Let $W = \{-t, ..., u\}$ for $t, u \in \mathbb{N}_0$.

i) We set $\alpha_W : (D(E))^W \to D(E^W)$, $g \mapsto \alpha_W(g)$ with

$$\alpha_W(g)(\psi) = \prod_{i \in W} g(i)(\psi(i))$$

ii) and $\beta_W : D(E^W) \to (D(E))^W$, $g \mapsto \hat{\beta}_W(g)$ with

$$\beta_W(g)(i)(e) = \sum_{\chi \in E^W \text{ s.t. } \chi(i)=e} g(\chi).$$

We want to keep the construction simple at this point and close the grid I to a torus $\mathbb{Z}_m$. This way we avoid boundary conditions in this special section.

Definition 4.7. A *cellular probabilistic automaton (CPA)* is a tuple (I, U, E, f_0), where for $m \in \mathbb{N}$, $2 \leq m$,

i) $I = \mathbb{Z}_m$ is a toroidal grid,
ii) $U = \{-1, 0\}$ is the neighborhood,
iii) E is a finite set of local states, and
iv) $f_0 : E^U \to D(E)$ is the local function.

The global function $f : (D(E))^I \to (D(E))^I, g \mapsto f(g)$ is given by

$$f(g)(i)(\psi) = \sum_{\varphi \in E^U} \alpha_U(\sigma_i(g)|_U)(\varphi) f_0(\varphi)(\psi).$$

The trajectory starting with $g^0 \in (D(E))^I$ is given by the sequence $(g^n)_{n\in\mathbb{N}}$, where $g^n = f(g^{n-1})$ for $n = 1, 2, ...$

A CPA can be used to evolve an input distribution $\beta_I(g)$ for $g \in D(E^I)$ via the global function. After n time steps the approximated global density is then given by $\alpha_I f^n \beta_I(g)$; see also Fig. 4.1 with $D^{\tilde{I}}_{dB} = D^{\tilde{I}}_{dBe} = (D(E))^I$, $\hat{\alpha} = \alpha_I$, and $\hat{\beta} = \beta_I$. The role model for the global function is the matrix operation with the discretized FPO P_B: The product of the transition probability with the probability of being in a preimage state is summed up over all possible preimage states. A probability is assigned to a preimage state $\varphi \in E^U$ by α_U.

Note that pattern superautomata of CNDA with $V = \{0\}$ on finite grids are equivalent to certain CPA of this form: assume that for all $\varphi \in E^U$ there is $c \in [0, 1]$ such that $f_0(\varphi)(e) = c$ or $f_0(\varphi)(e) = 0$ for all $e \in E$, and that the input is also a collection of states with equal non-zero or zero probability. We will see that an analogous argument also holds for larger patterns. Furthermore deterministic CA are special cases of CPA: assume that for all $\varphi \in E^U$ there is $e \in E$ such that $f_0(\varphi)(e) = 1$ and that the input is deterministic.

4.2.2 De Bruijn Calculus

To generalize the construction to arbitrary V we first study the relation between local and global objects in more depth. We introduce de Bruijn density calculus by further developing the de Bruijn calculus on power sets from Sec. 3.2.3. As before it is based on pattern ideas in CA theory [76,173], in the theory of de Bruijn graphs [162], and in pair approximation [96]. Ideas similar to our probabilistic extension are also at the heart of probabilistic graphical models [102].

Like in the special case we introduce a projection operator β_W that localizes the global information to densities on states of length $|V|$, this time $|V| \geq 1$. The precise definition of β_W is still rather straightforward: marginal distributions dismiss all information but that over a certain range V. We will find below that the reconstruction of global densities out of local information by α_W is more involved. However, let us first define β_W.

Definition 4.8. Let $V = \{-p, ..., q\}$ and $W = \{-t, ..., u\}$ for $p, q \in \mathbb{N}$ and $t, u \in \mathbb{Z}, -t \leq u$. $\beta_W : D(E^{V+W}) \to (D(E^V))^W$ is given by

$$\beta_W(g)(i)(\psi) = \sum_{\substack{\chi \in E^{V+W} \text{ s.t.} \\ \chi|_{i+V} = \sigma_{-i}(\psi)}} g(\chi).$$

Example 4.1. Let $E = V = W = \{0, 1\}$, $c \in (0, 1)$, and $g, \tilde{g} \in D(E^{V+W})$ be given by

$$g(\psi) = \begin{cases} c & \text{if } \psi = (001) \\ 1-c & \text{if } \psi = (100) \\ 0 & \text{else} \end{cases}, \qquad \tilde{g}(\psi) = \begin{cases} c(1-c) & \text{if } \psi = (000), \psi = (101) \\ c^2 & \text{if } \psi = (001) \\ (1-c)^2 & \text{if } \psi = (100) \\ 0 & \text{else} \end{cases}.$$

We find that $\beta_W(g) = \beta_W(\tilde{g})$ with

$$\beta_W(g)(0)(\psi) = \begin{cases} c & \text{if } \psi = (00) \\ 1-c & \text{if } \psi = (10) \\ 0 & \text{else} \end{cases}, \qquad \beta_W(g)(1)(\psi) = \begin{cases} 1-c & \text{if } \psi = (00) \\ c & \text{if } \psi = (01) \\ 0 & \text{else} \end{cases}.$$

Ex. 4.1 shows that information is lost under β_W, i.e., different global densities are mapped to the same collection of local densities. Now we are interested in the embedding $\alpha_W : (D(E^V))^W \to D(E^{V+W})$. Although the properties of β_W allow us to define α_W as the solution of a linear nonnegative least squares problem [108], this algebraic approach is not appropriate. We rather suggest a probabilistic approach that fulfills two natural requirements: first, our α_W degenerates to simple multiplication of local densities for $V = \{0\}$, and second, α_W and β_W are inverse of each other on important sets.

We first introduce several definitions that are central to our approach; see also Fig. 4.2a, Def. 3.16 and Def. 3.21.

Definition 4.9.

i) $X_{dB}^W = (\mathcal{P}(E^V))^W$ is the set of *de Bruijn states*, where $\mathcal{P}(E^V)$ is the power set of E^V. The elements of E^V are called patterns of length $|V|$.

ii) The subset of *extendable de Bruijn states* is given by

$$X_{dBe}^W = \{\Phi \in X_{dB} \,|\, \forall i \in W\, \forall \varphi \in \Phi(i)\, \exists \psi \in E^{V+W}\, \forall j \in W : \\ \sigma_j(\psi)|_V \in \Phi(j), \sigma_i(\psi)|_V = \varphi\}.$$

iii) $D_{dB}^W = (D(E^V))^W$ is called the set of *de Bruijn densities*.

iv) The subset of *extendable de Bruijn densities* is given by

$$D_{dBe}^W = \{g \in D_{dB}^W \mid \times_{i \in W} \text{supp } g(i) \in X_{dBe}^W\}.$$

The idea behind extendable de Bruijn states is that every pattern can be extended to a global state by gluing suitable patterns on it. An example of an extendable de Bruijn density is $\beta_W(g)$ in Ex. 4.1: for example, pattern (10) at site 0 can be extended by (01) at site 1 to the global state (101), because the patterns coincide in the overlapping state 0. We find that this observation can be generalized.

Lemma 4.10. $\text{im}\,(\beta_W) \subseteq D_{dBe}^W$.

Proof Let $g \in \text{im}\,(\beta_W)$, $j \in W$, and $\varphi \in E^V$ with $g(j)(\varphi) > 0$. Then there are $\tilde{g} \in D(E^{V+W})$ and $\psi \in E^{V+W}$ such that $\tilde{g}(\psi) > 0$ and $\sigma_j(\psi)|_V = \varphi$. But, furthermore, already $g(i)(\sigma_i(\psi)|_V) > 0$ for all $i \in W$, and therefore $\times_{i \in W} \text{supp } g(i) \in X_{dBe}^W$. □

Since only the extendable de Bruijn densities are addressed by β_W, we need only define α_W on D_{dBe}^W. In our choice of α_W the probability of a global state is calculated by using conditional probabilities while concatenating according patterns. Starting with a pattern at a site $i \in W$ we extend it to the left and right with suitable patterns by one more site and condition on the overlap; see also Fig. 4.2b. By repeating this procedure site by site to the left and right we cover the whole grid with the desired global state. More formally, our choice is motivated by the following calculation, for which we first introduce some notation.

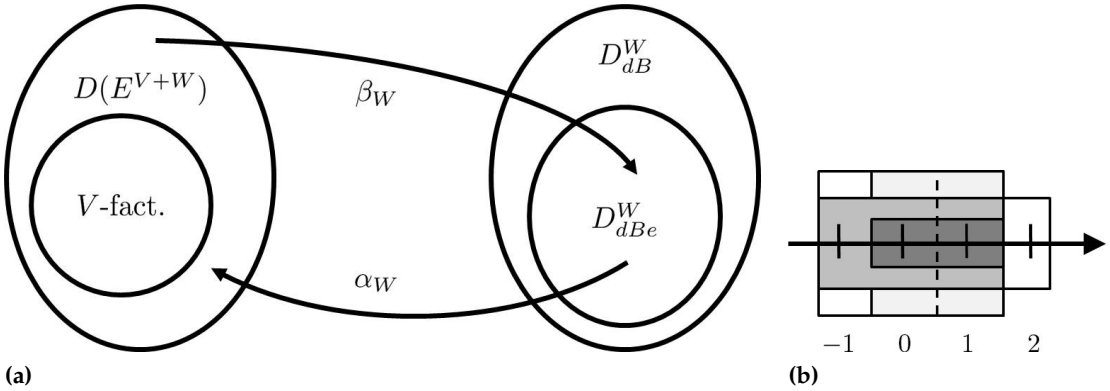


Figure 4.2: (a) The relation between global densities and de Bruijn densities. (b) An example of how a global density is approximated by local densities via an approximation of Markov type for $W = \{-1, 0, 1\}$, $V = \{0, 1\}$ and $i = 0$. The thin and dark grey box that covers sites 0 and 1 is the factor at site 0. The medium box that covers sites -1 to 2 is the factor at site 1: the local state at site 2 depends on the local states from -1 to 1 (medium grey part). In the approximation (dashed line) the local state at 2 depends only on that at 1. The thick box covering sites -1 to 1 is the factor at site -1. The local state at site -1 depends on the local states on sites 0 and 1 (light grey part). In the approximation the dependence stops again at the dashed line.

For finite $J \subseteq \mathbb{Z}$ and $A \subseteq E^J$, $\mu[A] = \sum_{\varphi \in A} g(\varphi)$ denotes the distribution associated with $g \in D(E^J)$. Furthermore,

$$\left\{\psi|^J_{\tilde{J}}\right\} = \{\varphi \in E^J \mid \varphi|_{\tilde{J}} = \psi|_{\tilde{J}}\}$$

for finite $\hat{J} \subseteq \mathbb{Z}$, $\tilde{J} \subseteq J \cap \hat{J}$ and $\psi \in E^{\hat{J}}$, and we also write $\{\psi|_{\tilde{J}}\} = \left\{\psi|^J_{\tilde{J}}\right\}$ if J is clear from the context. We calculate for $i \in W$ and $\psi \in E^{V+W}$ that

$$\begin{aligned}\mu[\{\psi\}] =& \mu[\{\psi|_{\{-t,\ldots,i\}+V}\}] \prod_{l=i+1}^{u} \frac{\mu[\{\psi|_{\{-t,\ldots,l\}+V}\}]}{\mu[\{\psi|_{\{-t,\ldots,l\}+V_-}\}]} \\ =& \prod_{k=-t}^{i-1} \mu[\{\psi|_{\{k-p\}}\} \mid \{\psi|_{\{k,\ldots,i\}+V_+}\}]\, \mu[\{\psi|_{i+V}\}] \prod_{l=i+1}^{u} \mu[\{\psi|_{\{l+q\}}\} \mid \{\psi|_{\{-t,\ldots,l\}+V_-}\}],\end{aligned}$$

where $V_+ = \{-p+1, \ldots, q\}$ and $V_- = \{-p, \ldots, q-1\}$. If there are no far-reaching correlations, we expect the following approximations to be suitable:

$$\begin{aligned}\mu[\{\psi|_{\{k-p\}}\} \mid \{\psi|_{\{k,\ldots,i\}+V_+}\}] &\approx \mu[\{\psi|_{\{k-p\}}\} \mid \{\psi|_{k+V_+}\}], \\ \mu[\{\psi|_{\{l+q\}}\} \mid \{\psi|_{\{-t,\ldots,l\}+V_-}\}] &\approx \mu[\{\psi|_{\{l+q\}}\} \mid \{\psi|_{l+V_-}\}]\end{aligned}$$

for $k \in \{-t, \ldots, i-1\}$ and $l \in \{i+1, \ldots, u\}$. They resemble a Markov property of order $|V| - 1$ in space; see also Fig. 4.2b: a site's state is independent from the states at sites that are more than $|V| - 1$ sites apart.

Lemma 4.11. Let μ_j be the distribution associated with $\beta_W(g)(j) \in D(E^V)$ for $j \in W$.

Then

$$\begin{aligned}\mu[\{\psi|_{\{k-p\}}\} \mid \{\psi|_{k+V_+}\}] &= \mu_k[\{\sigma_k(\psi)|_{\{-p\}}\} \mid \{\sigma_k(\psi)|_{V_+}\}],\\ \mu[\{\psi|_{i+V}\}] &= \mu_i[\{\sigma_i(\psi)|_V\}],\\ \mu[\{\psi|_{\{l-p\}}\} \mid \{\psi|_{l+V_-}\}] &= \mu_l[\{\sigma_l(\psi)|_{\{-p\}}\} \mid \{\sigma_l(\psi)|_{V_-}\}]\end{aligned}$$

for $i \in W$, $k \in \{-t, ..., i-1\}$ and $l \in \{i+1, ..., u\}$.

Proof Without loss of generality we prove the statement only for $k \in \{-t, ..., i-1\}$:

$$\begin{aligned}\mu[\{\psi|_{\{k-p\}}\} \mid \{\psi|_{k+V_+}\}] &= \frac{\sum_{\chi\in\{\psi|_{k+V}\}} g(\chi)}{\sum_{\chi\in\{\psi|_{k+V_+}\}} g(\chi)} = \frac{\beta_W(g)(k)(\sigma_k(\psi)|_V)}{\sum_{\varphi\in\left\{\sigma_k(\psi)|^V_{V_+}\right\}} \sum_{\chi\in\left\{\sigma_{-k}(\varphi)|^{V+W}_{k+V}\right\}} g(\chi)}\\ &= \frac{\beta_W(g)(k)(\sigma_k(\psi)|_V)}{\sum_{\varphi\in\left\{\sigma_k(\psi)|^V_{V_+}\right\}} \beta_W(g)(k)(\varphi)} = \mu_k[\{\sigma_k(\psi)|_{\{-p\}}\} \mid \{\sigma_k(\psi)|_{V_+}\}].\end{aligned}$$

□

Using the above approximations of Markov type and the lemma we find that

$$\begin{aligned}\mu[\{\psi\}] &\approx \prod_{k=-t}^{i-1} \mu[\{\psi|_{\{k-p\}}\} \mid \{\psi|_{k+V_+}\}]\, \mu[\{\psi|_{i+V}\}] \prod_{l=i+1}^{u} \mu[\{\psi|_{\{l+q\}}\} \mid \{\psi|_{l+V_-}\}]\\ &= \prod_{k=-t}^{i-1} \mu_k[\{\sigma_k(\psi)|_{\{-p\}}\} \mid \{\sigma_k(\psi)|_{V_+}\}]\, \mu_i[\{\sigma_i(\psi)|_V\}] \prod_{l=i+1}^{u} \mu_l[\{\sigma_l(\psi)|_{\{q\}}\} \mid \{\sigma_l(\psi)|_{V_-}\}].\end{aligned}$$

This way we are led to the following definition, which in general depends on the site i.

Definition 4.12. Let $i \in W$. Then $\alpha_{W,i} : D^W_{dBe} \to D(E^{V+W})$ is given by

$$\begin{aligned}\alpha_{W,i}(g)(\psi) = &\prod_{k=-t}^{i-1} \mu_k[\{\sigma_k(\psi)|_{\{-p\}}\} \mid \{\sigma_k(\psi)|_{V_+}\}]\ g(i)(\sigma_i(\psi)|_V)\\ &\prod_{l=i+1}^{u} \mu_l[\{\sigma_l(\psi)|_{\{q\}}\} \mid \{\sigma_l(\psi)|_{V_-}\}],\end{aligned}$$

where μ_j denotes the conditional distribution associated with $g(j) \in D(E^V)$ for $j \in W$, and $\psi \in E^{V+W}$.

For $W = \{u\}, u \in \mathbb{Z}$, the definition simplifies to $\alpha_{\{u\},u}(g)(\psi) = g(u)(\sigma_u(\psi))$. For $V = \{0\}$ the conditions vanish, and we get back simple multiplication, $\alpha_{W,i}(g)(\psi) = \prod_{k\in W} g(k)(\psi(k))$. We formally justify that our construction is well defined.

Lemma 4.13. Let $g \in D^W_{dBe}$ and $i \in W$. Then $\alpha_{W,i}(g) \in D(E^{V+W})$.

Proof Let $g \in D^W_{dBe}$ and $i \in W$. $\alpha_{W,i}(g) \not\equiv 0$, because there is at least one extendable pattern with nonzero probability. We prove that it is also normalized in the following. Without loss of generality we assume that $i = u$. Then

$$\begin{aligned}
\sum_{\varphi\in E^{V+W}} \alpha_{W,u}(g)(\varphi) &= \sum_{\varphi\in E^{V+W}} \prod_{k=-t}^{u-1} \mu_k[\{\sigma_k(\varphi)|_{\{-p\}}\} \mid \{\sigma_k(\varphi)|_{V_+}\}]\, \mu_u[\{\sigma_u(\varphi)|_V\}] \\
&= \sum_{\tilde\varphi\in E^{V+W\setminus\{-t\}}} \sum_{\varphi\in\left\{\tilde\varphi|^{V+W}_{V+W\setminus\{-t\}}\right\}} \mu_{-t}[\{\sigma_{-t}(\varphi)|_{\{-p\}}\} \mid \{\sigma_{-t}(\tilde\varphi)|_{V_+}\}] \\
&\quad \prod_{k=-t+1}^{u-1} \mu_k[\{\sigma_k(\tilde\varphi)|_{\{-p\}}\} \mid \{\sigma_k(\tilde\varphi)|_{V_+}\}]\, \mu_u[\{\sigma_u(\tilde\varphi)|_V\}] \\
&= \sum_{\tilde\varphi\in E^{V+W\setminus\{-t\}}} \prod_{k=-t+1}^{u-1} \mu_k[\{\sigma_k(\tilde\varphi)|_{\{-p\}}\} \mid \{\sigma_k(\tilde\varphi)|_{V_+}\}]\, \mu_u[\{\sigma_u(\tilde\varphi)|_V\}] \\
&= \ldots = \sum_{\varphi\in E^{u+V}} \mu_u[\{\sigma_u(\varphi)|_V\}] = \sum_{\varphi\in E^V} g(\varphi) = 1.
\end{aligned}$$

The steps indicated by ... follow by induction in $|W|$. □

Now we focus on the important set im (β_W) and investigate $\alpha_{W,i}$ on this subset of D^W_{dBe}. It turns out that enough information of the preimage state in $D(E^{V+W})$ is preserved under β_W to ensure that in this case $\alpha_{W,i}$ is independent of the starting site i.

Lemma 4.14. Let $i \in W$ and $g \in$ im (β_W). Then $\alpha_{W,i}(g) = \alpha_{W,j}(g)$ for all $i, j \in W$.

Proof We show that $\alpha_{W,i}(g) = \alpha_{W,i+1}(g)$ for all $i \in W\setminus\{-t\}$. An index shift to the left can be proven analogously.

Note that $\mu_i[\{\sigma_i(\psi)|_V\}] = g(i)(\sigma_i(\psi)|_V)$ and that for $k \in \{-t, ..., i-1\}$ and $l \in \{i+1, ..., u\}$,

$$\mu_k[\{\sigma_k(\psi)|_{\{-p\}}\} \mid \{\sigma_k(\psi)|_{V_+}\}] = \frac{g(k)(\sigma_k(\psi)|_V)}{\sum_{\chi\in\left\{\sigma_k(\psi)|^V_{V_+}\right\}} g(k)(\chi)},$$

$$\mu_l[\{\sigma_l(\psi)|_{\{q\}}\} \mid \{\sigma_l(\psi)|_{V_-}\}] = \frac{g(l)(\sigma_l(\psi)|_V)}{\sum_{\chi\in\left\{\sigma_l(\psi)|^V_{V_-}\right\}} g(l)(\chi)}.$$

Therefore $\alpha_{W,i}(g)(\psi)$ and $\alpha_{W,i+1}(g)(\psi)$ have the same numerator and differ only in the denominator. It is enough to show that a factor in the denominator may be shifted one step to the right: Let $i \in W\setminus\{u\}$, and let $g = \beta_W(\tilde g)$ for $\tilde g \in D(E^{V+W})$. Then

$$\begin{aligned}
&\sum_{\chi\in\left\{\sigma_i(\psi)|^V_{V_+}\right\}} \beta_W(\tilde g)(i)(\chi) && = \sum_{\chi\in\left\{\sigma_i(\psi)|^V_{V_+}\right\}} \sum_{\varphi\in\left\{\sigma_{-i}(\chi)|^{V+W}_{i+V}\right\}} \tilde g(\varphi) \\
&= \sum_{\varphi\in\left\{\psi|^{V+W}_{i+V_+}\right\}} \tilde g(\varphi) && = \sum_{\varphi\in\left\{\psi|^{V+W}_{i+1+V_-}\right\}} \tilde g(\varphi) \\
&= \sum_{\chi\in\left\{\sigma_{i+1}(\psi)|^V_{V_-}\right\}} \sum_{\varphi\in\left\{\sigma_{-(i+1)}(\chi)|^{V+W}_{i+1+V}\right\}} \tilde g(\varphi) && = \sum_{\chi\in\left\{\sigma_{i+1}(\psi)|^V_{V_-}\right\}} \beta_W(\tilde g)(i+1)(\chi) \qquad \square
\end{aligned}$$

Moreover, $\alpha_{W,i}$ and β_W are even inverse on this set.

Theorem 4.15. Let $g \in \mathrm{im}(\beta_W)$. Then $\beta_W \alpha_{W,i}(g) = g$ for all $i \in W$.

Proof Let $i \in W$ and $g \in \mathrm{im}(\beta_W)$. We prove $\beta_W \alpha_{W,i}(g)(j) = g(j)$ without loss of generality only for $j = u$.

With Lm. 4.14 and because a conditional distribution is a distribution as well, for $\psi \in E^V$,

$$\begin{aligned}
\beta_W \alpha_{W,i}(g)(u)(\psi) =& \beta_W \alpha_{W,u}(g)(u)(\psi) = \sum_{\varphi \in \left\{\sigma_{-u}(\psi)|_{u+V}^{V+W}\right\}} \alpha_{W,u}(g)(\varphi) \\
=& \sum_{\tilde{\varphi} \in \left\{\sigma_{-u}(\psi)|_{u+V}^{V+W\setminus\{-t\}}\right\}} \sum_{\varphi \in \left\{\tilde{\varphi}|_{V+W\setminus\{-t\}}^{V+W}\right\}} \mu_{-t}[\{\sigma_{-t}(\varphi)|_{\{-p\}}\} \mid \{\sigma_{-t}(\tilde{\varphi})|_{V_+}\}] \\
& \prod_{k=-t+1}^{u-1} \mu_k[\{\sigma_k(\tilde{\varphi})|_{\{-p\}}\} \mid \{\sigma_k(\tilde{\varphi})|_{V_+}\}]\, \mu_u[\{\sigma_u(\tilde{\varphi})|_V\}] \\
=& \sum_{\tilde{\varphi} \in \left\{\sigma_{-u}(\psi)|_{u+V}^{V+W\setminus\{-t\}}\right\}} \prod_{k=-t+1}^{u-1} \mu_k[\{\sigma_k(\tilde{\varphi})|_{\{-p\}}\} \mid \{\sigma_k(\tilde{\varphi})|_{V_+}\}]\, \mu_u[\{\sigma_u(\tilde{\varphi})|_V\}] \\
=& \ldots = g(u)(\psi).
\end{aligned}$$

The last steps follow by induction in $|W|$. □

Next we consider the opposite way and focus on global densities that are preserved under $\alpha_{W,i}\beta_W$ for all $i \in W$. We will see that they enable an algebraic interpretation of β_W.

Definition 4.16. $g \in D(E^{V+W})$ is called *V-factorizable* if $g = \alpha_{W,i}\beta_W(g)$ for all $i \in W$.

An example of a $\{0,1\}$-factorizable density is $\tilde{g}$ in Ex. 4.1. g in the same example, however, has correlations over more than 2 sites: a state has positive probability only if the local states at sites 0 and 2 differ. This long-range correlation cannot be preserved under mappings with pattern length $|\{0,1\}| = 2$, and therefore g is not $\{0,1\}$-factorizable. Obviously the set of V-factorizable states is the natural counterpart of im (β_W).

Lemma 4.17. Let $i \in W$ and $g \in$ im (β_W). Then $\alpha_{W,i}(g)$ is V-factorizable.

Proof Let $g \in$ im (β_W) and $i, j \in W$. Then by Lm. 4.14 and Thm. 4.15

$$\alpha_{W,i}(g) = \alpha_{W,j}(g) = \alpha_{W,j}\beta_W \alpha_{W,i}(g),$$

and hence $\alpha_{W,i}(g)$ is V-factorizable. □

The important role of V-factorizable states for the algebraic interpretation of β_W is highlighted by the next result.

Theorem 4.18. For all $g \in D(E^{V+W})$ there is a unique V-factorizable $\tilde{g} \in D(E^{V+W})$ such that $\beta_W(\tilde{g}) = \beta_W(g)$.

Proof Let $g \in D(E^{V+W})$, and choose $\tilde{g} = \alpha_{W,u}\beta_W(g)$. Then $\tilde{g}$ is V-factorizable by Lm. 4.17, and the definition does not depend on the site u. Furthermore,

$$\beta_W(\tilde{g}) = \beta_W \alpha_{W,u} \beta_W(g) = \beta_W(g)$$

by Thm. 4.15. We also know that there is at most one such density, because β_W is injective on the V-factorizable densities by definition. Hence $\tilde{g}$ is unique. □

β_W induces equivalence classes on $D(E^{V+W})$ by collecting all global densities with the same image in one class. According to Thm. 4.18 each equivalence class contains at least one V-factorizable density. Moreover, we know that there is exactly one such density, because β_W is injective on these densities by definition. It is given as the image under $\alpha_{W,i}\beta_W$ of any density in the class for any $i \in W$. Therefore it is possible to choose the V-factorizable densities as the representatives of the equivalence classes. These representatives are preserved under $\alpha_{W,i}\beta_W$ for any $i \in W$. Ex. 4.1 provides an example: g and $\tilde{g}$ are in the same equivalence class, and $\tilde{g} = \alpha_{W,0}\beta_W(g)$ is the unique V-factorizable representative of the class.

However, in general $\alpha_{W,i}(g)$ is not V-factorizable if $g \in D^W_{dBe} \backslash \text{im}\,(\beta_W)$. There is a degree of freedom in how to map a density collection to a global density on this set. We choose the arithmetic mean over all $\alpha_{W,i}$, where $i \in W$. Note that for $g \in \text{im}\,(\beta_W)$ the definition then coincides with any $\alpha_{W,i}$.

Definition 4.19. $\alpha_W : D^W_{dBe} \to D(E^{V+W})$ is given by $\alpha_W(g) = \frac{1}{|W|}\sum_{i\in W} \alpha_{W,i}(g)$.

It is clear that $\alpha_W(g)$ is a density by reasoning similar to that for $\alpha_{W,i}(g)$.

4.2.3 General Cellular Probabilistic Automata

With the de Bruijn calculus at hand we can now generalize the definition of CPA to general V. To cope with boundary conditions it is necessary that the global function operates only on $\tilde{I}$ instead of I, where $\tilde{I}$ contains the sites away from the boundary. We also have to adapt α_W and β_W to boundary conditions when they operate on the whole grid.

Definition 4.20. As before let $I = \{1, ..., m\}$, $U = \{-1, 0\}$, $V = \{-p, ..., q\}$ for $p, q, m \in \mathbb{N}_0$ with $2 + p + q \leq m$. We now set $i_l = 2 + p$, $i_r = m - q$, and $\tilde{I} = \{i_l, ..., i_r\}$.

i) We set $\hat{\alpha} : D^{\tilde{I}}_{dBe} \to D_{BC}(E^I)$, $g \mapsto \hat{\alpha}(g)$ with

$$\hat{\alpha}(g)(\psi) = \begin{cases} \alpha_{\tilde{I}}(g)(\psi|_{\{2,...,m\}}) & \text{if } \psi(1) = \rho \\ 0 & \text{else} \end{cases}$$

ii) and $\hat{\beta} : D_{BC}(E^I) \to D^{\tilde{I}}_{dBe}$, $g \mapsto \hat{\beta}(g)$ with

$$\hat{\beta}(g)(i)(\psi) = \sum_{\substack{\chi \in E^I \text{ s.t.} \\ \chi|_{i+V} = \sigma_{-i}(\psi)}} g(\chi).$$

Definition 4.21. A *cellular probabilistic automaton (CPA)* is a tuple (I, U, V, E, f_0), where for $m, p, q \in \mathbb{N}_0$ with $2 + p + q \leq m$

i) $I = \{1, ..., m\}$ is a finite grid,

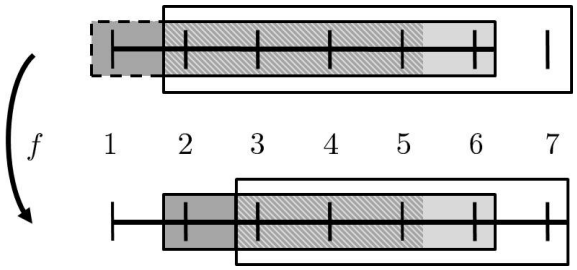


Figure 4.3: An example of a CPA with $I = \{1, ..., 7\}$ and $V = \{-2, ..., 2\}$. The global function considers patterns located at $\tilde{I} = \{4, 5\}$ and sketched in the image (lower part) by rectangles of small and large heights, respectively. The corresponding preimage patterns (upper part) are larger due to the neighborhood, and their probability of occurrence is influenced by the boundary condition (dashed part). From an implementational point of view, V may be constructed from $\tilde{V} = \{-2, ..., 1\}$ and $W = \{0, 1\}$; see Sec. 5.1.1: Focus on a pattern at site 4. The transition probability from a preimage pattern is calculated from the information about two subtransitions between light and dark grey subpatterns.

ii) $U = \{-1, 0\}$ is the neighborhood,
iii) $V = \{-p, ..., q\}$ gives rise to de Bruijn patterns,
iv) E is a finite set of local states, and
v) $f_0 : E^{U+V} \to D(E^V)$ is the local function.

With the boundary condition $\rho \in E$ the global function is given by

$$f : D_{dBe}^{\tilde{I}} \to D_{dBe}^{\tilde{I}}, \quad g \mapsto f(g),$$

$$f(g)(i)(\psi) = \begin{cases} \sum_{\varphi \in E^{U+V} \text{ s.t. } \varphi(p-1)=\rho} g(i)(\varphi|_V) \cdot f_0(\varphi)(\psi) & \text{if } i = i_l \\ \sum_{\varphi \in E^{U+V}} \alpha_U(\sigma_i(g)|_U)(\varphi|_{V+U}) \cdot f_0(\varphi)(\psi) & \text{if } i > i_l \end{cases}.$$

The trajectory starting with $g^0 \in (D(E^V))^{\tilde{I}}$ is given by the sequence $(g^n)_{n \in \mathbb{N}}$, where $g^n = f(g^{n-1})$ for $n = 1, 2, ...$

See Fig. 4.3 for a sketch of how the CPA works on general patterns. In contrast to the pattern superautomata in CNDA from Sec. 3.2.3 here we use pattern directly in the local function and not to approximate a global object after it has been constructed with minimal pattern length. Note that f_0 is not arbitrary but connected to a dynamical system with the locality property. By exploiting this relation we can assure that the global function is well-defined.

Lemma 4.22. $f(D_{dBe}^{\tilde{I}}) \subseteq D_{dBe}^{\tilde{I}}$.

Proof Since $D_{dBe}^{\tilde{I}} = (D(E^V))^{\tilde{I}}$ for $|V| = 1$, the statement is trivial in this case. So we focus on $|V| > 1$. We show that without loss of generality any pattern in the support of any site in the image can be extended to the right by a pattern in the support of the neighboring site.

Let $g \in D_{dBe}^{\tilde{I}}, i \in \{i_l, ..., i_r - 1\}$, and $\psi \in \text{supp } f(g)(i)$. By the construction of f and α_U we know that there is $\varphi \in E^{U+V}$ such that $\sigma_j(\varphi)|_V \in \text{supp}(g)(i + j)$ for all $j \in U$ and $f_0(\varphi)(\psi) > 0$. Because of the extension property of the preimage g we can find $\tilde{\varphi} \in E^{U+V}$ such that $\sigma_{-1}(\tilde{\varphi})|_V = \varphi|_V \in \text{supp } (g)(i)$ and $\tilde{\varphi}|_V \in \text{supp } (g)(i + 1)$.

This enables us to find $\tilde{\psi} \in \text{supp } f(g)(i+1)$ such that $\tilde{\psi}|_{V_-} = \sigma_1(\psi|_{V_+})$, as we will show in the following. Hence the pattern ψ may be extended to the right by $\tilde{\psi}$, and the proof is complete.

As $f_0(\varphi)(\psi) > 0$ and the partition is uniform, there is an ϵ-ball B_ϵ with respect to the 2-norm in $\mathbb{R}^{mn}$, $\epsilon > 0$, such that

$$B_\epsilon \subseteq \Omega_{\sigma_{-i}(\varphi)} \cap \Phi^{-\tau}(\Omega_{\sigma_{-i}(\psi)}).$$

Since the set is just restricted on sites $i + U + V$ due to the locality property, we may independently restrict at site $i + 1 + q$ and can still find $\epsilon' > 0$ with

$$\begin{aligned} B_{\epsilon'} &\subseteq \Omega_{\sigma_{-i}(\varphi)} \cap \Omega_{\sigma_{-(i+1)}(\tilde{\varphi}(q+s))} \cap \Phi^{-\tau}(\Omega_{\sigma_{-i}(\psi)}) \\ &\subseteq \Omega_{\sigma_{-i}(\varphi|_{V_+ + U})} \cap \Omega_{\sigma_{-(i+1)}(\tilde{\varphi}|_{q+s})} \cap \Phi^{-\tau}(\Omega_{\sigma_{-i}(\psi|_{V_+})}) \\ &= \Omega_{\sigma_{-(i+1)}(\tilde{\varphi})} \cap \Phi^{-\tau}(\Omega_{\sigma_{-i}(\psi|_{V_+})}). \end{aligned}$$

In the second line we have again used the locality property, and the equality sign holds due to $\sigma_{-i}(\varphi|_{V_+ + U}) = \sigma_{-(i+1)}(\tilde{\varphi}|_{V_- + U})$. We now define $\tilde{\psi} \in E^V$ by $\tilde{\psi}|_{V_-} = \sigma_1(\psi|_{V_+})$ and choose $\tilde{\psi}(q)$ such that there is $\epsilon'' > 0$ with

$$B_{\epsilon''} \subseteq B_{\epsilon'} \cap \Phi^{-\tau}(\Omega_{\sigma_{-(i+1)}(\psi(q))}).$$

Therefore

$$B_{\epsilon''} \subseteq \Omega_{\sigma_{-(i+1)}(\tilde{\varphi})} \cap \Phi^{-\tau}(\Omega_{\sigma_{-(i+1)}(\tilde{\psi})}),$$

$f_0(\tilde{\varphi})(\tilde{\psi}) > 0$, and $\tilde{\psi} \in \text{supp } f(g)(i+1)$. □

We denote the case of $i_l = i_r$ with $V_{\max}$ and find that $\hat{\alpha}$ and $\hat{\beta}$ are inverse in this case.

Lemma 4.23. If $i_l = i_r$,

i) $\hat{\alpha}\hat{\beta}(g) = g$ for all $g \in D_{BC}(E^I)$ and

ii) $\hat{\beta}\hat{\alpha}(g) = g$ for all $g \in (D(E^V))^{\tilde{I}}$.

Proof Claim i) Let $g \in D_{BC}(E^I)$ and $\psi \in E^I$ and note that $\sigma_{-i_l}(\sigma_{i_l}(\psi)|_V) = \psi|_{i_l+V}$. Therefore

$$\alpha_{\{i_l\}}(\hat{\beta}(g))(\psi|_{\{2,\dots,m\}}) = \hat{\beta}(g)(i_l)(\sigma_{i_l}(\psi)|_V) = \sum_{\substack{\chi \in E^I \text{ s.t.} \\ \chi|_{i_l+V} = \psi|_{i_l+V}}} g(\chi) = g(\psi).$$

If $\psi(1) = \rho$, we find that

$$\hat{\alpha}\hat{\beta}(g)(\psi) = \alpha_{\{i_l\}}(\hat{\beta}(g))(\psi|_{\{2,\dots,m\}}) = g(\psi),$$

and otherwise also

$$\hat{\alpha}\hat{\beta}(g)(\psi) = 0 = g(\psi).$$

Claim ii) Let now $g \in D(E^V)^{\tilde{I}}$ and $\psi \in E^V$. Then

$$\begin{aligned} \hat{\beta}\hat{\alpha}(g)(i_l)(\psi) &= \sum_{\substack{\chi \in E^I \text{ s.t.} \\ \chi|_{i+V} = \sigma_{-i}(\psi)}} \hat{\alpha}(g)(\chi) = \sum_{\substack{\chi \in E^{\tilde{I}} \text{ s.t.} \\ \chi|_{i+V} = \sigma_{-i}(\psi)}} \alpha_{\{i_l\}}(g)(\chi) \\ &= \sum_{\substack{\chi \in E^{\tilde{I}} \text{ s.t.} \\ \chi|_{i+V} = \sigma_{-i}(\psi)}} g(i_l)(\sigma_{i_l}(\chi)|_V) = g(i_l)(\psi). \end{aligned}$$ □

For $V_{\max}$ the evolution of the global density is calculated directly, and locality is completely omitted. It can be shown that then the CPA exactly corresponds to the discretized FPO.

Proposition 4.24. For $V = V_{\max}$ we find that $\hat{\alpha} f \hat{\beta} = P_B$.

Proof We prove that $\hat{\alpha} f \hat{\beta}(g)(\psi) = P_B(g)(\psi) = \sum_{\varphi \in E^I} g(\varphi) P_{B,\varphi,\psi}$ for all $g \in D_{BC}(E^I)$ and all $\psi \in E^I$.

If $\psi(1) \neq \rho$, the claim follows immediately. The reason is that in this case by definition $\hat{\alpha} f \hat{\beta}(g)(\psi) = 0$ and also $P_{B,\varphi,\psi} = \frac{\lambda(\Omega_\varphi \cap \Phi^{-\tau}(\Omega_\psi))}{\lambda(\Omega_\varphi)} = 0$ for all $\varphi \in E^I$ because of $\Phi^{-\tau}(\Omega_\psi) = \emptyset$.

So consider $\psi(1) = \rho$. For $\varphi \in E^I$ we calculate that $\hat{\beta}(g)(i_l)(\sigma_{i_l}(\varphi)|_V) = g(\varphi)$ and

$$\begin{aligned} f\hat{\beta}(g)(i_l)(\sigma_{i_l}(\psi)|_V) &= \sum_{\varphi \in E^I \text{ s.t. } \varphi(1)=\rho} \hat{\beta}(g)(i_l)(\sigma_{i_l}(\varphi)|_V) f_0(\sigma_{i_l}(\varphi))(\sigma_{i_l}(\psi)|_V) \\ &= \sum_{\varphi \in E^I} g(\varphi) f_0(\sigma_{i_l}(\varphi))(\sigma_{i_l}(\psi)|_V). \end{aligned}$$

To simplify the index of the sum in the first line it was observed that $E^{\{i_l\}+U+V} = E^I$ in the case at hand. Since for $\{i_l\}$ we do not have to normalize, furthermore

$$\begin{aligned} \hat{\alpha} f \hat{\beta}(g)(\psi) &= \alpha_{\{i_l\}}(f\hat{\beta}(g))(\psi|_{\{1+r,\dots,m-s\}}) = f\hat{\beta}(g)(i_l)(\sigma_{i_l}(\psi|_{\{1+r,\dots,m-s\}})|_V) \\ &= f\hat{\beta}(g)(i_l)(\sigma_{i_l}(\psi)|_V) = \sum_{\varphi \in E^I} g(\varphi) f_0(\sigma_{i_l}(\varphi))(\sigma_{i_l}(\psi)|_V) \\ &= \sum_{\varphi \in E^I} g(\varphi) \frac{\lambda(\Omega_\varphi \cap \Phi^{-\tau}(\Omega_{\psi|_{\{1+r,\dots,m-s\}}}))}{\lambda(\Omega_\varphi)}. \end{aligned}$$

If we show that $\Omega_\varphi \cap \Phi^{-\tau}(\Omega_{\psi|_{\{2,\dots,m\}}}) = \Omega_\varphi \cap \Phi^{-\tau}(\Omega_\psi)$, the proof is complete. So let $v \in \Omega_\varphi \cap \Phi^{-\tau}(\Omega_{\psi|_{\{2,\dots,m\}}})$. Then $\hat{T}(\Phi^\tau(v))|_{\{2,\dots,m\}} = \psi|_{\{2,\dots,m\}}$ and $\hat{T}(v)(1) = \rho = \psi(1)$. But therefore already $\hat{T}(\Phi^\tau(v)) = \psi$, and also $v \in \Omega_\varphi \cap \Phi^{-\tau}(\Omega_\psi)$. The other subset relation follows immediately, because $\Omega_\psi \subseteq \Omega_{\psi|_{\{2,\dots,m\}}}$. □

4.3 Consistency and Locality Errors

Up to technical postprocessing, time evolution of a density with an FPO is approximated by evolution with the according CPA. The accuracy of the approximation is determined by two parameters: state space resolution and de Bruijn pattern length. In this section we first show that for maximal pattern length the approximation can be made arbitrarily close. A subsequent investigation of potential locality errors in the case of smaller pattern length complements this consistency result.

4.3.1 Consistency

There are many ways to study distances between probability measures [63]. In our density based formulation we use the $\mathcal{L}^1$-norm for probability densities, which leads to the notion

of strong convergence in the literature [106]. A result by Li ensures that in this norm the discretized FPO converges pointwise to the original FPO for increasing state space resolution [111]. For sake of completeness we give a proof for the given case.

Theorem 4.25. Let $g \in D(\mathbb{R}^{mn})$ with $\text{supp}(g) \subseteq \Omega^m$, and let $T : \Omega \to E$ be a uniform partition with resolution $\Delta\Omega$. Then R converges pointwise to the identity with respect to the $\mathcal{L}^1$-norm,

$$\lim_{\Delta\Omega \to 0} \|R(g) - g\|_1 = 0.$$

Note that the 1-dimensional proof by Li can immediately be extended to general dimension mn of the phase space. We present an alternative proof here, which needs some tools:

A subset $M \subseteq X$ of a metric space (X, d) is called totally bounded, if for every $\epsilon > 0$ there exist $n \in \mathbb{N}$ and $x_1, ..., x_n \in M$ such that

$$M \subseteq \bigcup_{i=1}^{n} \{x \in X \mid d(x, x_i) < \epsilon\}.$$

Totally bounded subsets of $\mathcal{L}^p(\mathbb{R}^n)$ can be alternatively characterized in a functional analytical sense by the theorem of Kolmogorov-Riesz [77,182]. In the following $|.|$ denotes the 2-norm in $\mathbb{R}^n$.

Theorem 4.26. (Kolmogorov-Riesz)

A set $M \subset \mathcal{L}^p(\mathbb{R}^n)$, $1 \leq p < \infty$ is totally bounded if and only if the following criteria are fulfilled:

i) M is bounded in $\mathcal{L}^p(\mathbb{R}^n)$, i.e., $\sup_{g \in M} \|g\|_p < \infty$.

ii) For every $\epsilon > 0$ there is some R so that for every $g \in M$

$$\int_{|v|>R} |g(v)|^p dv < \epsilon^p.$$

iii) For every $\epsilon > 0$ there is $\delta > 0$ so that for every $g \in M$ and $w \in \mathbb{R}^n$ with $|w| < \delta$

$$\int_{\mathbb{R}^n} |g(v + w) - g(v)|^p dv < \epsilon^p.$$

Proof of Thm. 4.25:

$$\begin{aligned}
\|R(g) - g\|_1 &= \int_{v \in \Omega^m} \left| \sum_{\varphi \in E^I} \chi_{\Omega_\varphi}(v) \frac{1}{\Delta\Omega^{mn}} \int_{w \in \Omega_\varphi} g(w)dw - g(v) \right| dv \\
&= \int_{v \in \Omega^m} \left| \sum_{\varphi \in E^I} \chi_{\Omega_\varphi}(v) \frac{1}{\Delta\Omega^{mn}} \int_{w \in \Omega_\varphi} g(w)dw - \sum_{\varphi \in E^I} \chi_{\Omega_\varphi}(v) g(v) \right| dv \\
&\leq \int_{v \in \Omega^m} \sum_{\varphi \in E^I} \chi_{\Omega_\varphi}(v) \left| \frac{1}{\Delta\Omega^{mn}} \int_{w \in \Omega_\varphi} g(w)dw - g(v) \right| dv \\
&= \sum_{\varphi \in E^I} \int_{v \in \Omega_\varphi} \left| \frac{1}{\Delta\Omega^{mn}} \int_{w \in \Omega_\varphi} g(w)dw - g(v) \right| dv
\end{aligned}$$

$$\leq \sum_{\varphi \in E^I} \frac{1}{\Delta\Omega^{mn}} \int_{v \in \Omega_\varphi} \int_{w \in \Omega_\varphi} |g(w) - g(v)| \, dwdv$$
$$= \sum_{\varphi \in E^I} \frac{1}{\Delta\Omega^{mn}} \int_{v \in \Omega_\varphi} \int_{u \in \Omega_\varphi - v} |g(v+u) - g(v)| \, dudv$$

The last step involves a change of variables from w to $u := w - v$, and $\Omega_\varphi - v := \{u - v \,|\, u \in \Omega_\varphi\}$. As

$$\Omega_\varphi - v \subseteq [-\Delta\Omega, \Delta\Omega]^{mn}$$

for $v \in \Omega_\varphi$ with $\varphi \in E^I$, we calculate with Fubini's theorem

$$\begin{aligned} \|R(g) - g\|_1 &\leq \sum_{\varphi \in E^I} \frac{1}{\Delta\Omega^{mn}} \int_{v \in \Omega_\varphi} \int_{u \in [-\Delta\Omega, \Delta\Omega]^{mn}} |g(v+u) - g(v)| \, dudv \\ &= \frac{1}{\Delta\Omega^{mn}} \int_{v \in \Omega^m} \int_{u \in [-\Delta\Omega, \Delta\Omega]^{mn}} |g(v+u) - g(v)| \, dudv \\ &\leq \frac{1}{\Delta\Omega^{mn}} \int_{u \in [-\Delta\Omega, \Delta\Omega]^{mn}} \int_{v \in \mathbb{R}^{mn}} |g(v+u) - g(v)| \, dvdu. \end{aligned}$$

Let $\epsilon > 0$. Since the set $\{g\} \subset \mathcal{L}^1(\mathbb{R}^n)$ is totally bounded, the theorem of Kolmogorov-Riesz, Thm. 4.26, guarantees that there is $\delta > 0$ such that

$$\int_{v \in \mathbb{R}^{mn}} |g(v+u) - g(v)| \, dv < \frac{\epsilon}{2^{mn}}$$

for $|u| < \delta$. If we choose $\Delta\Omega$ such that $\Delta\Omega < \frac{\delta}{\sqrt{mn}}$ we ensure that

$$|u| = \sqrt{\sum_{i=1}^{mn} |u_i|^2} \leq \sqrt{\sum_{i=1}^{mn} \Delta\Omega^2} = \sqrt{mn}\Delta\Omega < \delta$$

for all $u \in [-\Delta\Omega, \Delta\Omega]^{mn}$ and hence

$$\|R(g) - g\|_1 < \frac{1}{\Delta\Omega^{mn}} \int_{u \in [-\Delta\Omega, \Delta\Omega]^{mn}} \frac{\epsilon}{2^{mn}} du = \epsilon.$$

Therefore $\|R(g) - g\|_1 \to 0$ for $\Delta\Omega \to 0$. □

With Prop. 4.24 we thus find that for maximal pattern length $|V_{\max}|$ and a uniform partition with resolution $\Delta\Omega$ our algorithm is consistent: For $g \in \kappa_B^{-1}(D_{BC}(E^I))$,

$$\begin{aligned} &\lim_{\Delta\Omega \to 0} \|\kappa_B^{-1} \hat{\alpha} f \hat{\beta} \kappa_B(g) - P_{\Phi^\tau}(g)\|_1 \\ &= \lim_{\Delta\Omega \to 0} \|\kappa_B^{-1} P_B \kappa_B(g) - P_{\Phi^\tau}(g)\|_1 \\ &= \lim_{\Delta\Omega \to 0} \|R P_{\Phi^\tau}(g) - P_{\Phi^\tau}(g)\|_1 \\ &= 0. \end{aligned}$$

4.3.2 Locality Errors

Now we investigate in more depth the role of locality in approximating the discretized FPO P_B by a CPA. It turns out that the CPA covers the dynamics of the underlying P_B if only the support is considered. This is expected in light of Lm. 3.30, as CPA are constructed in analogy to pattern superautomata of CNDA.

Lemma 4.27. For all $n \in \mathbb{N}$, all $g \in D_{BC}(E^I)$, and all $i \in \tilde{I}$ it holds that

$$\text{supp}\,(\hat{\beta}P_B^n(g)(i)) \subseteq \text{supp}\,(f^n\hat{\beta}(g)(i)).$$

Proof Let $n \in \mathbb{N}$, $g \in D_{BC}(E^I)$, $i \in \tilde{I}$, and $\chi \in \text{supp}\,(\hat{\beta}P_B^n(g)(i)) \in E^V$. Then there is $\psi \in E^I$ such that $P_B^n(g)(\psi) > 0$ and $\sigma_i(\psi)|_V = \chi$. Let $\varphi_n = \psi$. Per induction it can be shown that we can find $\varphi_0, ..., \varphi_{n-1} \in E^I$ such that $P_{B,\varphi_{k-1},\varphi_k} = \frac{\lambda(\Omega_{\varphi_{k-1}} \cap \Phi^{-\tau}(\Omega_{\varphi_n}))}{\lambda(\Omega_{\varphi_{n-1}})} > 0$ and $P_B^{k-1}(g)(\varphi_{k-1}) > 0$ for $k \in \{1, ..., n\}$. Since for all $j \in \tilde{I}$

$$\Omega_{\varphi_{k-1}} \cap \Phi^{-\tau}(\Omega_{\varphi_k}) \subseteq \Omega_{\varphi_{k-1}|_{j+U+V}} \cap \Phi^{-\tau}(\Omega_{\varphi_n|_{k+U+V}}),$$
$$\Omega_{\varphi_{k-1}} \subseteq \Omega_{\varphi_{k-1}|_{j+U+V}},$$

we conclude that $f_0(\sigma_j(\varphi_{k-1})|_{U+V})(\sigma_j(\varphi_k)|_V) > 0$ for all $j \in \tilde{I}$.

Furthermore, $\hat{\beta}(g)(j)(\sigma_j(\varphi_0)) > 0$ for all $j \in \tilde{I}$, and therefore $f\hat{\beta}(g)(j)(\sigma_j(\varphi_1)|_V) > 0$ for all $j \in \tilde{I}$. This induces $f^2\hat{\beta}(g)(j)(\sigma_j(\varphi_2)|_V) > 0$ for all $j \in \tilde{I}$ and so on, and therefore $f^n\hat{\beta}(g)(j)(\sigma_j(\varphi_n)|_V) > 0$ for all $j \in \tilde{I}$. Recalling that $\sigma_i(\varphi_n)|_V = \chi$, we conclude that $\chi \in \text{supp}\,(f^n\hat{\beta}(g)(i))$. □

However, we cannot recover the precise global behavior of the discretized FPO from a CPA in general. The errors that can occur are twofold, and we will provide examples for both types here. On the one hand, it may happen that correlations over $|V|$ sites are not preserved, because we work on patterns of length $|V|$. This is independent of the actual dynamics and a direct consequence of our approximation space; see Sec. 4.2.2. On the other hand, we will see that even for $U \subseteq V$ in general there are locally allowed transitions of a global state that are not allowed in a global consideration with P_B. This is remarkable, since such behavior was ruled out for the underlying dynamical system by the locality property. This may also lead to errors. While the first error type is a true locality effect, the second arises from the interplay of locality and state space discretization. We note that the first type has in principle also been studied in the context of pattern superautomata for CNDA in Sec. 3.2.3, but the second type is due to the construction of the objects from an underlying dynamical system and new at this point.

Example 4.2. This example shows that in general correlations over $|V|$ sites are not preserved. We compare one CPA time step to one time step with the discretized FPO. Consider $I = \{1, 2, 3, 4\}$ and the dynamical system that is given by the identity on $\Omega^m = [0, 1]^4$ with $U = \{-1, 0\}$. We choose the partition $\Omega_0 = [0, 0.5)$ and $\Omega_1 = [0.5, 1]$, i.e., $E = \{0, 1\}$, and look at the CPA with $V = \{0, 1\}$ and $\tilde{I} = \{2, 3\}$. We find that $f_0(\varphi)(\psi) = \delta_{\varphi|_V,\psi}$ for all $\varphi \in E^{U+V}$ and $\psi \in E^V$.

Consider $\rho = 0$ and a density $\hat{g} \in D_{BC}(E^I)$ as the extension of $g \in D(E^{V+W})$ in Ex. 4.1 by the boundary condition, i.e., such that $\hat{\beta}(\hat{g}) = \beta_{\tilde{I}}(g)$. Then also $f\hat{\beta}(\hat{g}) = \beta_{\tilde{I}}(g)$, and so

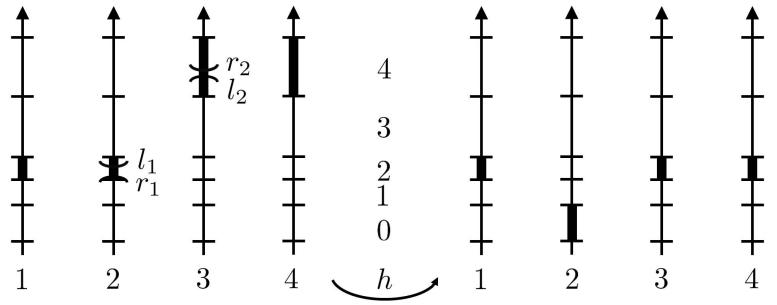


Figure 4.4: Illustration of a transition from $\chi = (2, 2, 4, 4)$ to $\psi = (2, 0, 2, 2)$ in Ex. 4.3. The left side shows the preimage, the right the image state. The horizontal numbers correspond to the respective sites, while the vertical numbers display $E = \{0, ..., 4\}$. The states χ and ψ are marked by black rectangles at the corresponding sites.

$\hat{\alpha} f \hat{\beta}(g)$ is the extension of $\tilde{g}$ from Ex. 4.1 by the boundary condition. However, $P_B(\hat{g}) = \hat{g}$, and so $\hat{\alpha} f \hat{\beta}(\hat{g}) \neq P_B(\hat{g})$.

Example 4.3. This example shows that transitions at different sites are not independent in general. By comparing one CPA time step to one time step with the discretized FPO we see that a specific local transition at one site cannot take place if another specific local transition happens at a neighboring site, although both transitions are allowed separately. Consider $I = \{1, ..., 4\}$ and the system on dynamically invariant state space $\Omega^m = [0, 1]^4$ given for all $n \in \mathbb{N}$ by $v_i^{n+1} = h(v_{i-1}^n, v_i^n) = \frac{v_{i-1}^n + v_i^n}{3.75}$ for $i \in \{2, ..., m\}$. We define 5 intervals

$$\Omega_j = [w_j, w_{j+1}), \quad \text{for } j \in \{0, ..., 3\}, \quad \Omega_4 = [w_4, w_5]$$

with

$$w_0 = 0, \quad w_1 = 0.183, \quad w_2 = 0.31, \quad w_3 = 0.4, \quad w_4 = 0.7, \quad w_5 = 1,$$

name them by their index, and obtain a partition of Ω with $E = \{0, ..., 4\}$; see Fig. 4.4. The induced flow is denoted by Φ^1 for one time step. We consider the CPA with $V = U = \{-1, 0\}$ for deterministic input $g \in D_{BC}(E^I)$ given by $g(\varphi) = \delta_{\chi,\varphi}$, where $\chi = (2, 2, 4, 4) \in E^I$. We focus on the image state $\psi = (2, 0, 2, 2) \in E^I$ and determine

$$l_1 = 0.3625, \quad l_2 = 0.8, \quad r_1 = 0.32375, \quad r_2 = 0.83875$$

as the solution of the equations

$$h(l_2, w_4) = w_3, \quad h(l_1, l_2) = w_2, \quad h(w_2, r_1) = w_1, \quad h(r_1, r_2) = w_2.$$

It is possible to show that

$$\{\Omega_{\chi|_{4+U+V}} \cap \Phi^{-1}(\Omega_{\psi|_{4+V}})\} \subseteq \{v \in \Omega \,|\, l_1 \leq v_2 \leq w_3, w_4 \leq v_3 \leq l_2\},$$
$$\{\Omega_{\sigma_{-1}(\chi|_{3+U+V})} \cap \Phi^{-1}(\Omega_{\sigma_{-1}(\psi|_{3+V})})\} \subseteq \{v \in \Omega \,|\, w_2 \leq v_2 \leq r_1, r_2 \leq v_3 \leq 1\},$$

$f_0(\sigma_4(\chi|_{4+U+V}))(\sigma_4(\psi|_{4+V}))) > 0$, and $f_0(\sigma_3(\chi|_{3+U+V}))(\sigma_3(\psi|_{3+V})) > 0$.

Hence $\hat{\alpha} f \hat{\beta}(g)(\psi) > 0$, but

$$\begin{aligned} P_B(g)(\psi) = \sum_{\varphi \in E^I} g(\varphi) P_{B,\varphi,\psi} &= \frac{\lambda(\Omega_\chi \cap \Phi^{-1}(\Omega_\psi))}{\lambda(\Omega_\chi)} \\ &\leq \frac{\lambda((\Omega_{\chi|_{4+U+V}} \cap \Phi^{-1}(\Omega_{\psi|_{4+V}})) \cap (\Omega_{\chi|_{3+U+V}} \cap \Phi^{-1}(\Omega_{\psi|_{3+V}})))}{\lambda(\Omega_\chi)} \\ &= \frac{\lambda(\emptyset)}{\lambda(\Omega_\chi)} = 0, \end{aligned}$$

and so $\hat{\alpha} f \hat{\beta}(g) \neq P_B(g)$.

Both examples are scalable in the sense that we can find analogous partitions of $[0, c], c \in (0, 1)$, with the above properties by dividing all phase space coordinates by c and complete the partition in $[c, 1]$ arbitrarily. So for decreasing size of the coding domains we can still find a partition of $[0, 1]$ with the above effects: locality errors are independent from resolution errors.

4.4 Conclusion and Outlook

We have introduced a numerical scheme for density based uncertainty propagation in distinct PDEs. The PDE is translated into a cellular probabilistic automaton (CPA), which then approximately evolves a probability density for given boundary conditions. The translation is based on state space discretization like in set oriented numerics and on the de Bruijn state idea from cellular automata theory. There are two parameters that allow to control the approximation of the exact density evolution: state space resolution and de Bruijn pattern length. We have presented the method for uncertain initial conditions under deterministic dynamics and shown consistency.

In the next chapter we extend the method to white noise boundary conditions as an example of more involved stochastic influence. We also consider implementational issues and validate the algorithm in several water grid applications. In Chap. 6 then we adapt the CPA framework to include measurement data in the knowledge about a system's state by establishing a connection to dynamic Bayesian networks.

From a theoretical perspective two main questions remain for future research. First, we are interested in really quantifying and improving the approximation error for given de Bruijn pattern length beyond our consistency result. This appears to be a demanding problem, because spatial correlations have to be quantified somehow. Second we suggest to focus on random parameters beyond initial value uncertainties and the white noise boundary conditions introduced in the next chapter. It seems difficult to preserve temporal correlations in such parameters with our algorithm, but our ideas can be used for white noise parameters. We are confident that white noise parameters only extend the translation of PDEs to CPA, whereas the simulation itself is not changed. In this sense CPA promise to overcome the curse of dimension in parameter space.

5 Applications and Extensions of Cellular Probabilistic Automata

In the previous chapter cellular probabilistic automata (CPA) have been theoretically introduced as a method for uncertainty propagation. This chapter complements the introduction with a practical perspective. It contains implementational considerations, applications and adaptations or extensions of the method. Especially the interpretation of the discrete state space in terms of risk levels is highlighted in the applications. Notation and results of the theory chapter are used here.

The chapter is divided into 5 sections. We start in Sec. 5.1 with considerations about the implementation of CPA. In particular we suggest a computational method to calculate the local function of a CPA, conduct a complexity analysis and deduce some practical tips for the design of an efficient algorithm. Furthermore we extend the CPA concept slightly so that CPA can cope with stochastic boundary conditions. We want to provide evidence this way that with CPA the treatment of more general stochastic spatio-temporal processes seems feasible. They are used in Sec. 5.2, where CPA are applied to advection-reaction equations for modeling contamination in water pipes. In particular we study two examples, arsenate adsorption to the pipe wall and bacterial regrowth with chlorine inhibition. The results are interpreted and validated with a Monte Carlo simulation. However, in further computer experiments for longer pipes we observe difficulties with conservation laws for CPA simulations. We propose an adaptation of the method in Sec. 5.3 to encounter these problems. In Sec. 5.4 we then extend the method to the simulation of contamination in whole water grids. The main contribution of this section is a way to deal with coupled partial differential equations (PDEs) in a discrete setup. Finally, in Sec. 5.5 we address more complex hyperbolic PDEs, in particular the isothermal Euler equation of fluid dynamics.

5.1 Implementation of Cellular Probabilistic Automata

From an implementational point of view two steps of uncertainty propagation with CPA have to be distinguished. Step one is the translation of the completely continuous system into a CPA. This is independent of initial or boundary conditions and can be achieved in a preprocessing procedure. Step two consists of the CPA evolution with given initial and boundary values. In Sec. 5.1.1 we suggest a method to conduct step one in practice. Afterwards a complexity analysis of both steps is given in Sec. 5.1.2. In Sec. 5.1.3 practical implications for the implementation are discussed; in particular we suggest a sparse implementation that can improve the performance significantly. Although the CPA theory has been developed for uncertainties in initial conditions, in Sec. 5.1.4 we extend the concept slightly such that CPA can cope with certain stochastic boundary conditions.

5.1.1 Calculating Local Transition Probabilities

The preprocessing basically consists of the approximation of local transition probabilities. We propose a local version of the standard Monte Carlo quadrature approach in set oriented numerics for this purpose [87]. We note that also advanced adaptive methods have been suggested; see e.g. [73].

i) For $\varphi \in E^{U+V}$ choose W_φ test vectors $w_i = (w_{i,-r-p}, ..., w_{i,s+q}) \in (\mathbb{R}^n)^{U+V}$, where $\{w_{i,j}\}_{i \leq W_\varphi}$ is randomly distributed over coding domain $\Omega_{\varphi(j)} \subseteq \Omega$, respectively.

ii) Compute for all $i \leq W_\varphi$ the image points

$$\tilde{w}_i = (h(\tau, w_{i,-r-p}, ..., w_{i,s-p}), h(\tau, w_{i,-r-p+1}, ..., w_{i,s-p+1}), ..., h(\tau, w_{i,-r+q}, ..., w_{i,s+q})).$$

iii) Determine $\psi_1, ..., \psi_L \in E^V$ such that there is $l \leq L$ and $\tilde{w}_i$ with $T((\tilde{w}_i)_j) = (\psi_l)_j$ for all $j \in V$. Let the number of image points in the specific coding domain be denoted by W_{ψ_l}, i.e., $\sum_{l=1}^{L} W_{\psi_l} = W_\varphi$. The local transition function is then approximated by

$$f_0(\varphi)(\psi) = \begin{cases} W_\psi / W_\varphi & \text{for all } \psi \in \{\psi_1, ..., \psi_L\} \\ 0 & \text{else} \end{cases}.$$

We will see in Sec. 5.1.2 that the number of numerical map evaluations grows exponentially in the patterns length, which is a problem for large V. We suggest to use de Bruijn calculus to determine transition probabilities for large V by concatenating transition probabilities for smaller $\tilde{V}$; see Fig. 4.3. For given $\tilde{f}_0 : E^{U+\tilde{V}} \to D(E^{\tilde{V}})$ and W given by $V = \tilde{V} + W$,

$$f_0 : E^{V+U} \to D(E^V), \quad \varphi \mapsto f_0(\varphi),$$

where

$$f_0(\varphi)(\psi) = \alpha_W(\hat{g})(\psi),$$
$$\hat{g} \in (D(E^{\tilde{V}}))^W, \quad \hat{g}(i) = \tilde{f}_0(\varphi|_{i+\tilde{V}+U}).$$

It can be shown with an example similar to Ex. 4.3 that this is again just an approximation of the directly calculated f_0.

5.1.2 Complexity Analysis

In this section we analyze the complexity of the CPA algorithm in three steps. First, we consider the concatenation of patterns, second the preprocessing as introduced in Sec. 5.1.1 and third the simulation. For simplicity we assume an identical number of symbols in each dimension of the phase space, but the ideas are easily extendable to the general case. We use the following notation in this section:

i		...	# test points per domain	$m = \|I\|$	...	# sites,
j	$= \|V\|$	...	length of pattern,	n	...	dimension of phase space
k	$= \|W\|$	...	length of pattern extension,	p	...	# domains per dimension,
l	$= \|U\|$	...	length of neighborhood,	t	...	# time steps

5.1.2.1 Concatenating Patterns

Many repetitive tasks have to be performed when designing an algorithm for the determination of $\alpha_W(g)$, where $\alpha_W : (D(E^V))^W \to D(E^{V+W})$ and $g \in (D(E^V))^W$. The idea of *dynamic programming* is to perform each repetitive computation just once, store the result and look it up every time it is needed [36]. It hence optimizes calculation time at the cost of memory. Depending on how much one uses this concept, one can control the tradeoff between efficiency in time and memory.

α_W is defined as the mean over all $\alpha_{W,a}$ for $a \in W$. For $\psi \in E^{V+W}$ each $\alpha_{W,a}(g)(\psi)$ is given as a fraction; see also the proof of Lm. 4.14. The numerator is a product of k factors that is independent of a, and the denominator is a factor of probabilities from left and right marginal densities of the local densities $(D(E^V))^W$:

$$\alpha_{W,a}(g)(\psi) = \frac{\prod_{b\in W} g(b)(\sigma_b(\psi)|_V)}{\prod_{b=-\hat{t}}^{a-1} g_l(b)(\sigma_b(\psi)|_{V_+}) \prod_{b=a+1}^{\hat{u}} g_r(b)(\sigma_b(\psi)|_{V_-})},$$

where we use the notation of Sec. 4.2.2 with $W = \{-\hat{t}, ..., \hat{u}\}$. The left and right marginal densities $g_{l/r} \in D(E^{V_{+/-}})$ are given by

$$g_{l/r}(b)(\varphi) = \sum_{\chi\in E^V \text{ s.t.} \chi|_{V_{+/-}}=\varphi} g(b)(\chi).$$

For global states E^{V+W} that have some subpatterns in common, the same marginal probabilities are needed. In the sense of dynamic programming we therefore determine all $k-1$ left and right marginal densities at all sites in W in advance. The complete algorithm then consists of 4 steps.

First, we compute all marginal densities in a straightforward way: The probability for each of the $p^{n(j-1)}$ states in a marginal distribution at one site can be calculated by p^n additions and $(j-1)p^{nj}$ comparisons of states at single sites. This leads to algorithmic complexity $O(kjp^{2nj-n})$ of the first step, because the latter term dominates the former asymptotically. In a second step we determine the numerators for all global states with $(k-1)p^{n(j+k-1)}$ multiplications. The algorithmic complexity for the secon step is hence $O(kp^{n(j+k-1)})$. The third step then calculates the mean of the denominators over all $a \in W$ for all global states. For each global state this calculation requires averaging over k terms, of which each is calculated by multiplication of $k-1$ marginal probabilities. So this step leads to complexity $O(k^2p^{n(j+k-1)})$. In the last step we determine the fraction of the numerators with the respective denominators in $O(p^{n(j+k-1)})$ divisions.

One might improve the efficiency in all steps further by lower-level dynamic programming. We use the second step for $V = \{0\}$, i.e., $j = 1$, as an example: subpatterns can be grown iteratively from single site states to global states, and new single state probabilities are multiplied to probabilities of the subpatterns. This requires all together $p^{2n} + ... + p^{kn}$ multiplications, which asymptotically leads to $O(p^{kn})$ instead of $O(kp^{kn})$ multiplications.

Furthermore we note that in practical applications it might be enough to only calculate with $\alpha_{W,i}$ for a specific $i \in W$ instead of the average over all of them. The results for the different sites are in many cases not too different from each other, and thus a lot of the numerical effort can be avoided.

5.1.2.2 Preprocessing: Local Function

Calculating $f_0 : D(E^{U+V}) \to D(E^V)$ requires to map i^{j+l-1} local point combinations and determine their respective image domain for $p^{n(j+l-1)}$ preimage patterns. This amounts to $p^{n(j+l-1)}i^{j+l-1}$ calls of the numerical solver to map the local point combinations and $jnp^{n(j+l-1)}i^{j+l-1}$ modulo calculations in the case of a uniform partition to determine the image domains. So our standard preprocessing algorithm has complexity $O(jnp^{n(j+l-1)}$ $i^{j+l-1})$.

The algorithm faces the well known curse of dimension of set-oriented methods: it scales exponentially bad in phase space dimension, pattern and neighborhood length. However, the calculations for all preimage patterns can be done in parallel, potentially on a cluster. Furthermore we can use the more advanced technique of Sec. 5.1.1 for large patterns and determine transition probabilities between them by merging transition probabilities for smaller $\tilde{V}$. If the partition is not uniform, more complex algorithms have to be applied instead of the modulo evaluations.

5.1.2.3 Simulation: Global Function

In the simulation we have to apply the global function $f : D(E^V)^{\tilde{I}} \to D(E^V)^{\tilde{I}}$ in every time step. Boundary effects are neglected here for simplicity. First, at every site in $\tilde{I}$ the preimages have to be concatenated in a neighborhood with α_U, and second, the local function has to be applied to it. The first step can be conducted in tm times the whole complexity of α_U, and for the second step for all image and all preimage patterns multiplications and additions of probabilities are necessary. This can be achieved in $O(mtp^{n(j+l-1)}p^{nj}) = O(mtp^{n(2j+l-1)})$. Like in the preprocessing we find the curse of dimension of set-oriented numerics, i.e., the exponentially bad scaling in phase space dimension, pattern and neighborhood length. However, the simulation scales linearly in space and time - like a conventional CA it is even parallelizable in space. In practice it turns out that the preprocessing is numerically more expensive than the simulation. So for industrial applications the CPA method points towards real-time uncertainty quantification, because the slow preprocessing only has to be performed once before the actual simulation. The method is especially interesting when many simulations with different initial and boundary conditions have to be performed.

5.1.3 Practical Considerations

The performance of the CPA algorithm depends very much on implementational details. It turns out that some considerations for the efficient implementation of algorithms for probabilistic graphical models are also relevant for the implementation of CPA. A basic decision is, for example, whether the multi-dimensional arrays, e.g. the local function or the collections of local densities, are flattened for storage or not [102]. Also for multi-dimensional phase spaces it has to be decided whether to use one- or multi-dimensional symbols.

In this section we focus on sparsity to encounter the exponentially bad scaling in phase space dimension, pattern and neighborhood length. Similar ideas have been used in the context of filtering in dynamic Bayesian networks [17]. We only store and process states

with probability larger than a specified threshold whenever possible. Otherwise already for reasonably high discretization and pattern length the calculations are not feasible. Instead of an array of probabilities, where indices correspond to states, we store an array of non-zero probabilities together with the according states. Then the indices do not have a meaning any more, but the information is compressed by orders of magnitude. We apply this idea in the preprocessing, the concatenation and the simulation.

In the preprocessing we can make up for the loss of index information by saving in an additional list how many image states are allowed for each preimage pattern. The number of allowed image states is typically much smaller than E^V. Transition probabilities for a given preimage state can then be easily accessed by using the strides as artificial indices. However, in concatenation and simulation in the worst case the sparse computation may require more operations than with the non-sparse algorithm investigated above. The reason is that in many steps of the algorithm (subpatterns of) states have to be compared e.g. for concatenation, because we do not know the content of the arrays in advance. Obeying a predetermined order for states in the arrays can relieve this problem, a requirement that deserves special caution during parallelization. In any case, usually the set of states with positive probability is much smaller than E^V, although it typically first grows and then shrinks again in the transient phase of dynamics. Therefore normally sparsity leads to implementations that are orders of magnitude faster than conventional implementations. We note that our de Bruijn choice of α_W enables such sparse calculations, whereas the whole space is needed to solve, for example, a linear nonnegative least squares problem.

5.1.4 Stochastic Boundary Conditions

The CPA theory has only been developed for uncertainties in initial conditions. Here we extend the concept slightly to stochastic boundary conditions. This is, for example, important in contaminant transport modeling [177], and we will directly apply the ideas in Sec. 5.2. With this first generalization we want to provide evidence that with CPA the treatment of more general stochastic spatio-temporal processes seems feasible.

We use the notation of Chap. 4 and again only consider $U = \{-1, 0\}$. Stationary temporal white noise boundary conditions $g_l \in D(E)$ now replace the deterministic $\rho \in E$. With stationary we mean that the densities do not change in time, and the term temporal white noise indicates that there are no correlations in the boundary random variable's realizations at different times. We can say that the set of boundary condition random variables at each time step is iid.

For this purpose the global function in Def. 4.21 is extendend to $f : D_{dBe}^{\tilde{I}} \to D_{dBe}^{\tilde{I}}, \quad g \mapsto f(g)$,

$$f(g)(i)(\psi) = \begin{cases} \sum_{\varphi \in E^{U+V}} g_l(i)(\varphi(-p-1)) \cdot g(i)(\varphi|_V) \cdot f_0(\varphi)(\psi) & \text{if } i = i_l \\ \sum_{\varphi \in E^{U+V}} \alpha_U(\sigma_i(g)|_U)(\varphi|_{V+U}) \cdot f_0(\varphi)(\psi) & \text{if } i > i_l \end{cases} .$$

The relations between global and de Bruijn densities have to be generalized in the stochastic case to $\hat{\alpha} : D(E) \times D_{dBe}^{\tilde{I}} \to D(E^I)$, $(g_l \times g) \mapsto \hat{\alpha}((g_l \times g))$ with

$$\hat{\alpha}((g_l \times g))(\psi) = g_l(\psi(1))\alpha_{\tilde{I}}(g)(\psi|_{\{2,\dots,m\}})$$

and $\hat{\beta} : D(E^I) \to D(E) \times D_{dBe}^{\tilde{I}}$, $g \mapsto \hat{\beta}(g) = (g_l, g_{\tilde{I}})$ with

$$g_l(e) = \sum_{\chi \in E^I \text{ s.t. } \chi(1)=e} g(\chi), \qquad g_{\tilde{I}}(i)(\psi) = \sum_{\substack{\chi \in E^I \text{ s.t.} \\ \chi|_{i+V} = \sigma_{-i}(\psi)}} g(\chi).$$

CPA with stochastic boundary conditions may be used to approximate spatio-temporal processes with deterministic dynamics, in which the initial and boundary conditions are stochastic. We note that it is straightforward to use time-dependent stochastic boundary conditions instead of stationary ones.

5.2 Application I: Advection-Reaction Equations

Our first application of uncertainty propagation with CPA is the class of advection-reaction equations. The theory in Chap. 4 has been formulated exactly for initial value uncertainties in this class. However, to show that CPA can be used to deal with more general stochastic influence we apply the extension to stochastic boundary conditions of Sec. 5.1.4 right away for contaminant sources [177].

We choose two standard problems of contamination in drinking water pipes, that have attracted a lot of attention in the water supply community lately [150]. The first is the rather simple but already non-linear arsenate adsorption, for which we investigate the different CPA parameters in depth. Here we also perform a Monte Carlo calculation to compare our results with. The second problem is the more complex bacterial regrowth with chlorine inhibition, for which we investigate a special case.

5.2.1 Adsorption of Arsenate

Consider the advection and adsorption of arsenate in drinking water pipes, a standard problem in the water safety community [101, 150]. This application will serve as the standard test problem also in later sections. In this section we describe a water tank on a hill and a pipe to a consumer in a valley. Report locations to observe the arsenate concentrations are installed in a distance of Δx; see Fig. 5.1a. The physics is described by the Langmuir adsorption model [103]

$$\begin{aligned} \partial_t D + v \partial_x D &= -\frac{1}{r_h} f(D, A), \\ \partial_t A &= f(D, A) \end{aligned} \tag{5.1}$$

with

$$f(D, A) = \frac{1}{\frac{1}{k_1} + \frac{1}{k_f}(S_{\max} - A)} (D(S_{\max} - A) - K_{\text{eq}} A),$$

where D is the concentration of dissolved arsenate and A the concentration of arsenate adsorbed at the pipe wall. We adopt realistic parameter values from [101, 142]

$$\begin{aligned} v &= 10\frac{\mathrm{m}}{\mathrm{min}}, & r_h &= 50\frac{l}{\mathrm{m}^2}, \\ k_1 &= 0.2\frac{\mathrm{l}}{\mathrm{mg\,min}}, & S_{\max} &= 100\frac{\mathrm{mg}}{\mathrm{m}^2}, \\ K_{\mathrm{eq}} &= 0.0537\frac{\mathrm{mg}}{l}, & k_f &= 2.4\frac{l}{\mathrm{m}^2\,\mathrm{min}}, \end{aligned} \tag{5.2}$$

and consider the system on the approximately positively invariant Ω given by $D \in [0,1]\frac{\mathrm{mg}}{l}$ and $A \in [0,100]\frac{\mathrm{mg}}{\mathrm{m}^2}$. To discretize time and space we decouple the advection and the reaction step with the Trotter formula [163]. For the advection the method of characteristics [109] is used with $U = \{-1, 0\}$ as the backward difference, $\Delta x = 100\mathrm{m}$ and $\Delta t = \Delta x / v = 10\mathrm{min}$:

$$\begin{aligned} D_i^{n+1} &= D_{i-1}^n - \Delta t \frac{1}{r_h} f(D_{i-1}^n, A_i^n) \\ A_i^{n+1} &= A_i^n + \Delta t f(D_{i-1}^n, A_i^n) \end{aligned} \tag{5.3}$$

To obtain the local function of a CPA we map test points by using intermediate steps with a less coarse discretization $\Delta x' = 1\mathrm{m}$ and $\Delta t' = 0.1\mathrm{min}$. For this purpose we assume continuity and linearly interpolate the states at the sites between the ends from the states at the ends. We use 75 randomly distributed test points for the coding domain at site 0 and 37 at site -1 to approximate the local function, and a probability threshold of 0.00005. We use $V = \tilde{V} = \{0\}$ and partition the phase space equidistantly with 5 symbols in each of the $n = 2$ directions. If we label the coding domains from 0 to 4 in each direction, the corresponding CPA results from transition probabilities like

$$f_0(((1,4),(2,4)))(\psi) = \begin{cases} 0.806 & \text{if } \psi = (1,4) \\ 0.194 & \text{if } \psi = (2,4) \\ 0 & \text{else} \end{cases};$$

see Fig. 5.1b. White noise boundary conditions are applied to describe a random arsenate source in the tank, and deterministic initial values represent a pipe which is completely empty in the beginning. Note that we can choose the A-direction of the boundary condition at will, since it has no influence on the system. The observed dynamics is shown in Fig. 5.1c-5.1e: Dissolved arsenate is transported along the pipe, and over time the walls are covered more and more with adsorbed arsenate. After 24 hours a steady state is reached, and we compare it to a Monte Carlo calculation; see Fig. 5.2a-5.2b. The latter has been obtained on the basis of the same time and space discretization with $\Delta t' = 0.1\mathrm{min}$ and $\Delta x' = 1\mathrm{m}$ for 20000 evaluations. The boundary condition has been drawn from the stationary boundary distribution every $10\mathrm{min}$ and held constant in the meantime. Our example features an interesting probabilistic effect due to the nonlinearity of the reaction equations. Although the boundary values are distributed in the D-domains $1-3$, the consumer mostly observes dissolved arsenate at a concentration of domain 2 in the steady

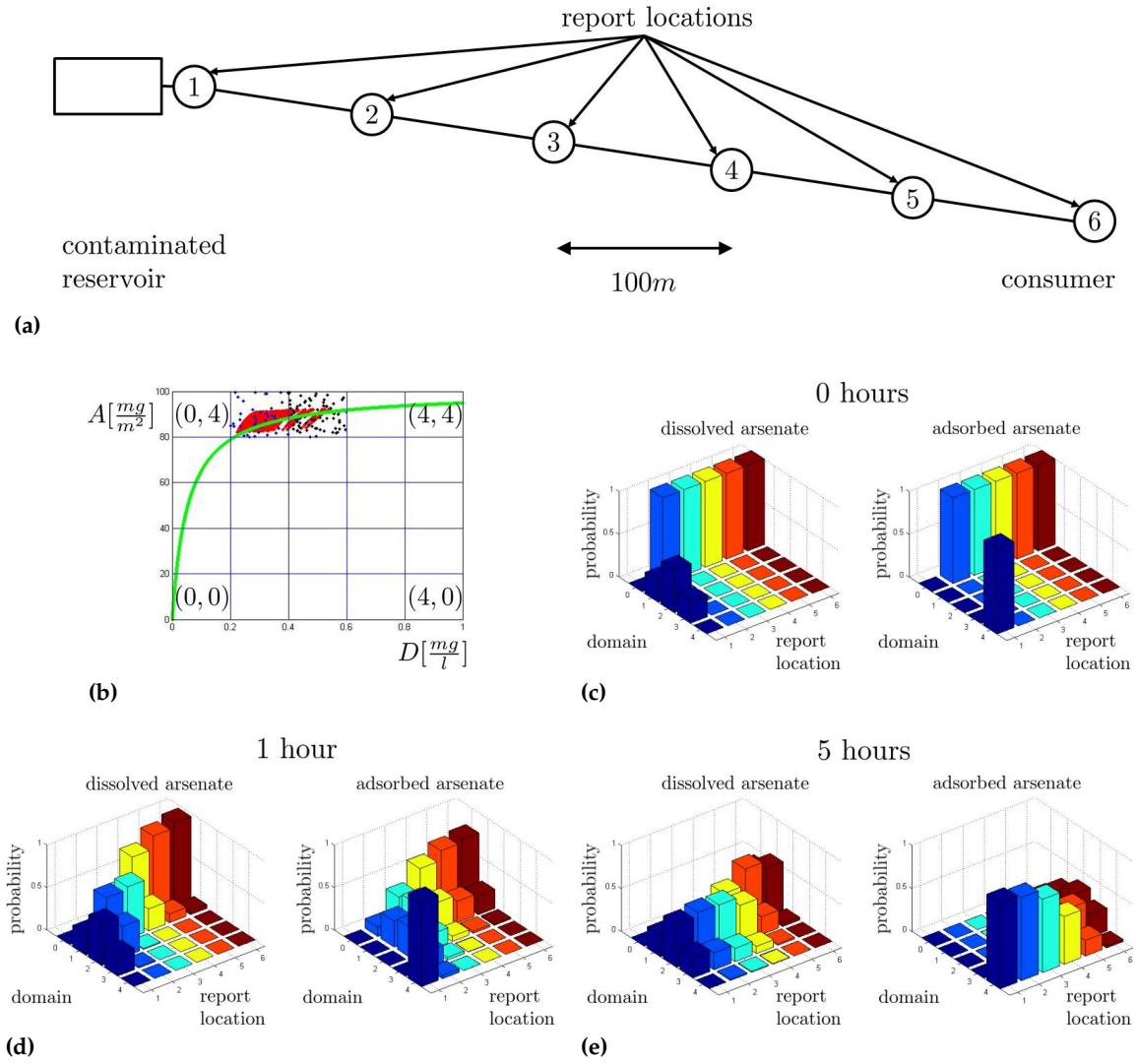


Figure 5.1: (a) A reservoir on a hill is connected to a consumer in a valley through a pipe with 6 report locations. (b) The phase space at every report location is divided into 5×5 coding domains, and the steady states are drawn in green. For example, $f_0(((1,4),(2,4)))$ can be approximated by the transition of blue test points in domain $(1,4)$ and black ones in domain $(2,4)$ to the set of red points. (c) shows the initial conditions for an exemplary simulation with the according CPA, and the results after 1 and 5 hours are shown in (d) and (e), respectively. See also Fig. 5.2b.

state. A possible explanation is that the adsorbed arsenate at the pipe wall balances fast changes in the concentration of the dissolved arsenate.

Furthermore, we plot the steady state results from CPA, for which the approximation parameters are altered. In Fig. 5.2c the result is plotted for $\tilde{V} = V = \{0\}$ with an equidistant phase space partition of 5 domains in the D- and 15 domains in the A-direction, whereas in Fig. 5.2d the patterns are extended by $W = \{-1, 0\}$ to $V = \{-1, 0\}$. It is observed that in this example increasing the pattern length does not improve the result if compared to the Monte Carlo case, but increasing the state space resolution has a notable effect.

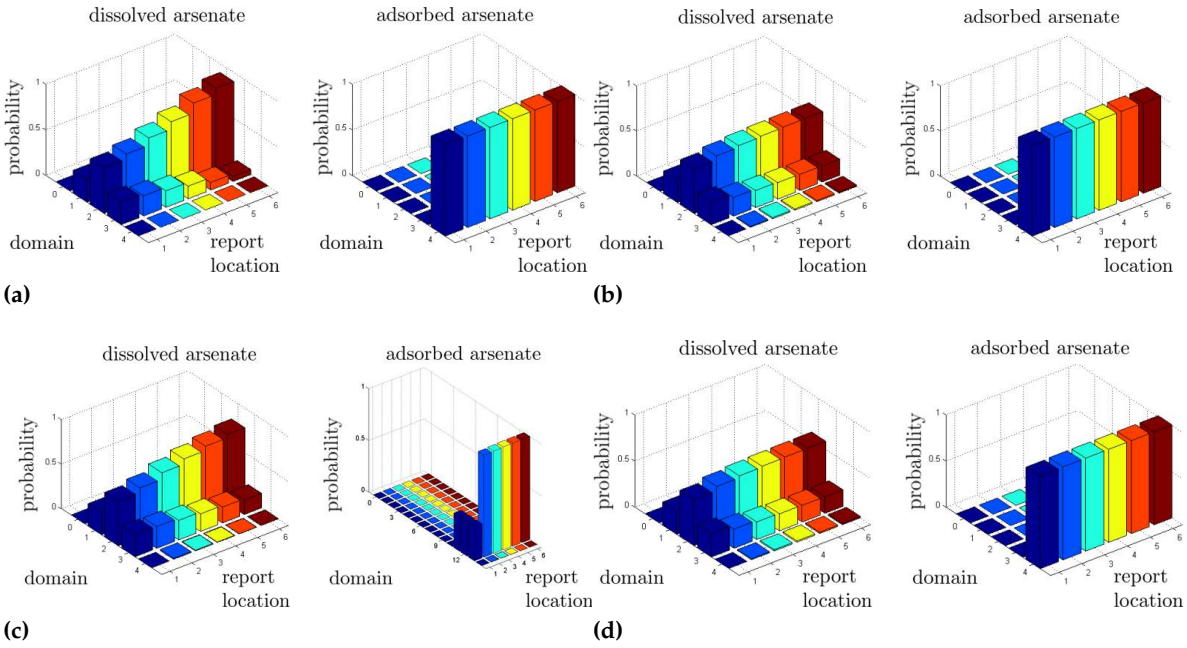


Figure 5.2: Steady states after 24 hours. (a) shows the result of a Monte Carlo computation. In (b) one finds the result for a CPA with a state space resolution of 5×5 and patterns with $\tilde{V} = V = \{0\}$, in (c) the results for 5×15 and $\tilde{V} = V = \{0\}$, and in (d) those for 5×5 and $\tilde{V} = \{0\}, V = \{-1, 0\}$.

We note that there is often no interest in global results and accordingly no need to transform between local and fully global states with $\hat{\alpha}$. Local information like that indicated in the graphs can be directly extracted from the CPA result. Similarly, in practice the information about initial values is often given locally, such that there is no need to use the full $\hat{\beta}$. Besides, the discrete state space information is often completely sufficient in practice. In the example a consumer is interested rather in risk level or threshold information about water contamination than in information in the form of exact concentrations. In some biological systems even the Boolean case, $|E| = 2$, is enough [31].

5.2.2 Bacterial Regrowth with Chlorine Inhibition

In this section we apply the CPA approach to a more complex advection-reaction problem, a model of bacterial regrowth with chlorine inhibition in a drinking water pipe.[1] First, we briefly review the model and then apply a CPA.

5.2.2.1 Bacterial Regrowth Model

The model is taken from [188] and has been considered as a standard example in the water safety community [149,150]. In the literature also an analysis of parameter uncertainty has been conducted [40].

[1] I thank Alexander Reiss for performing computations in his Bachelor's thesis. Details of this section can be found in [143].

It consists of the interaction of 4 species, that is chlorine Cl_2, the biodegradable fraction of dissolved organic carbon (BDOC) or substrate S, free bulk bacteria X_b and bacteria attached to the pipe wall X_a. The variables describe according concentrations. The dissolved species are advected with velocity v, and dispersion could be added also in the CPA method but is left out for simplicity here.

$$\begin{aligned}
\partial_t Cl_2 + v\partial_x Cl_2 =& -k_b Cl_2 - \frac{k_w}{r_h} \\
\partial_t S + v\partial_x S =& -\frac{1}{Y_g \beta}(\frac{\mu_a}{r_h} X_a + \mu_b X_b) \\
\partial_t X_b + v\partial_x X_b =& (\mu_b - k_d)X_b - k_{\text{dep}} X_b + \frac{k_{\text{det}} v}{r_h} X_a \\
\partial_t X_a =& (\mu_a - k_d)X_a - k_{\text{det}} v X_a + k_{\text{dep}} r_h X_b
\end{aligned}$$

The parameter r_h is the hydraulic radius and needed for conversion of bulk to wall concentrations like in the arsenate model. All other parameters can be assigned to one of the three processes that take place in this model.

i) Chlorine decays in the bulk water at rate k_b and is consumed at the pipe wall at rate k_w.

ii) The net growth of bacteria is given by a specific growth rate μ_i and a mortality rate k_d, where $i = a, b$. The growth rate depends on the substrate and the chlorine concentration

$$\mu_i = \begin{cases} \mu_{\max,i} \frac{S}{S+K_S} \exp\left(-\frac{Cl_2 - Cl_{2,t,i}}{Cl_{2,c}}\right) & \text{if } Cl_2 > Cl_{2,t,i} \\ \mu_{\max,i} \frac{S}{S+K_S} & \text{if } Cl_2 \leq Cl_{2,t,i} \end{cases} .$$

Here K_S is a saturation constant, $Cl_{2,t,i}$ a threshold for chlorine impact on bacteria and $Cl_{2,c}$ a characteristic chlorine concentration. To be more realistic, μ_i can be assumed to depend on the temperature. We neglect this dependence. The growth yield coefficient of bacteria Y_g and the conversion factor β are needed to describe the influence of bacterial growth on the substrate.

iii) k_{dep} and k_{det} are the rate constants for deposition of bulk bacteria at the pipe wall and detachment of attached bacteria, respectively. The detachment process depends on the bulk velocity v.

By studying only the reaction dynamics it is possible to get a better understanding of the phase space dynamics. For this purpose a chemostat with constant in- and outflow of the bulk species can be considered. We find that there are three stationary points, one of which is in the unphysical part of phase space. The second point is trivial, it is the state without any bacteria, and the third point is non-trivial. Depending on the parameters the points undergo bifurcations and can hence be stable or unstable.

We choose the parameters

$$\begin{aligned} k_{\text{det}} &= 0.18\text{h}^{-1}, & k_{\text{dep}} &= 0.25\frac{\text{s}}{\text{mh}}, \\ \mu_{\max,a} &= 0.2\text{h}^{-1}, & \mu_{\max,b} &= 0.2\text{h}^{-1}, \\ Cl_{2,t,a} &= 0.1\frac{\text{mg}}{l}, & Cl_{2,t,b} &= 0.03\frac{\text{mg}}{l}, \\ Cl_{2,c} &= 0.2\frac{\text{mg}}{l}, & K_S &= 0.4\frac{\text{mg}}{l}, \\ Y_g &= 0.15, & \beta &= 10^9\frac{\text{cells}}{\text{mg}}, \\ k_d &= 0.06\text{h}^{-1}, & v &= 0.05\frac{\text{m}}{\text{s}}. \end{aligned}$$

The reaction parameters are adopted for the most part from [188], only the literature value of k_{det} is actually 0.03h^{-1}.

5.2.2.2 Application of Cellular Probabilistic Automata

A critical parameter for CPA efficiency is the dimension of the phase space. In this application the dimension is with 4 species rather high. So we suggest to drop in this case study the chlorine equation and only consider the remaining 3-dimensional phase space at a constant chlorine concentration of $0.6\frac{\text{mg}}{l}$. With this assumption we can also drop the case-by-case analysis for μ_i and only calculate with the upper case.

For time and space discretization we proceed like in Sec. 5.2.1 with the Trotter formula and the backward difference ($U = \{-1, 0\}$) in the method of characteristics. The CPA time and space steps are $\Delta t = 15\text{min}$ and $\Delta x = v\Delta t = 45\text{m}$, and the test point mapping in the preprocessing is conducted with the less coarse discretization $\Delta t' = \frac{\Delta t}{25}$ and $\Delta x' = \frac{\Delta x}{25}$.

We consider the approximately invariant phase space Ω given by $S \in [0, 0.5]\frac{\text{mg}}{l}$, $X_b \in [0, 10^5]\frac{\text{cells}}{\text{ml}}$ and $X_b \in [0, 10^5]\frac{\text{cells}}{\text{cm}^2}$ and partition it uniformly with 5 symbols in S- and 7 symbols in each bacteria direction. The domains are labeled by (i, j, k), where $i = 0, ..., 4$ is the index in S-, $j = 0, ..., 6$ the one in X_b- and $k = 0, ..., 4$ the one in X_a-direction. The pattern length is chosen to be minimal, $V = \tilde{V} = \{0\}$, and in the preprocessing we use 100 test points in each domain.

As an example we conduct a simulation with a tubular pipe of length $l = 1500\text{m}$ and radius $r = 0.15\text{m}$, i.e., $I = \{1, ..., 33\}$ and hydraulic radius $r_h = 50\frac{l}{\text{m}^2}$. As initial conditions we choose symobl $(0, 0, 0)$ at all sites with probability 1, i.e., a completely empty pipe. At site 1 we assume a white noise boundary condition, in particular that with probability $\frac{1}{2}$ the bulk and the attached bacteria are in domains 2 and 3 each, and that the substrate is with probability $\frac{1}{2}$ in domain 2 and with probability $\frac{1}{4}$ in domain 1 and 3 each. The results after 24h are shown in Fig. 5.3.

The concentration of free bacteria gets lower with increasing distance from the boundary. Also, the concentration of attached bacteria remains in the lowest symbol with probability 1. We can trace back this behavior to the preprocessing, where e.g.

$$f_0((3,3,3),(0,0,0))((3,2,0)) = 0.2256, \quad f_0((3,3,3),(0,0,0))((3,3,0)) = 0.7744.$$

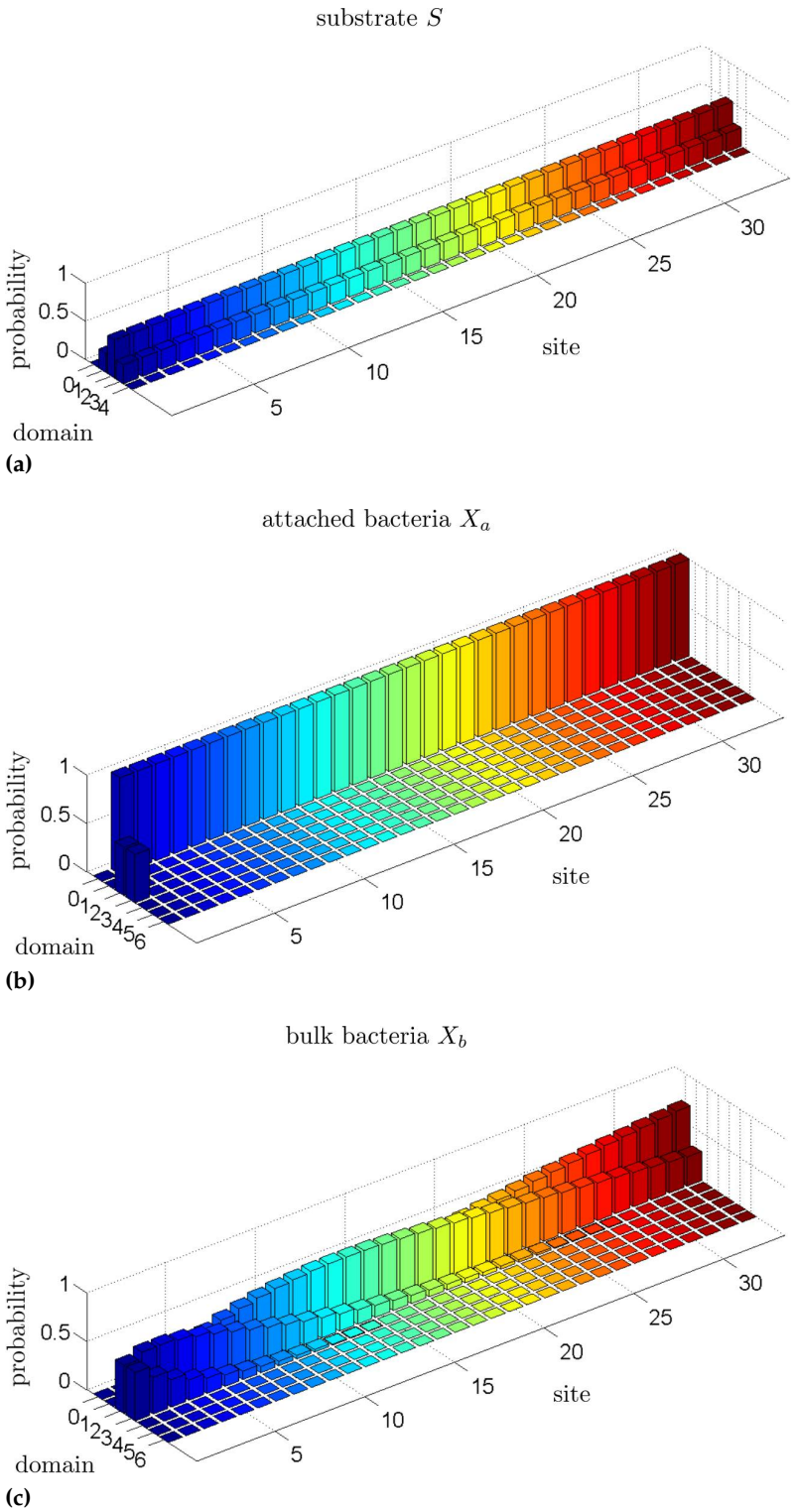


Figure 5.3: Results for a CPA simulation of the bacterial regrowth model after 24 hours. (a) contains the substrate concentration, (b) the concentration of attached bacteria and (c) the one of dissolved bacteria. Chlorine is set to a cconstant during the simulation.

Here no transition to a higher domain in the attached direction takes place, and the concentration of the free bacteria is diminished. This can be explained by the chlorine concentration which is held constant and relatively high throughout the whole pipe. As a result the bacteria death is higher than the growth rate. Also, the detachment rate for attached bacteria is relatively high in this simulation, so that it is even harder for the population of attached bacteria to grow. The substrate, in contrast, has almost the same distribution at all sites. The reason is that the concentration of bacteria is with maximally $10^5 \frac{\text{cells}}{\text{cm}^2}$ in the attached and $10^5 \frac{\text{cells}}{\text{ml}}$ in the free case relatively low. Not much substrate is consumed throughout the pipe, even more as the free bacteria are diminished over the pipe.

We conclude that in principle it is possible to apply the CPA approach also to a more complex advection-reaction problem. In our simple exemplary simulation we obtain results that meet our intuition. However, we made some model simplifications in this case study, most importantly we dismissed chlorine dynamics due to complexity issues. Further work is required to adapt the CPA method better to the bacteria regrowth model. For example the partition has to be adapted better to the different orders of magnitude in the various concentrations to really resolve the dynamics. Also additional simulations for more complex initial and boundary values or a comparison to other uncertainty propagation methods are desirable. Finally, it would be very interesting to include the chlorine dynamics. To overcome complexity issues a parallelization of the preprocessing and a dimensionality reduction of the relevant model dynamics to a lower-dimensional subspace [145] may be conducted.

5.3 Conservation Laws in Cellular Probabilistic Automata

Now we consider CPA with respect to conservation laws. The study is motivated by the arsenate application, for which we expect conservation of the total arsenate concentration. In the first section we show an example in which the CPA method violates the conservation law. We suggest an extension to the regular CPA method to impose conservation in the second section. Finally, in the third section the extension is applied to solve the problem in the arsenate example.

5.3.1 Violation of Conservation in Cellular Probabilistic Automata

Reconsider the transport and advection of arsenate of Sec. 5.2.1, that means the system of Eq. 5.1 with the parameters of Eq. 5.2. We apply the Trotter formula and the method of characteristics to the differential equations, Eq. 5.3, and choose the same space and time steps. The state space is discretized in 5 domains in D-, and in 15 domains in A-direction, and we only consider the minimal pattern length, $V = \tilde{V} = \{0\}$. The only difference to Sec. 5.2.1 is that we simulate a longer pipe with 60 sites now, and that we impose the stochastic boundary condition

$$g_l(e) = \begin{cases} 0.5 & \text{if } e = (0, 9) \\ 0.5 & \text{if } e = (1, 13) \\ 0 & \text{else} \end{cases},$$

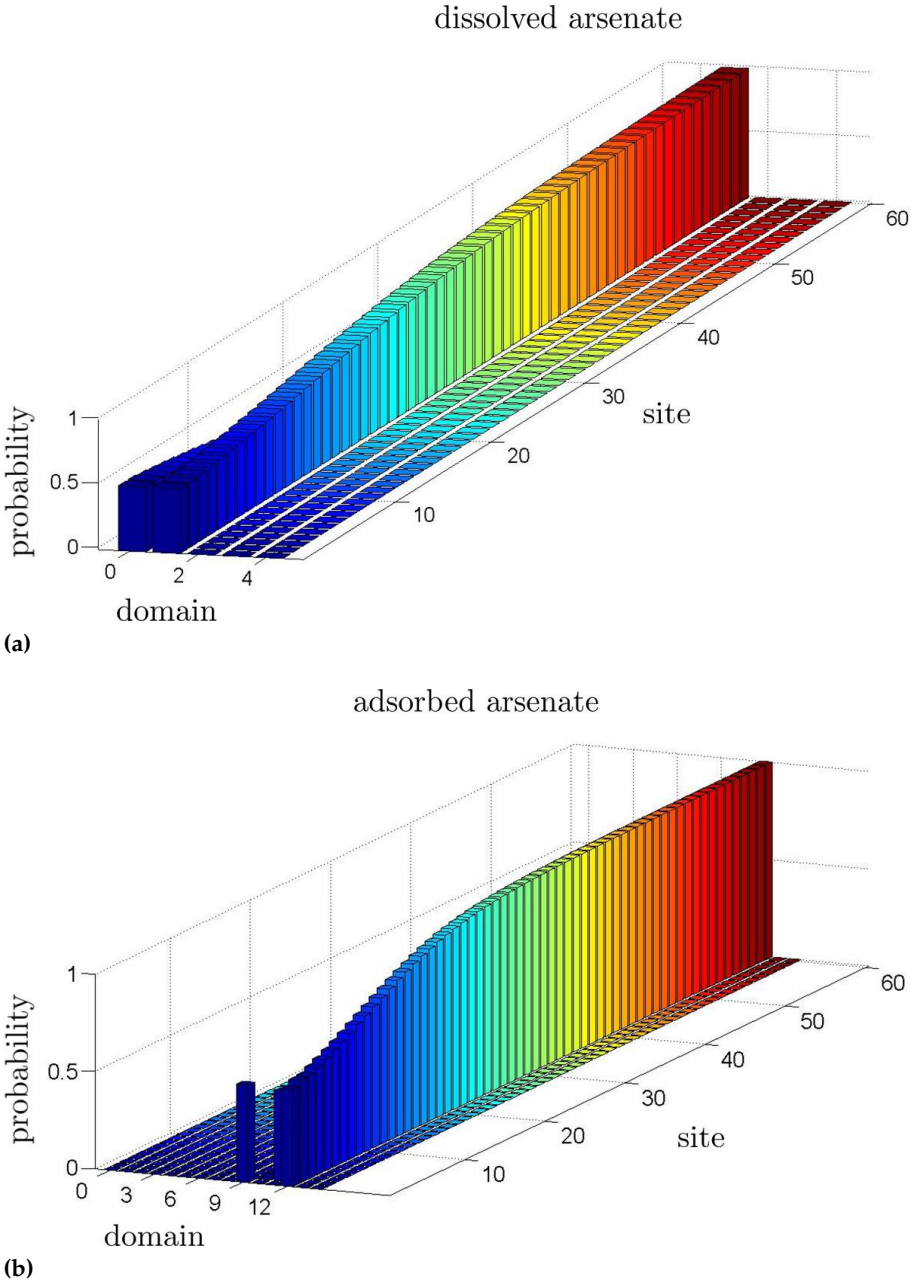


Figure 5.4: Simulation of arsenate transportation and adsorption in a long pipe with a regular local function. a) shows the dissolved arsenate and b) the adsorbed arsenate in the steady state. The total arsenate concentration is not conserved due to numerical errors; see the text.

where the first symbol is the D-, and the second the A-domain. The steady state result after 4 days has been calculated with probability threshold 0.0001 and is shown in Fig. 5.4.

In both the adsorbed and the dissolved arsenate direction the probabilitiy accumulates in two neighboring domains throughout the whole pipe. We observe that for the respective domain with lower concentration the probability at the entrance of the pipe is higher than

the corresponding probability at the end of the pipe. In contrast, for the respective domain with higher concentration the probability at the inflow is smaller than that at the outflow. This means that in the steady state in average the total arsenate inflow is higher than the outflow. But the PDE model respects the conservation of total arsenate concentration. We find the hyperbolic conservation law $\partial_t(D + \frac{1}{r_h}A) + v\partial_x D = 0$, where the conserved quantity is the total concentration $D + \frac{1}{r_h}A$, and where the flow only consists of dissolved arsenate. So the conservation law is violated in the pipe. We will investigate that more formally in the next section.

We recap that uncertainty propagation with CPA consists of two steps, the preprocessing and the simulation. It is in the preprocessing, where the transition probabilities are determined to translate the PDE into a CPA, whereas the simulation step is completely independent of the PDE. We conclude that the violation of conservation is a matter of the preprocessing, and the errors from the preprocessing then accumulate in the simulation. Our approach is therefore to adapt the preprocessing to conservation laws.

5.3.2 Imposing Conservation in the Preprocessing

In this section we suggest a method to adapt the CPA transition probabilities to conservation laws, i.e., a postprocessing of the preprocessing before the simulation. We only consider the minimal pattern length, $V = \tilde{V} = \{0\}$. There are various possible definitions of conservation in a probabilistic setting. We show that the expectation value of a conserved quantity is preserved during time evolution of a probability density, if it is preserved by the trajectories of all realizations. Therefore we require our extension of the CPA preprocessing to assure that the expected values of the conserved quantities are preserved. We expect that the method can be extended to higher moments.

Formally we consider the time- and space-discrete dynamical system with the locality property of Def. 4.1. We assume that the dynamical system is derived from a PDE by time and space discretization, and that a conservation law holds locally. In the following lemma V^n_{i+U} denotes the random variable that is the natural restriction of the random variable $V^n : X \to \Omega^m$ to the $(i+U)$-coordinates of the image, where $i \in \tilde{I}$. $E[V]$ is the expected value of a real-valued random variable V; see Sec. 2.4.

Lemma 5.1. Let $n \in \mathbb{N}_0$, $i \in \tilde{I}$, measurable $k : \Omega^U \to \mathbb{R}$ and $k' : \Omega \to \mathbb{R}$ such that

$$k'(v_i^{n+1}) = k(v_{i+U}^n)$$

is fulfilled for every $v^n, v^{n+1} = \Phi^\tau(v^n) \in \Omega^m$. Consider random variables $V^n, V^{n+1} = \Phi^\tau(V^n) : X \to \Omega^m$ with densities $g^n, g^{n+1} = P_{\Phi^\tau}(g^n) \in D(\mathbb{R}^{mn})$, respectively. Assume that Φ^τ is injective and continuously differentiable. Then

$$E\left[k'(V_i^{n+1})\right] = E\left[k(V_{i+U}^n)\right].$$

Proof It is well known [106] that under the given assumptions $g^{n+1}(v) = P_{\Phi^\tau}(g^n)(v) = g^n(\Phi^{-\tau}(v))|\det \mathrm{D}(\Phi^{-\tau})|$, where $\det \mathrm{D}(\Phi^{-\tau})$ is the determinant of the Jacobian matrix of

$\Phi^{-\tau}$. Therefore

$$\begin{aligned} E\left[k'(V_i^{n+1})\right] &= \int k'(v_i) g^{n+1}(v) dv \\ &= \int k'(v_i) (g^n(\Phi^{-\tau}(v))| \det \mathrm{D}(\Phi^{-\tau})| dv \\ &= \int k'(\Phi^{\tau}(w)_i) g^n(w) dw \\ &= \int k(w_{i+U}) g^n(w) dw \\ &= E\left[k(V_{i+U}^n)\right] \end{aligned}$$

by substitution $w = \Phi^{-\tau}(v)$ and locality of Φ^{τ}. □

Note that g^n and g^{n+1} may be replaced by suitable marginal distributions when calculating the expected values of the local variables.

Since a conservation law holds locally, the locality approximation in the CPA construction cannot be the cause of its violation. There remain just two steps in the translation from PDEs to CPA, where things can go wrong: the numerical approximation of the Frobenius-Perron operator (FPO) by test point mapping, and the subsequent restriction of image densities to the subspace of piecewise constant densities; see Def. 4.4.

We suggest an algorithm that corrects the errors originating from both steps with respect to the expected values of the conserved quantities. The idea is to constrain the set of local functions to those which give rise to the exact expected values in the image distribution. Note that we know the exact expected value according to the above lemma. The transition probabilities are then chosen by an optimization such that they are as close as possible to the approximate image point distribution from the regular preprocessing.

We state the algorithm for notational simplicity only for one conserved quantity. It is straightforward to extend it to several of them by adding more constraints. For a given partition E of Ω and for all $\varphi \in E^U$ we conduct the following steps:

i) determine vector $m = (m_e)_{e \in E}$ by $m_e = E[k'(V)]$, where $V : X \to \Omega$ has uniform distribution over Ω_e,

ii) determine vector of non-zero preliminary transition probabilities p' from φ to $e \in E$ by mapping of test points with standard preprocessing of Sec. 5.1.1,

iii) determine for all image domains Ω_e with non-zero probability the mean $\tilde{M}_e$ of the conserved quantity k' with respect to the image points that it contains,

iv) determine exact expected value $M = E\left[k(V)\right]$ of the conserved quantity, where $V : X \to \Omega^U$ has uniform distribution over Ω_φ,

v) determine a diagonal weight matrix W with diagonal elements $(\frac{1}{d_e})$, where $d_e = |m_e - \tilde{M}_e|$,

vi) solve quadratic program for new non-zero transition probabilities p

$$\begin{aligned} (p - p')^T W (p - p') &\to \min \\ M = m^T p, \quad \|p\|_1 = 1, \quad p &\geq 0, \end{aligned}$$

where m just contains only the indices $e \in E$ of the vector that appear in p'.

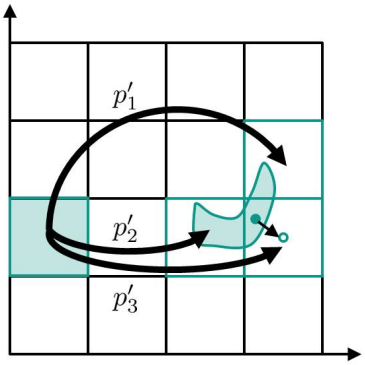


Figure 5.5: We show one site with an exemplary 2-dimensional phase space partitioned in 16 boxes to illustrate the postprocessing; see the text. The preimage state at this site is the left green box, and the standard preprocessing assigns in this example probabilities p'_1, p'_2 and p'_3 to three image domains (boxes with green boundaries). All domains that have an intersection with the images of the test points (green shape) are considered as image domains. The exact conserved expected value M is calculated from the preimage state. In one of the image boxes we exemplary sketch $\tilde{M}_e$ by a filled green circle and m_e by a green circle with white interior. The quadratic program redistributes the probabilities according to the differences in $\tilde{M}_e$ and m_e such that the expected value is conserved exactly.

Fig. 5.5 gives some intuition on what the algorithm does. We add five comments on the algorithm. First, we note that step i) is independent of φ, so it has to be computed only once. Second, the effect of the weight matrix W constructed in steps iv) and v) is to better maintain the approximate p' when restricting the image density to be piecewise constant. The better the expected value of the conserved quantity with respect to the image points in one domain already matches the domain's expected value, the more expensive it is to change the probability in the optimization step vi). A simple variant of the algorithm is of course to omit steps iv) and v) and choose W as the identity matrix in step vi). Then the quadratic program turns into a least-squares problem. Third, we note that by only redistributing the probabilities for the domains with non-zero probability we prevent the optimization from allocating non-zero probabilities to domains that are actually not in the support of the image density. However, for some $\varphi \in E^U$ there might be no feasible solution to the quadratic program, because then the constraints for normalization and expected value are too strict. This happens especially for boundary domains. A practical workaround is to also include the neighboring domains in the optimization procedure. If we then still do not find any feasible solution, we keep the probabilities from the regular preprocessing. Fourth, we mention that there are other options to define the optimization problem, e.g. by including the constraint for the expected value with a big constant in the minimization procedure. This also would help to find feasible solutions more frequently. And fifth, we note that the additional computational effort of the postprocessing is negligible in practice. We only have to execute one optimization for every preimage pattern, which is much less expensive than the test point mapping.

5.3.3 Application: Adsorption of Arsenate

In the arsenate application the conserved quantity is the total arsenate concentration. With the time and space discretization of Eq. 5.3 we find that

$$D_i^{n+1} + \frac{1}{r_h} A_i^{n+1} = D_{i-1}^n + \frac{1}{r_h} A_i^n.$$

So we can identify

$$k(v_{i+U}^n) = D_{i-1}^n + \frac{1}{r_h} A_i^n, \qquad k'(v_i^{n+1}) = D_i^{n+1} + \frac{1}{r_h} A_i^{n+1}.$$

When we conduct an extended preprocessing with the parameters of Sec. 5.3.1, we include also the next neighbors in the optimization. As a result we find a feasible solution for all preimage states, and the new transition probabilities differ slightly from the regular preprocessing. A remarkable difference is that now $f_0(0,0)(0) = 1$, whereas without the postprocessing we had $f_0(0,0)(0) = 0.9446$ and $f_0(0,0)(1) = 0.0554$. With the extended preprocessing it is not any longer possible by numerical mistakes that arsenate is created in an empty pipe.

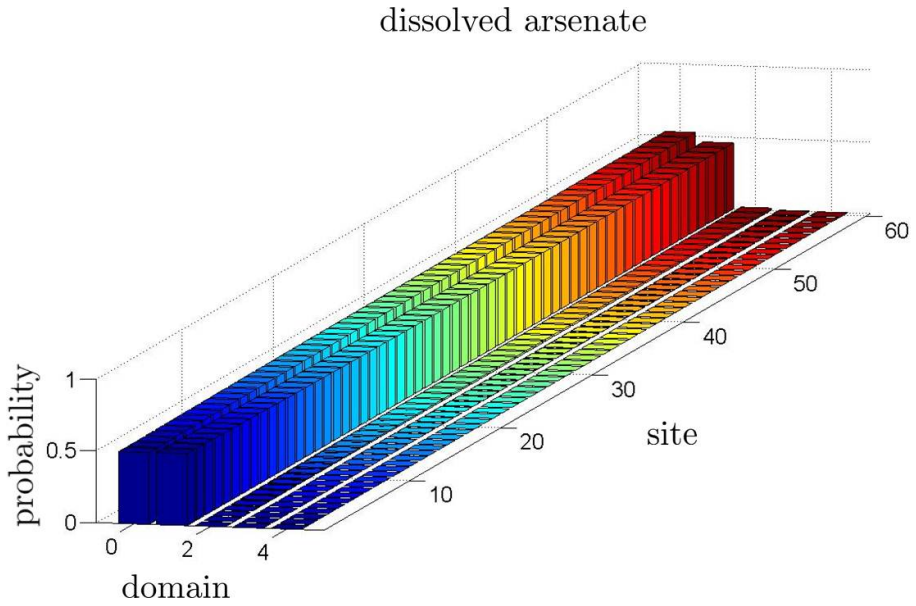
With the new local function we repeat the same simulation as in Sec. 5.3.1. The results for the steady state after 4 days are shown in Fig. 5.6. Even in a long pipe we can now guarantee conservation of the expected value of the total arsenate concentration. The unphysical numerical phenomenon described in Sec. 5.3.1 no longer appears. Furthermore, simulations in the setup of Sec. 5.2.1 with the new local function show very similar results to the old local function. Especially the accumulation of dissolved arsenate in risk level 3 can still be observed; see Fig. 5.8 for an example with similar parameters. We note that in practice it is critical to choose the threshold probability in the simulation for considering a probability to be non-zero quite low, so that small errors cannot accumulate. Here we used 0.0001.

5.4 Application II: Contamination in Water Grids

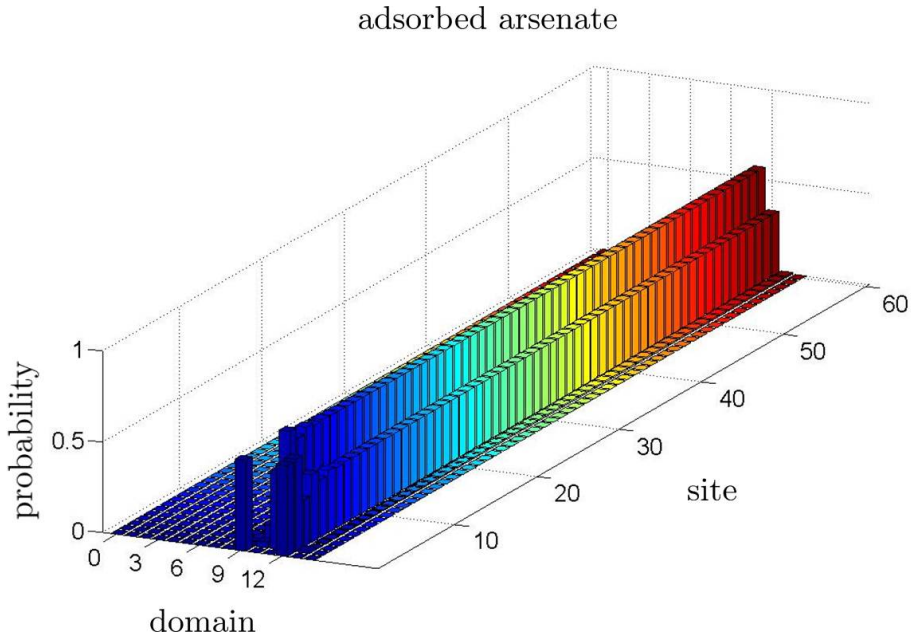
In this section we aim at simulating contaminant fate in large water grids. First, we introduce our general modeling approach and discuss how grid simulation can be addressed with CPA. Second, we derive a method to handle junctions as the characteristic elements of a grid in a CPA setup, and third we apply our ideas to arsenate adsorption in an exemplary municipal drinking water grid.

5.4.1 Modeling Approach

The simlation of contaminant fate in water grids is a multiphysics problem, where hydraulics and chemical kinetics have to be combined [83, 150]. We decouple both phenomena and focus only on the chemical kinetics on the basis of a simplified water model. The simplified model is the result of a sophisticated hydraulic simulation with SIWA, a commercial software package from Siemens [153]. We use SIWA to model, illustrate and simulate the hydraulics of a water grid defined by reservoirs, pumps, junctions, valves, pipes,



(a)



(b)

Figure 5.6: Simulation of arsenate transportation and adsorption in a long pipe with an improved local function. a) shows the dissolved arsenate and b) the adsorbed arsenate in the steady state. In contrast to Fig. 5.4 now the expected value of the total arsenate concentration is conserved; see the text.

consumers and a given topology. We assume tubular pipes throughout the whole network. Only the water velocity in each pipe is extracted, which together with length, radius and topology of the pipes constitutes the simplified model. The chemical kinetics can then be described by advection-reaction equations for pipe dynamics and algebraic couplings for

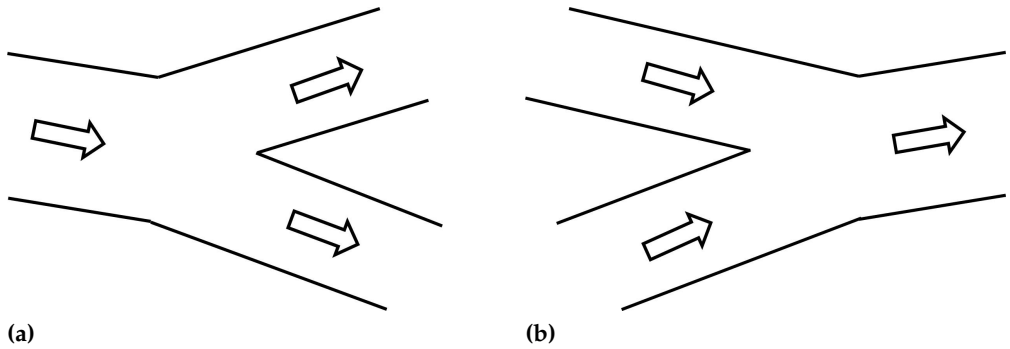


Figure 5.7: Two types of junctions have to be distinguished for CPA simulation of an advection-reaction problem. In type (a) the hydraulics is such that we have only one inflow pipe that splits up, and in type (b) we have several mixing inflows.

conservation at junctions. In practice this simplification might be enough, when we only consider risk levels anyway.

Concentrations of dissolved species are usually given in units of mass per water volume, while species at the pipe walls are described in units of mass per wall area. The interaction of both types can be described with the help of the quotient of the cross sectional pipe area and the wetted perimeter of the cross-section, the hydraulic radius r_h. For tubular pipes the hydraulic radius is just half of the usual radius r, $r_h = \frac{r}{2}$. The chemical kinetics thus differs with the radius of the pipe in which it takes place. When describing the chemical kinetics with CPA, in principle the preprocessing has to be conducted for every pipe radius in the grid. However, as pipes are usually standardized, only a few preprocessing runs are necessary in practice.

Besides the radius there is one more parameter which may vary from pipe to pipe in the simplified advection-reaction model: water velocity. But different water velocities do not require several CPA preprocessings. We use a global time step for all pipes and calculate an individual space step according to a pipe's velocity. By adjusting the number of sites per pipe to the length and space step of the pipe we can therefore approximately use the same CPA for all different velocities.

Our ideas immediately work for hydraulic steady states on which we focus in this work. However, it is straightforward to extend this idea to simple dynamic hydraulics, i.e., changing velocities. We only have to vary the number of sites per pipe to account for changing velocities. When the number of sites per pipe is high, this is certainly a good approximation. Also, the space-discrete contaminant variables have to be adapted proportionately. Furthermore, the mixing rule at junctions has to be adapted dynamically, as it depends on velocity; see the next section. However, for very short pipes or small velocities, perhaps even changing flow directions, we encounter problems with this approach. We note that for models with dispersion we may also run into difficulties with these ideas because the second spatial derivative leads to additional terms that scale differently from the advection terms.

5.4.2 State-Discrete Modeling of Junctions

To simulate whole water grids we need to extend the CPA algorithm to junctions. In this section we derive an according algorithm for advection-reaction problems involving 3 pipes. For notational simplicity we derive our idea just for a single dissolved species and minimal pattern length, but generalizations are straightforward. Applications are given as part of the grid simulation in the next section.

First, we note that when we use a time and space discretization like in Eq. 5.3, mixing is a local phenomenon. We only have to define a new local update rule for the first site of each output pipe. In principle we have to distinguish between two dynamically different types of junctions; see also Fig. 5.7. If the flow direction is such that there is just one input pipe, Fig. 5.7a, concentrations can be handled easily. We do not have to think of any mixing for concentrations, since both water volume and species' mass are split. So we can just apply the regular CPA rule on the neighborhoods consisting of the last site of the input pipe and the first site of each output pipe, respectively.

We focus on the second case, Fig. 5.7b, where there are two input pipes a and b and one output pipe c.

V_i... volume of water	M_i... mass of dissolved species
u_i... water velocity	D_i... concentration of dissolved species
A_i... cross sectional area	

denote the respective variables at the last site in pipe $i \in \{a, b\}$ and at the first site of pipe $i = c$. The conservation of the species' mass and the conservation of volume of the incompressible water read

$$\begin{aligned} M_c &= M_a + M_b, \\ V_c &= V_a + V_b \end{aligned}$$

and can be used to derive an equation for the concentration of the dissolved species in pipe c. We consider only the 1-dimensional phase space $\tilde{\Omega}$ of the dissolved species with partition $\{\tilde{\Omega}_e\}_{e\in\tilde{E}}$. Although $\tilde{\Omega}$ and $\tilde{E}$ are the same at every site we will attach indices a, b, c to them to clarify to which site we refer. Then $m : \tilde{\Omega}_a \times \tilde{\Omega}_b \to \tilde{\Omega}_c$ is given by

$$\begin{aligned} D_c = m(D_a, D_b) = \frac{M_c}{V_c} = \frac{M_a + M_b}{V_a + V_b} = \frac{V_a D_a + V_b D_b}{V_a + V_b} \\ = \underbrace{\frac{A_a u_a}{A_a u_a + A_b u_b}}_{\alpha} D_a + \underbrace{\frac{A_b u_b}{A_a u_a + A_b u_b}}_{\beta} D_b \end{aligned}$$

and has to be translated into discrete phase space. All we can get is again a probability distribution $m_0 : \tilde{E}_a \times \tilde{E}_b \to D(\tilde{E}_c)$ given by

$$m_0(e_a, e_b)(e_c) = \frac{\lambda(\Omega_{e_a} \times \Omega_{e_b} \cap m^{-1}(\Omega_{e_c}))}{\lambda(\Omega_{e_a} \times \Omega_{e_b})},$$

Lemma 5.2. If the partition $\{\tilde{\Omega}_e\}_{e\in E}$ is uniform with resolution $\Delta\tilde{\Omega}$, then $m(\tilde{\Omega}_{e_a}, \tilde{\Omega}_{e_b})$ is also an interval of length $\Delta\tilde{\Omega}$ for all $e_a, e_b \in E$.

Proof Let $\tilde{\Omega}_{e_i} = [v_i, w_i]$ for $i \in \{a, b\}$, and denote $m([v_a, w_a], [v_b, w_b]) = [v_c, w_c]$. Note that by definition $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$. Therefore $v_c = \alpha v_a + \beta v_b$ and $w_c = \alpha w_a + \beta w_b$, and so

$$w_c - v_c = \alpha(w_a - v_a) + \beta(w_b - v_b) = (\alpha + \beta)\Delta\tilde{\Omega} = \Delta\tilde{\Omega}.$$

□

In case of a uniform partition with resolution $\Delta\tilde{\Omega}$ therefore there are $e_c, e_c' \in E$ with

$$m_0(e_a, e_b)(e) = \begin{cases} p & \text{if } e = e_c \\ 1-p & \text{if } e = e_c' \\ 0 & \text{else} \end{cases}$$

for $p = \frac{w_c - \Delta\tilde{\Omega}\lfloor w_c/\Delta\tilde{\Omega}\rfloor}{\Delta\tilde{\Omega}}$, where $\lfloor * \rfloor$ is the floor function and $w_c = m(w_a, w_b)$ for $\tilde{\Omega}_{e_i} = [v_i, w_i]$, $i \in \{a, b\}$. e_c is given by the coding domain $\tilde{\Omega}_{e_c}$ that contains w_c, and e_c' encodes the neighboring coding domain.

Similarly to Sec. 5.3 m_0 might also be extended to include conservation considerations. Furthermore m_0 has to be embedded in the n-dimensional phase space Ω of dissolved and adsorbed species to get the whole mixing rule.

5.4.3 An Exemplary Municipal Grid

We apply our ideas to a part of a municipal drinking water grid given in Fig. 5.8. It consists of many elements, of which we extract for CPA simulation nine long pipes, three consumers and two sources, the reservoirs. The rest of the elements is just needed for the hydraulic simulation: Reservoir 2 is attached to pipes 1 and 2 by very short connection pipes a,b and c. However, they are so short that their influence on chemical kinetics can be neglected. Pumps and elevation parameters for all objects establish a pressure profile that leads to water flow to the consumers. Some valves can be used to control the hydraulics.

For simplicity all setup parameters like demand and pump profiles are assumed to be constant in time such that the system is in a hydraulic steady state. A SIWA simulation in a realistic exemplary setting leads to the following parameters for a simplified advection-reaction model:

pipe	1	2	3	4	5	6	7	8	9
radius [m]	0.1500	0.300	0.300	0.1625	0.300	0.1500	0.300	0.300	0.300
length [m]	1500	1345	2305	500	1730	500	2690	4000	100
velocity $[\frac{\text{m}}{\text{s}}]$	2.150	0.690	0.690	0.300	0.605	0.120	0.575	0.575	0.575
# sites	5	13	22	12	19	29	31	46	2
CPA	1	2	2	3	2	1	2	2	2

pipe	a	b	c
radius [m]	0.1500	0.300	0.150
length [m]	10	10	5
velocity $[\frac{\text{m}}{\text{s}}]$	2.150	0.690	0.620

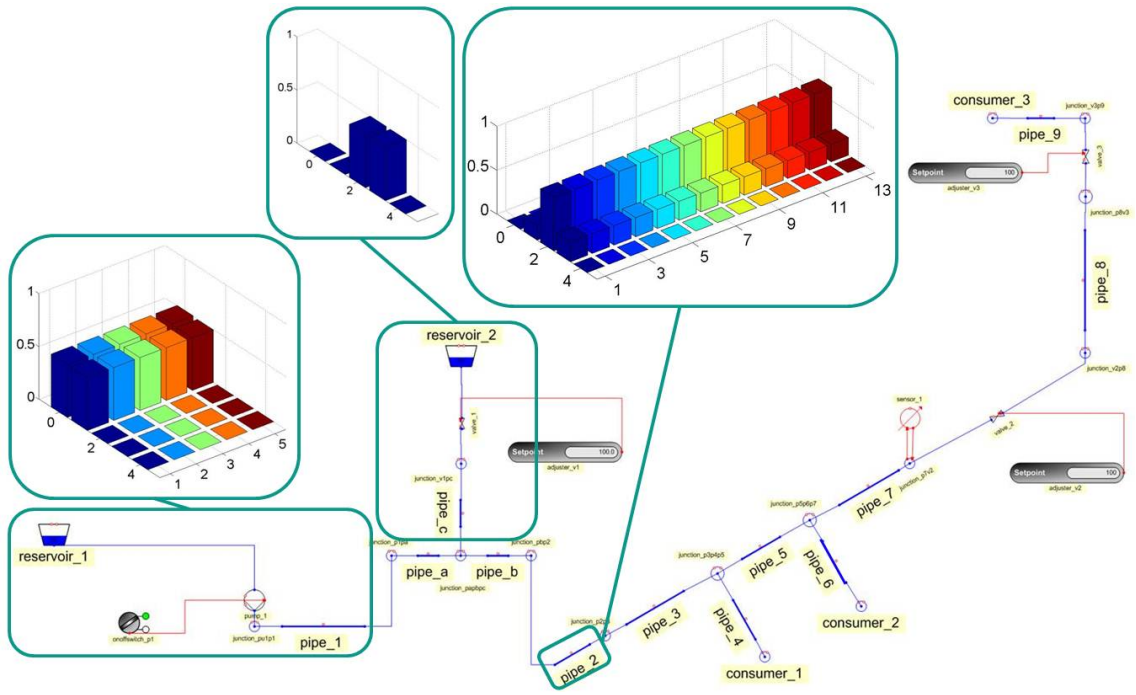
We describe advection and adsorption of arsenate in the grid like in Sec. 5.2.1 with the model of Eq. 5.1. The chemical parameters are given in Eq. 5.2, but the hydraulic and the pipe parameters are of course different. We consider the system on the approximately positively invariant state space domain given by $D \in [0,1]\frac{\text{mg}}{l}$ and $A \in [0,100]\frac{\text{mg}}{\text{m}^2}$. The translation to CPA with the postprocessing of Sec. 5.3 is conducted 3 times according to the 3 different pipe radii; see the above table and the discussion in Sec. 5.4.1. To discretize time and space we decouple again the advection and the reaction step with the Trotter formula. For the advection the method of characteristics is used with $U = \{-1, 0\}$ as the backward difference and the global time step $\Delta t = 2.5$min. We map 75×37 randomly distributed test points by using intermediate steps with the less coarse discretization $\Delta t' = 0.1$min. The pattern length is minimal, $V = \tilde{V} = \{0\}$, and the phase space partition equidistant with 5 domains in the D- and 15 in the A-direction.

In the CPA simulation we choose the threshold probability 0.005. To describe the mixing in the formalism of Sec. 5.4.2 we model reservoir 2 as a pipe with one site, the boundary condition, with the characteristics of pipe c. The number of sites in each pipe is calculated by dividing the length by the global time step and the velocity. The inflow boundary conditions of pipes 3, 8 and 9 are just the last sites of the previous pipes 2, 7 and 8, respectively. At the inflow of pipe 4, 5, 6 and 7 the trivial mixing rule is applied, and at the inflow of pipe 2 the non-trivial. The reservoirs are modeled as external stochastic boundary conditions for pipes 1 and c, respectively.

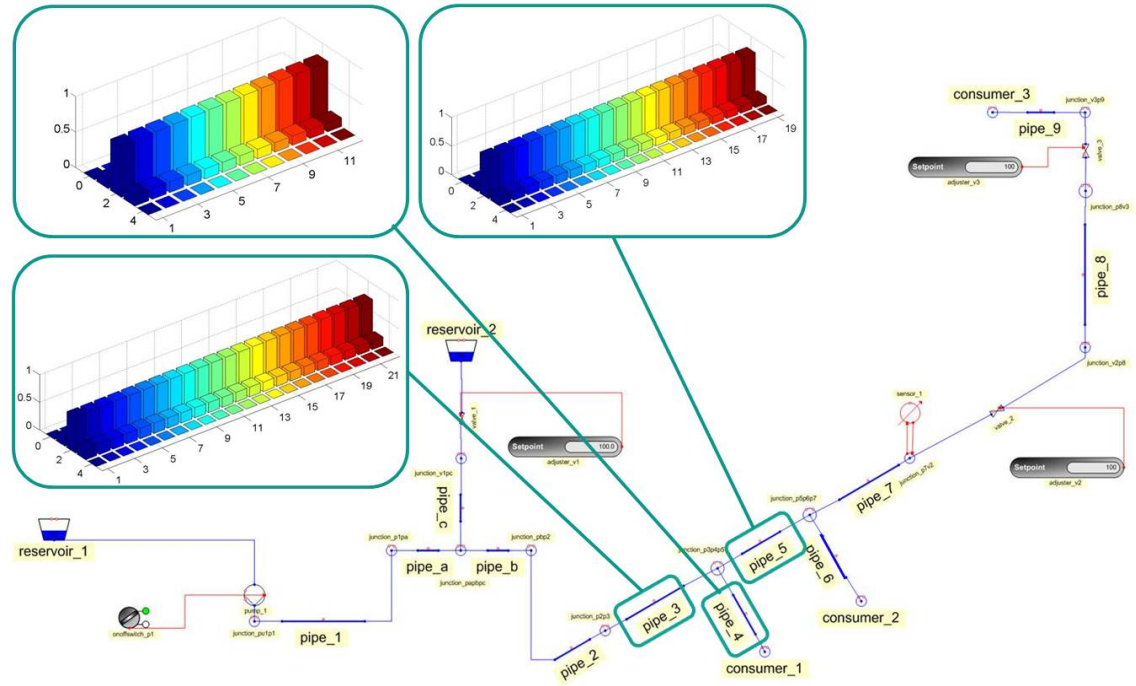
We want to simulate the spread of arsenate in reservoir 2 through the network over time, while the rest of the grid and in particular reservoir 1 is not contaminated in the beginning. As initial conditions we thus choose the lowest state with probability 1 for dissolved and adsorbed arsenate throughout the whole network except for reservoir 2. In reservoir 1 the boundary condition for the dissolved arsenate is the lowest state with probability 1, and in reservoir 2 a distribution over higher states; see Fig. 5.8a. Because of the Trotter discretization we do not have to specify a boundary condition for the adsorbed arsenate.

The time evolution is simulated over 4 days, and it can be displayed at every site. We only show as an exemplary result in Fig. 5.8 the final steady state for dissolved arsenate in the pipes involved in mixing. In the steady state the adsorbed arsenate is everywhere in the grid approximately in state 13 with probability 1. An exception is of course pipe 1 which behaves like in Fig. 5.6. The simulation time is with about 6 hours on a common laptop much shorter than the simulated time. The preprocessing time was approximately 3 times 1,25 hours.

We are able to simulate under uncertainties a large grid faster than real-time by calculating directly on a simplified state space. The state space allows for a direct practical interpretation in terms of contamination risk levels. Our example comprises the features introduced in the former sections: the unexpected accumulation of probability in state 3 for the dissolved arsenate due to the non-linearity of the interaction, the mass conservation and the mixing behavior. Already in this simple setup we gather interesting information for the consumer which would not be available in a deterministic simulation. In chemically and topologically more involved examples we expect the consumer to really benefit from results that cannot be simulated easily by other means.



(a)



(b)

Figure 5.8: The topology of a part of an exemplary municipal grid. The grid mainly consists of 9 long and three short pipes, 2 reservoirs and 3 consumers. For the hydraulic simulation with SIWA also the pumps and valves are necessary, which can in principle be controlled online with the red boxes. The steady state after 4 days is shown for the dissolved arsenate and both junction types, in a) for reservoir 2 and pipes 1 and 3, and in b) for pipes 3, 4 and 5. In each diagram the vertical axis depicts the probability of being in the domain 0 to 4 at the respective site.

5.5 Application III: Euler Equation of Fluid Dynamics

The applications so far have involved advection-reaction equations, where the difficulty lies in mastering the reaction dynamics. We are also interested in how the CPA method performs for more complex hyperbolic PDEs, where the spatial dynamics is the challenge. Here we pick the isothermal Euler equation of fluid dynamics for a case study.

In the first section we state the equations and motivate our work with water and gas grid modeling. Secondly, we describe how an according CPA can be constructed and highlight sensible points in this process. In the end we give several numerical examples and comment on the opportunities and limits of the approach.

5.5.1 The Isothermal Euler Equations

The isothermal Euler equations are basic hyperbolic equations of fluid dynamics [109,110]. They describe the conservation of mass and momentum of a fluid:

$$\begin{aligned} \partial_t \rho + \partial_x q &= 0, \\ \partial_t q + \partial_x (\frac{q^2}{\rho} + a^2 \rho) &= 0, \end{aligned}$$

where t and x are time and one-dimensional space coordinate, respectively, a is the speed of sound, $\rho(x,t)$ the fluid density and $q(x,t) = \rho(x,t)v(x,t)$ the flow given by the product of density and velocity $v(x,t)$. Acutally $p = a^2\rho$ is the pressure in the ideal gas law. According Riemann problems can be solved even analytically and lead to shocks or rarefaction waves. The characteristic speeds are the eigenvalues $\lambda_{1,2} = \frac{q}{\rho} \pm a$. In the following we will use the short-hand notation

$$\partial_t v + \partial_x f(v) = 0$$

for the isothermal Euler eqution, where $v = (\rho, q)^T$ and $f(v) = (q, \frac{q^2}{\rho} + a^2\rho)^T$.

We motivate the study of the isothermal Euler equations with two common applications. One is hydraulic simulation of drinking or waste water grids, and the other is gas grid simulation. In general municipal drinking water grids there are both pressurized and free surface water flow. In one pipe pressurized flow is described by the general Euler equation [110], and free surface flow by the Saint-Venant or shallow water equations [160]. In realistic applications a unified model for both types can be used [19,158]:

$$\begin{aligned} \partial_t(\rho A) + \partial_x(\rho Q) &= 0, \\ \partial_t(\rho Q) + \partial_x \frac{(\rho Q)^2}{\rho A} + g(\rho A)\partial_x z + A\partial_x p &= -g(\rho A)S_f, \end{aligned}$$

where $\rho(x,t)$ is again the density, $A(x,t)$ the wetted cross-section, $Q(x,t)$ the volume flow along the pipe, g the gravitational acceleration and $z(x,t) = z(x, A(x,t))$ the water

surface elevation above a reference level. For the friction S_f either the Manning-Strickler or Darcy-Weisbach ansatz can be used, and the pressure p is given by the Boussinesq closure

$$p = \rho g h + \frac{1}{\beta}\frac{\rho - \rho_0}{\rho_0},$$

where the first term is the hydrostatic pressure at water depth $h = h(z(x,t))$, β the isothermal compressibility and ρ_0 the water density under free flowing conditions. Like in the isothermal Euler equation this closure is linear in ρ. Actually the isothermal Euler equations are the special case of a pressure-driven flow with constant cross-section, no elevation and no friction, where $q := \rho\frac{Q}{A}$ and $a := (\beta\rho_0)^{-\frac{1}{2}}$. Application of CPA to them can hence be viewed as a case study for more complex water problems. Here we face severe uncertainties in the boundary conditions as lack of knowledge about consumer behavior, and also in weather forecasts in case of waste water simulation. Especially in the waste water application there is also a natural interpretation of the discretized phase space in terms of threshold values for amounts of water, e.g. for the filling of overflow chambers.

Also in gas grid simulation there has been a lot of interest in dynamic simulations recently. Dynamic extend steady-state models and are able to grasp more aspects of reality [4]. Many models have been introduced; see e.g. [23, 141] and other publications of the Pipeline Simulation Interest Group (PSIG) [70], and different numerical solution techniques have been discussed; see e.g. [6]. The isothermal Euler equations are a widely used model for gas pipelines, and many other existing models are extensions thereof.

5.5.2 A Cellular Probabilistic Automata Approach

From now on we work with the parameters for methane pipelines. In this section we describe in five steps how a CPA can be constructed for the isothermal Euler equations. First, we choose a finite volume method to discretize time and space. This decides on the neighborhood and pattern structure. Second, we fix and partition an invariant part of phase space. Third, we pick the time and space step size in accordance with the CFL-condition for the fixed phase space doamin. Fourth, we explain how to deal with stochastic boundary conditions. In this step we need to conceptually extend the CPA method, because the boundary conditions are more difficult than in the transport equations considered so far. In the last step we describe how the ideas of Sec. 5.3 can be applied to ensure conservation properties of the expected value.

5.5.2.1 Choosing a Finite Volume Method

A CPA first requires discretization of time and space with step sizes Δt and Δx, respectively. Since the isothermal Euler equation is hyperbolic, we suggest to use a finite volume method for this purpose. An introduction and further information regarding the according concepts can be found in [110]. Note that the interpretation of the variables differs a little from the finite difference schemes used so far: variables at each discretized site and time are now spatial averages over intervals around the sites and not any more the values at the sites itself. With this said we will yet use the same variable names as before.

In Sec. 5.1.2 we have seen that CPA scale exponentially bad in the stencil or neigborhood length $|U|$. Therefore we choose the Lax-Friedrichs method, which with $U = \{-1, 1\}$ only

contains two sites,

$$v_i^{n+1} = \frac{1}{2}(v_{i-1}^n + v_{i+1}^n) - \frac{\Delta t}{2\Delta x}[f(v_{i+1}^n) + f(v_{i-1}^n)]. \tag{5.4}$$

It updates the even sites separately from the uneven sites. Hence patterns for de Bruijn densities are now only meaningful when they just consist of states at sites with even indices.

5.5.2.2 Choosing a Phase Space Domain

Formally for CPA an invariant bounded part of phase space has to be identified. This is not possible for the isothermal Euler equation, as shocks and rarefaction waves leave no bounded domain invariant. However, the phase space is clearly bounded in practice. Typical methane pipes are operated at room temperature with pressure $p \approx 1 - 100$bar and with velocities below a certain erosional velocity of about $v_e = 15\frac{\mathrm{m}}{\mathrm{s}}$ [120]. With the ideal gas law we can calculate from this information that in practice $\rho \in [0.7, 70]\frac{\mathrm{kg}}{\mathrm{m}^3}$ and $q \in [-1000, 1000]\frac{\mathrm{kg}}{\mathrm{m}^2\mathrm{s}}$. In the following we will investigate a subspace Ω thereof for numerical simplicity: $\rho \in [0.7, 10.7]\frac{kg}{m^3}$ and $q \in [-1000, 1000]\frac{kg}{m^2 s}$. When determining the local CPA rule for a given partition there are some preimage domain combinations that are (partially) mapped outside by the dynamics. Typically this happens at boundary domains. We have to decide what to do with according test points. We can discard them or count them in the nearest boundary domain when determining the transition probabilities. In both cases, errors are introduced - but only for transitions that are unlikely to happen in practice.

5.5.2.3 The CFL-Condition

The CPA method is explicit, and hence time and space discretization of a hyperbolic PDE have to fulfill the CFL-condition for stability. For two phase space dimensions and one space dimension it reads

$$c := \max_{v \in \Omega}(\max(|\lambda_1(v)|, |\lambda_2(v)|)) \leq \frac{\Delta x}{\Delta t}.$$

It is based on considerations about the sites of variables that can influence variables at other sites in the future. This is completely independent of whether the phase space is continuous or discrete, and hence the consideratins can be transferred to CPA. For the Lax-Friedrichs scheme the CFL-condition will be necessary and sufficient for stability, if the system is linear. For non-linear systems like the isothermal Euler equation stability statements are much more difficult, but the CFL-condition is still our heuristic to determine orders of magnitude for our CPA discretization. For methane we calculate $a \approx 396\frac{\mathrm{m}}{\mathrm{s}}$ [109], and with the maximal erosional velocity $c = v_e + a = 411\frac{\mathrm{m}}{\mathrm{s}}$. We choose $\Delta x = 0.25$m and $\Delta t = 6 * 10^{-4}$s in accordance with the CFL-condition.

There are two ways to relieve the time step restriction. By enlarging the neighborhood length we can alter the CFL-condition. But this will lead to complexity issues as discussed in Sec. 5.1.2. Alternatively, we can increase Δx and Δt at the same time while leaving their fraction constant. However, this may lead to a direct loss of accuracy, even if we use a finer mesh in the preprocessing.

5.5.2.4 Boundary Conditions

So far we have only considered boundary conditions for the advection-reaction dynamics with $U = \{-1, 0\}$, where it was enough to just consider the inflow site. For the isothermal Euler equation with $U = \{-1, 1\}$ the definition of proper boundary conditions is more complicated [81]. We follow a common approach used e.g. in [55]. First, we describe the approach for regular finite volume methods and then adapt it to the stochastic case in CPA.

Consider a grid $I = \{1, ..., m\}$. The number of conditions that can be set at a boundary depends on the properties of the system's characteristics. We are in the subsonic regime $(v < a)$, where the characteristic speeds have opposite signs. Hence we need to impose one physical boundary condition at the inlet site 1, and one at the outlet site m. Instead of the characteristic variables we can also use other sets of variables as long as they provide enough information about the characteristics. We choose to fix the conserved variables, in particular the denisty ρ and thereby the pressure at the inlet, and the flow q at the outlet. This choice describes a pipe, where the pressure is imposed by a pump at the inflow, whereas a consumer at the outflow decides on the flow by his consumption. However, when we want to calculate the variables at sites 2 and $m - 1$ with our numerical scheme of Eq. 5.4, we need to provide additional information in form of numerical boundary conditions, i.e., for q at the inlet and ρ at the oulet. For this purpose we extrapolate the values at sites 2 and $m - 1$ to the sites 1 and m with order 0, i.e., $q_1^n = q_2^n$ and $\rho_m^n = \rho_{m-1}^n$ for all $n \in \mathbb{N}^+$. In the literature these sites are also known as ghost cells. They can potentially be used to couple other elements to the pipe. We note that we use the conservative variables for the physical boundary conditions and the extrapolation for simplicity, but there are many non-equivalent other methods; see [81] and references therein. Often the boundary conditions are also formulated by imposing the values directly at the cell boundaries between sites 1 and 2, and $m - 1$ and m, respectively; see e.g. [158]. This is conceptually not different from our scheme, because our extrapolation is of order 0 and constant throughout cells 1 and 2 anyway.

Now we describe how this approach can be transferred to stochastic boundary conditions in CPA. The concept is the same as in Sec. 5.1.4 for the advection equation. We consider only the inlet site 1, the outlet site m can be handled analogously. Assume that we are given a probability distribution $g_l \in D(E)$ over the density symbols as the physical boundary condition, and focus on minimal pattern length first. Our goal is to construct a distribution over the 2-dimensional discrete states at site 1. According to the above 0-order scheme we extract a probability distribution over the flow symbols from site 2 and copy it as the numerical boundary condition to site 1. This distribution can be calculated by marginalization. The product of the physical and the numerical boundary distribution then serves as the distribution over the 2-dimensional state space at site 1. Whenever we update site 2, we can now simply determine the product of the distributions at sites 1 and 3 as the preimage distribution and apply the regular local function to it.

We describe the idea also for larger de Bruijn densities with $V = \{-2, 0\}$. The concept can be extended to larger pattern lengths in a straightforward way. We define the distribution over patterns on V at site 3 as the product of a distribution over patterns with minimal length at site 1 and site 3. The distribution at site 3 is determined by marginalization of the distribution over patterns on V at site 5. The distribution at site 1 is determined as described above from the physical distribution at site 1 and the numerical distribution copied

from site 2. The numerical distribution of site 2 in turn can be determined by marginalization of the distribution over the pattern on V at site 4. The distribution at site 4 can then be updated with the local function, when the distributions at sites 3 and 5 are merged by α_U to a preimage distribution.

5.5.2.5 Imposing Conservation Properties

We also apply the postprocessing of Sec. 5.3 for the transition probabilities of the local function. Now we have 2 conserved quantities, ρ and q, and the conservation laws can be expressed with the discretization of Eq. 5.4 as

$$k_1(v^n_{i+U}) = \frac{1}{2}(\rho^n_{i-1} + \rho^n_{i+1}) - \frac{\Delta t}{2\Delta x}[f_1(v^n_{i+1}) + f_1(v^n_{i-1})], \quad k_1'(v^{n+1}_i) = \rho^{n+1}_i,$$
$$k_2(v^n_{i+U}) = \frac{1}{2}(q^n_{i-1} + q^n_{i+1}) - \frac{\Delta t}{2\Delta x}[f_2(v^n_{i+1}) + f_2(v^n_{i-1})], \quad k_2'(v^{n+1}_i) = q^{n+1}_i.$$

The algorithm is adapted slightly to our specific problem. Steps i), iii) and iv) have to be conducted for both k_1' and k_2', and k_1 and k_2, respectively. For step iv) we furthermore assume a uniform partition E of Ω and calculate

$$M_\rho = E[k_1(\rho^n_{i+U})] = \frac{1}{2}(E[\rho^n_{i-1}] + E[\rho^n_{i+1}]) - \frac{\Delta t}{2\Delta x}(E[q^n_{i+1}] - E[q^n_{i-1}]),$$
$$M_q = E[k_2(q^n_{i+U})] = \frac{1}{2}(E[q^n_{i-1}] + E[q^n_{i+1}]) -$$
$$- \frac{\Delta t}{2\Delta x}(E[(q^n_{i+1})^2]E[\frac{1}{\rho^n_{i+1}}] + a^2 E[\rho^n_{i+1}] - E[(q^n_{i-1})^2]E[\frac{1}{\rho^n_{i-1}}] + a^2 E[\rho^n_{i-1}]),$$

where all expected values $E[...]$ are taken with respect to the uniform distribution over Ω_φ for $\varphi \in E^U$. In the last line we used that ρ_i and q_i are independent for all $i \in I$ due to the joint uniform distribution. Note that in this notation we do not distinguish between the deterministic vectors ρ_i and q_i and the according random variables. The computation can be completed with the formulas

$$E[V^2] = \frac{1}{3(c-b)}(c^3 - b^3), \quad E[\frac{1}{V}] = \frac{\log(|c|) - \log(|b|)}{c-b}$$

for an arbitrary random variable $V : X \to \mathbb{R}$ with uniform distribution in $[b, c]$, where $b, c \in \mathbb{R}$. In step v) we choose d_e as the average of the formulas for the single conserved quantities, and in step vi) we impose two constraints for the exptected values, $M_\rho = m_\rho^T p$ and $M_q = m_q^T p$.

5.5.3 Numerical Results

In this section we give some numerial results of the CPA method for the isothermal Euler equation. The analysis in this part is divided into three steps. First, we investigate at a short pipe the phase space resolution, the pattern length and how to proceed with test points that are mapped outside of Ω. Second, we compare the scheme with and without postprocessing of the local function to a Monte Carlo computation of a long pipe. All these simulations are conducted at the example of a simple steady state. In the third step we then apply the CPA method to a wave.

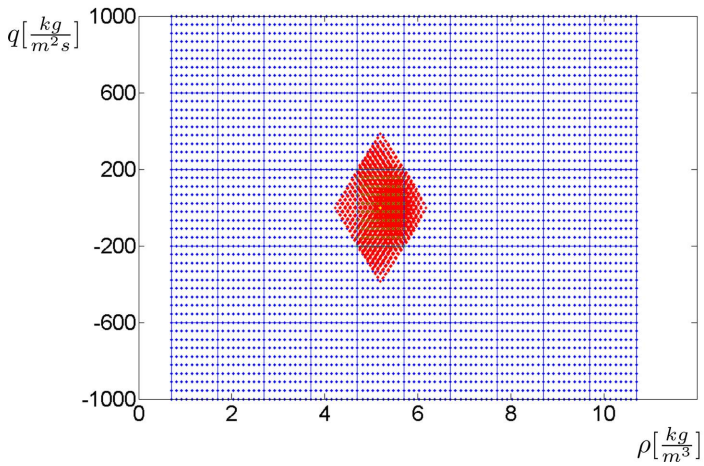


Figure 5.9: The invariant phase space Ω is partitioned with 10 symbols in ρ and 5 symobls in q-direction. The blue points are the uniformly distributed test points in each domain. The red points are the image points of the preprocessing after one time step for the exemplary initial configuration $\varphi = ((4,2),(4,2))$.

5.5.3.1 Short Pipes

Throughout the whole section we always choose a uniform parition and distribute 100 test points equidistantly in each coding domain in the preprocessing. Unless stated otherwise, furthermore we use threshold probability 0.0005 in the simulation and the discretization step sizes $\Delta t = 6 * 10^{-4}$s and $\Delta x = 0.25$m. Note that unlike in previous examples we calculate with the same step sizes for the preprocessing and the CPA, because we want to avoid that we miss potential shocks which can accumulate also on a small scale.

In the first step we consider a short pipe with $I = \{1, ..., 8\}$ and do not apply the post-processing yet. Our goal is to compare two types of phase space resolution and pattern lengths, and to investigate how to best proceed with image points that are mapped out of Ω in the preprocessing. We start with minimal patterns, $V = \tilde{V} = \{0\}$, and a partition consisting of 10 symbols in ρ- and 5 in q-direction; see Fig. 5.9. We label the domains with (i, j), where i encodes the q- and j the ρ-direction, and call the resolution a 10×5-resolution. Fig. 5.10a- 5.10b show the simulation results of the steady state after $t \approx 0.25$s with the first resolution. In both cases we choose as physical boundary conditions the 4. symbol with probability 1 in the q-direction at the inlet, site 1, and the 2. symbol with probability 1 in the ρ-direction at the outlet, site 8. The initial condition is symbol $(4, 2)$ at all interior sites. The difference between both simulations is the underlying local function. In Fig. 5.10a the test points that are mapped out of Ω are discarded, and in Fig. 5.10b they are counted in the nearest coding domain. No difference can be observed for both variants, the distinction is not relevant in this part of phase space. We only expect different behavior in the boundary domains.

Now we change the setup to the higher 10×10-resolution, but keep the minimal pattern length. Test points that are mapped out of Ω are not discarded, but that is not important here. We choose the 4. and 5. symbol, each with probability 0.5, as the physical boundary condition in ρ-direction. Fig. 5.10c shows the simulation result. It leads to a similar result as with the lower resolution, but the distribution in ρ-direction is sharper. It will turn out

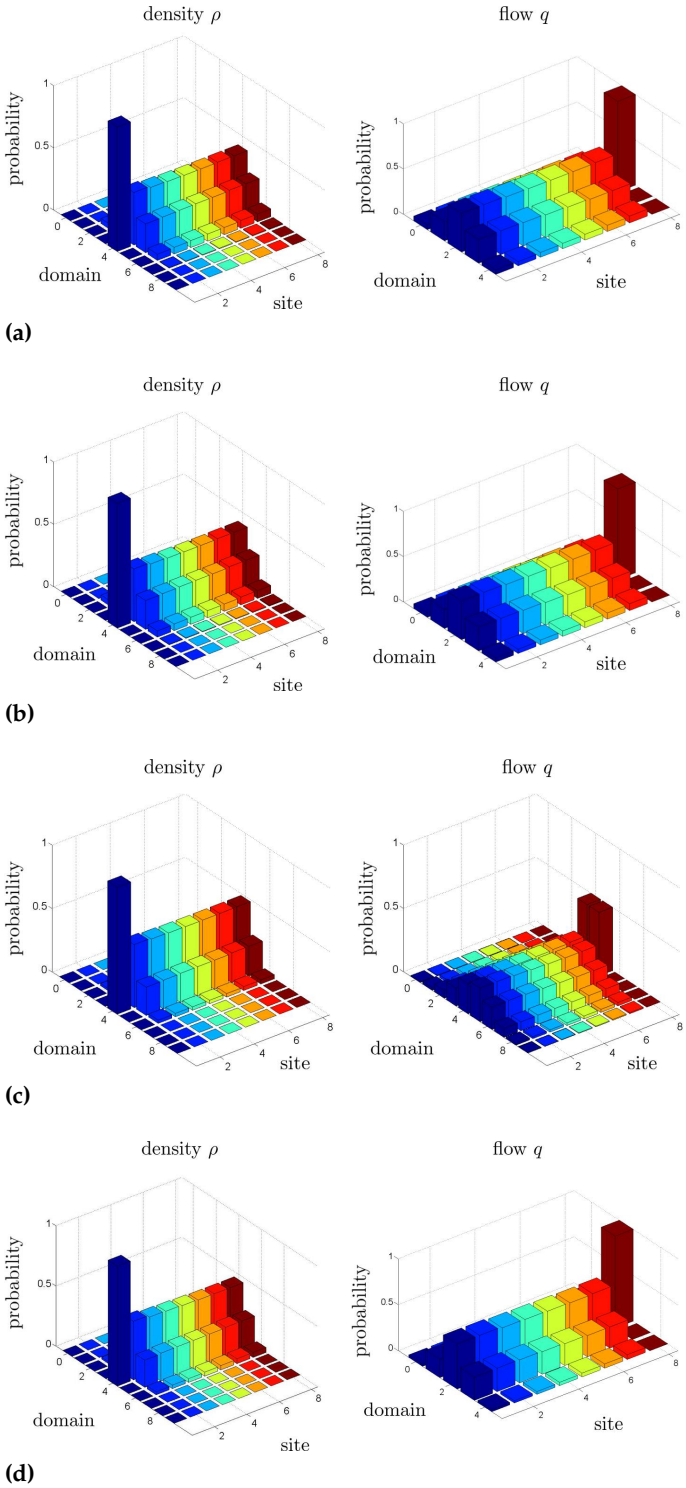


Figure 5.10: Steady state results from different CPA simulations for short pipes; see the text. The left diagram always depicts the density, the right the flow distribution. (a) 10×5-resolution, test points are discarded, (b) 10×5-resolution, test points counted in nearest coding domain, if mapped oustide of Ω, (c) 10×10-resolution, (d) 10×5-resolution, patterns on $V = \{-2, 0\}$

in the next step that this is closer to the Monte Carlo result for longer pipes.

We also analyze the influence of the pattern length for the lower phase space resolution. In particular, we conduct a CPA simulation with $V = \{-2, 0\}$ and $\tilde{V} = \{0\}$, which is the next bigger pattern length in the Lax-Friedrichs-scheme. The results are shown in Fig. 5.10d. They agree qualitatively with the results for the minimal pattern length, but the distributions are again a little sharper.

In conclusion of the first step the precise handling of test points does not have significant impact on the simulation result. But increasing state space resolution and pattern length improves the results. In all following simulations in this section we discard the test points and simulate with the minimal pattern length.

5.5.3.2 Long Pipes

In the second step we simulate longer pipes with $I = \{1, ..., 32\}$ and $t \approx 0.25s$ to study conservation properties in the steady state. For this purpose we compare CPA simulations with and without postprocessing to a Monte Carlo computation. The Monte Carlo computation also allows for a validatio of the results of the first step.

In particular we simulate with a 10×5-phase space resolution and a probability threshold of 0.0005. First, we do not apply the postprocessing for conservation of the mean. As physical boundary conditions we choose with probability 1 the 4. symbol in the ρ- and the 2. symbol in the q-direction. Initially the configuration at all sites is $(4, 2)$, where we use the same notation as above. The steady state result after $t = 0.25$s is shown in Fig. 5.11a. We observe that we encounter problems with conservation in ρ at sites far away from the physical boundary condition in ρ at site 1. An analogous observation can be made for q. We repeat the simulation, this time with a local function that was corrected by the postprocessing and a probability threshold of 10^{-7} in the simulation. We note that for many preimages that consist of boundary domains no feasible solution is found in the optimization; in these cases we keep the probabilities from the regular preprocessing. The result is shown in Fig. 5.11b: it is almost identical with the result of a simulation without postprocessing.

We also conduct a simulation with the higher 20×10-resolution and a probability threshold of 10^{-7}. To make it comparable with the former simulations, this time we apply stochastic boundary conditions: symbols 8 and 9 of the ρ-component at site 1 have probability $\frac{1}{2}$ each, and symobls 4 and 5 of the q-component at site m have probability $\frac{1}{2}$ each. Initially we set the state at all interior sites to $(8, 4), (8, 5), (9, 4)$ and $(9, 5)$ with probability $\frac{1}{4}$. The result of a simulation for $t = 0.25$s is shown in Fig. 5.11c. We represent the results in the lower 10×5-resolution by merging neighboring domains. The problem with conservation of the expected value vanishes in the ρ-, but stays in the q-direction. We provide two potential explanations for this effect. The first explanation is based on the observation that the boundary domains are occupied with much lower probabilities in the simulation with higher resolution. It could be that the errors from boundary domains accumulate and lead to the conservation problems. Second it is possible that our straightforward treatment of boundary conditions gives rise to the effect, and that more sophisticated schemes have to be applied. This would especially explain, why the errors appear at the boundaries.

As the last element of the second step we conduct a Monte Carlo computation with the identical finite volume scheme and time and space discretization as for CPA. The initial

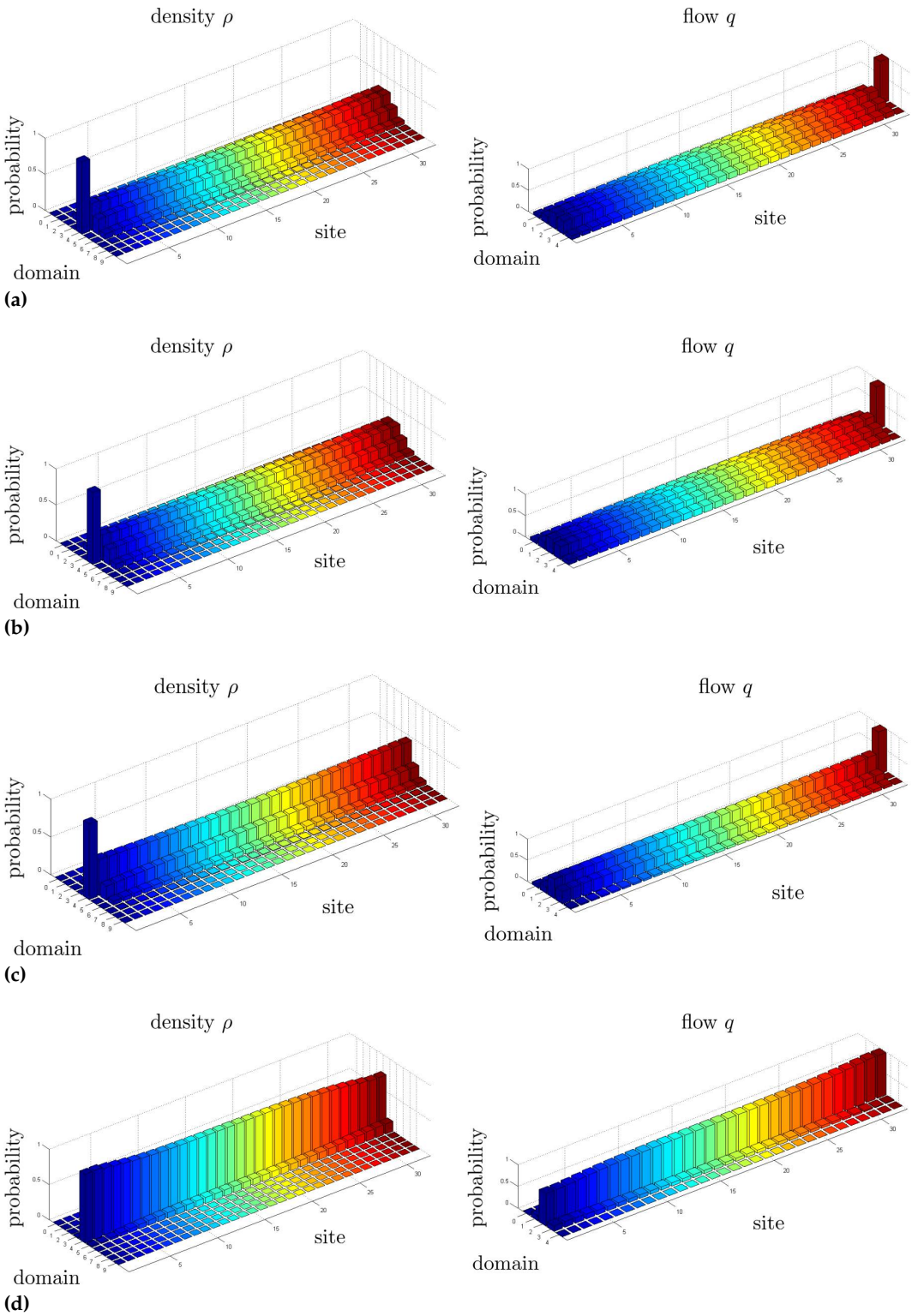


Figure 5.11: Steady state results from different CPA simulations for short pipes; see the text. In (a) the original CPA method is used, in (b) the postprocessing of the transition probabilities is applied. (c) is calculated with a higher resolution, but represented in the same as the other graphs. (d) is a Monte Carlo computation.

and boundary values are the same as in the experiments in the second step. One deterministic run takes as an input a realization from the uniform distribution in the initial domain at each interior site, and also one from the boundary domains at each time step. The result for 10^6 samples can be found in Fig. 5.11d. We observe that the expected value is conserved in both directions, and that the variance is much smaller than in the results of the CPA method. This is very different from the advection-reaction applications discussed previously and can be explained by the very different phase space structure. In the isothermal Euler equation every point in phase space is a steady state, and hence the phase space is not contracting. Even more, little differences in the states at neighboring sites can lead to shocks. This is observed in Fig. 5.9 for a small time step and also in the Monte Carlo computation as a blow-up for longer simulation times. In fact, some of the deterministic runs do not converge, i.e., the values for ρ and q go to infinity. This behavior, however, does not appear in practice; i.e., CPA impose in this application more numerical uncertainty that exists in practice. Nevertheless, with increasing phase space resolution and pattern length the variance decreases as expected from the convergence result in Sec. 4.3.

5.5.3.3 Waves

Although there are some problems of the CPA approach for the isothermal Euler equations, CPA are capable of recovering shocks or rarefaction waves as their specific dynamical features. This is shown now in the third step. We choose a 20×10-discretization and simulate with a probability threshold of 0.002. The initial and boundary conditions are depicted in Fig. 5.12a. They are identical at the sites 1 to 15, and also at sites 16 to 32, but there is a discontinuity in the density between site 15 and 16. In the simulation the wave with positive speed starts to travel to the right, and after 13 time steps it has almost reached the boundary site 32; see Fig. 5.12b. There it is reflected and travels back in the other direction. A snapshot after 35 time steps is depicted in Fig. 5.12c. In contrast, the wave with negative speed almost vanishes. Furthermore, in Fig. 5.13 we show a snapshot at $t = 13\Delta t$ of the solution of the deterministic problem given by the mean of the initial values of the CPA problem,

$$\rho(x,0) = \begin{cases} 5.2\frac{\mathrm{kg}}{\mathrm{m}^3} & \text{if } 0\mathrm{m} \leq x \leq 4\mathrm{m} \\ 6.2\frac{\mathrm{kg}}{\mathrm{m}^3} & \text{if } 4\mathrm{m} < x <\leq 8\mathrm{m} \end{cases}, \quad q(x,0) = \begin{cases} 0\frac{\mathrm{kg}}{\mathrm{m}^3} & \text{if } 0\mathrm{m} \leq x \leq 4\mathrm{m} \\ 400\frac{\mathrm{kg}}{\mathrm{m}^2\mathrm{s}} & \text{if } 4\mathrm{m} < x \leq 8\mathrm{m} \end{cases}$$

and according boundary values. It is computed with a local Lax-Friedrich scheme with CWENO-reconstruction [158] and the same time and space steps as in the CPA computations. The result confirms at least the qualitative structure of the CPA solution.

5.6 Conclusion and Outlook

This section contains practical aspects of uncertainty propagation with CPA. We have applied the method successfully to advection-reaction systems with two examples, arsenate adsorption and bacterial regrowth in drinking water pipes. Interesting non-linear behavior has been observed, and the results have been confirmed by Monte Carlo simulations. Then we have suggested a way to couple pipes and have extended the method to water

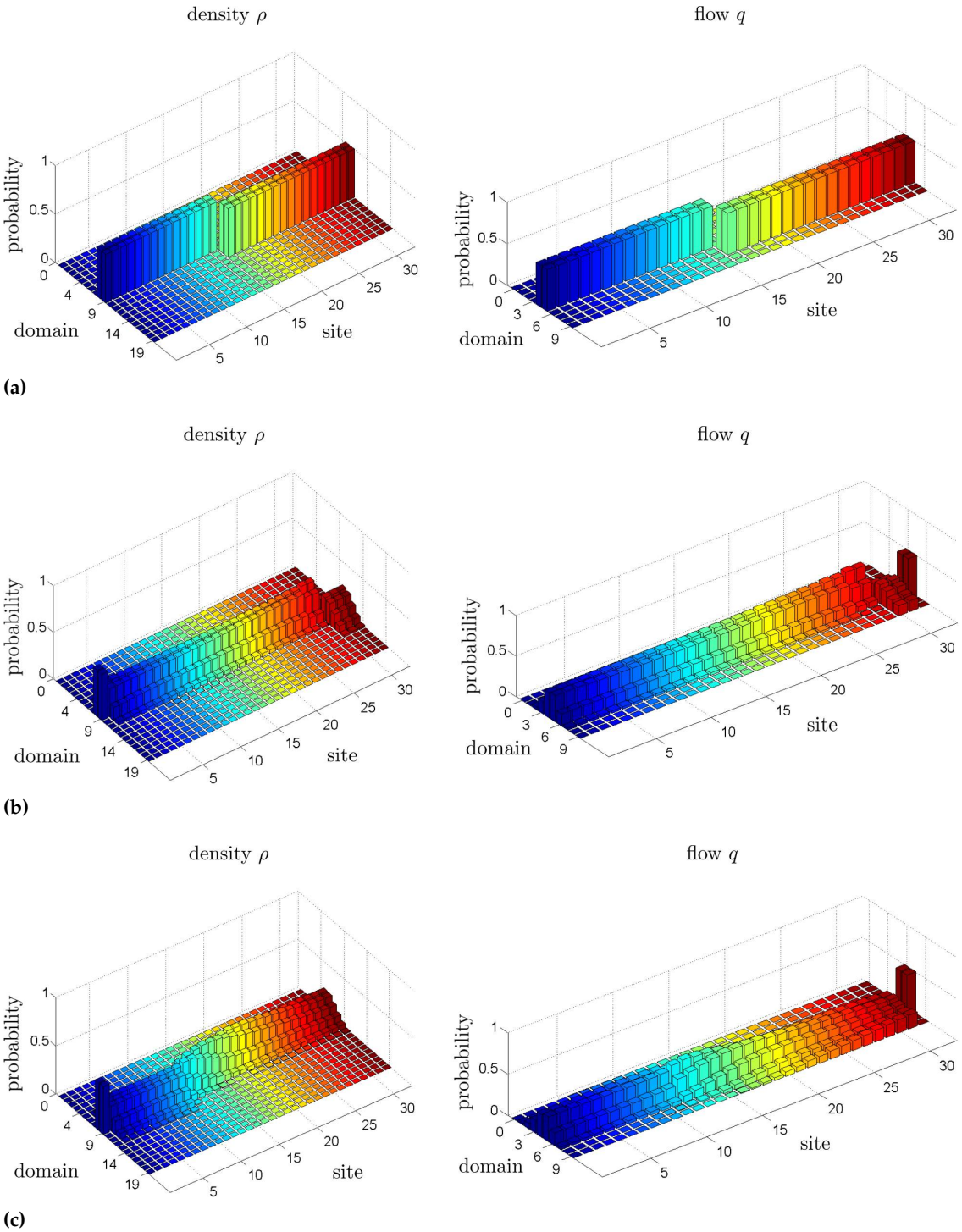


Figure 5.12: A wave is considered with a CPA. (a) shows the initial state, (b) the state after 13 time steps and (c) the state after 35 time steps.

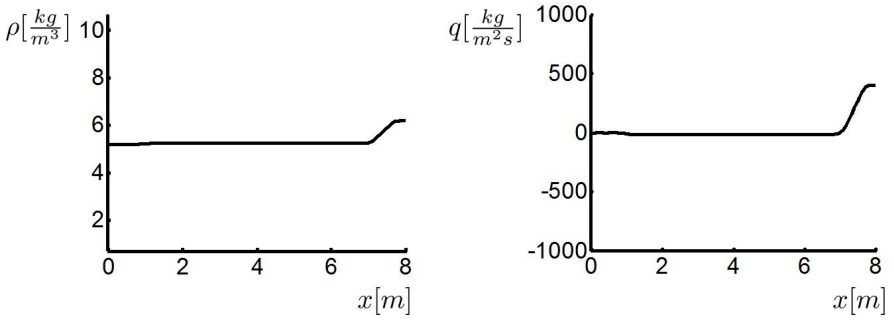


Figure 5.13: The figure shows the deterministic solution of a wave after 13 time steps. The configuration corresponds to the mean of the CPA setup; see the text.

grids. Also, we have addressed a more complex hyperbolic equation with the method, the isothermal Euler equation of fluid dynamics.

The applications highlight three main strenghts of the CPA approach. First, the CPA method can be divided into a preprocessing step, in which the underlying PDE system is translated into a cellular probabilistic automaton, and a simulation step, in which initial and boundary conditions are evolved. The simulation is fast in relation to the preprocessing. In the reaction-advection applications the whole method is faster than real-time, but slower for the isothermal Euler equation. A Monte Carlo evaluation needs approximately the same time as the preprocessing, but has to be conducted from the scratch for each new configuration of initial or boundary values. Second, CPA operate on discrete states instead of on the continuous phase space. Since the discrete states can be interpreted as risk levels, fast uncertainty propagation directly on this simplified state space suits industrial demands. And third, we have paved the way towards the handling of spatio-temporal processes with more involved stochastic influence. More precisely, it has been shown how to deal with white noise boundary conditions, an important topic, for example, in contaminant transport modeling.

However, in the applications we have alo encountered some practical difficulties of the approach. They are fourfold. First, it has turned out that simulations of long pipes with the CPA method violate conservation laws. With an appropriate adaptation we have managed to overcome this problem. Second, the method suffers from complexity problems for high phase space dimensions as well as for neighborhood and pattern lengths. The problem can be mitigated by using a sparse implementation or by formulating a low-dimensional surrogate model before translating a system into a CPA [145]. Moreover, sophisticated ideas from set oriented numerics can be adapted to the preprocessing [73]. Third, at the hand of the isothermal equation we found that a rich phase space structure can have limitations on the accuracy of the method. To increase precision we suggest to pursue the idea of non-rectangular state-space domains like it was introduced for CNDA in 3.3.3. And fourth the last application has also underlined that we face strong time step restrictions with our explicit scheme.

From a practical point of view it will be interesting to further work out the simulation of contamination in water grids in future work. An application can be a monitoring device. If we find a way to improve the simulation of the isothermal Euler equation, it will be interesting to combine it with the extension to grids. A decision support system for

load relieving of overflow chambers in waste water grids can be developed. Also, these ideas can lead to a leak detection system for gas, drinking or process water grids, where traditional approaches face difficulties with severe demand uncertainties [60,148].

6 Bayesian Inference with Cellular Probabilistic Automata

This chapter shows how measurement data can be integrated in the completely discrete and probabilistic framework developed in the preceding chapters. More precisely, we introduce a novel numerical approach to inverse problems in spatio-temporal partial differential equations (PDEs). The problem is considered in a Bayesian inference context, and the main idea is to translate the PDE into a state-discrete dynamic Bayesian network (DBN). For this translation we suggest to use the method of cellular probabilistic automata (CPA) that was proposed in Chap. 4. First, the PDE is discretized in time and space by a finite difference method, and then tools of set oriented numerics and cellular automata theory are used to also discretize the state space while exploiting the locality of the interaction. The result is a completely discrete system which consists of graph-based time slices with local transition probabilities between them. It approximates the time evolution of the PDE and can be interpreted as a state-discrete DBN.

In an inverse problem one aims at deriving knowledge about model parameters or the system's state from measurement results. We are interested in parameter estimation, which we recast as a filtering problem. Filtering is a special case of inference, where data are integrated online into a belief state about a state variable. This can be achieved conveniently in a DBN model, because there a vast pool of sophisticated Bayesian inference algorithms exists. We identify certain requirements for a filtering algorithm in our special DBN to optimize its performance. As it is closely related to the locality idea of the CPA construction we then choose the BK algorithm. It approximates the exact belief state in each time step by a factorized one.

However, it can be read independently of the previous chapters and is structured as follows. In Sec. 6.1 we review DBNs, CPA and the Bayesian setting for inference in a unified notation. A particular focus is on the BK algorithm. Sec. 6.2 is the main part, in which we bring together these concepts to solve an exemplary parameter estimation problem for PDEs. To demonstrate our ideas we apply the scheme to the problem of arsenate advection and adsorption in a water pipe with a contamination source at the entrance in Sec. 6.3. From measurements of the downstream boundary condition we infer the strength of the upstream arsenate source. For this purpose we generate measurement data with a CPA forward simulation and use the BNT implementation [132] of the BK algorithm for inference. In Sec. 6.4 we finally give our conclusions.

6.1 Methods

In this section we first introduce DBNs. Then we review cellular probabilistic automata (CPA) as introduced in Chap. 4 and compare them to DBNs. It will turn out that the structure of CPA is very similar to the structure of DBNs. Third we give a short introduction

to Bayesian inference and then focus on a special inference algorithm in DBNs, the BK algorithm, in the fourth part.

We adopt the notational conventions of the probabilistic graphical models community in this chapter: We write $\mu(V) = \mu_V$ and $\mu(V|V') = \mu_{V|V'}$ for the (conditional) distributions of random variables V, V' on a probability space $(X, \mathcal{A}, \mu)$. Similarly, $\mu(V|V' = v')$ means the *conditional distribution of V given $V' = v'$*. Products of distributions are understood as products of real numbers when the distributions are evaluated for according events. Technical details can be worked out along the lines of Sec. 2.4.

6.1.1 Dynamic Bayesian Networks

We introduce DBNs along the lines of [102]. Let $(X, \mathcal{A}, \mu)$ be a probability space and $(E_i, \mathcal{P}(E_i))$ measurable spaces for finite E_i, where $\mathcal{P}(E)$ is the power set of E and $i \in \{1, ..., m\}$. Consider the stationary Markovian stochastic process $(V^t)_{t\in\mathbb{N}_0}$, where $V^t = (V_1^t, ..., V_m^t)$ and each V_i^t is a discrete random variable over $(X, \mathcal{A}, \mu)$ with values in $(E_i, \mathcal{P}(E_i))$; see Def. 2.14. The time-independent transition distribution is given by a conditional distribution $\mu(V|V')$. We focus first on the structure of one time slice.

Definition 6.1. Let $V = (V_1, ..., V_m)$ be a set of random variables and G a directed acyclic graph, in which the nodes correspond to the variables. If a probability distribution $\mu(V)$ can be expressed as a product of conditional distributions

$$\mu(V) = \prod_{i=1}^{m} \mu(V_i|\mathrm{Pa}(V_i)),$$

the pair (G, μ_G) is called a *Bayesian network*. Here $\mathrm{Pa}(V_i)$ is short-hand for the variables at the parent nodes of node i, and μ_G denotes the set of conditional distributions associated with G's nodes.

It is enough to specify an initial distribution and a transition model to describe infinite time evolutions.

Definition 6.2. A *dynamic Bayesian network (DBN)* is a pair $(B_0, \mu(V|V'))$. B_0 is a Bayesian network over random variables $V^0 = (V_1^0, ..., V_m^0)$ and specifies an initial distribution. $\mu(V|V')$ is a time-independent and factorized transition distribution between also m-dimensional random variables V, V', each associated with m nodes,

$$\mu(V|V') = \prod_{i=1}^{m} \mu(V_i|\mathrm{Pa}(V_i)).$$

$\mathrm{Pa}(V_i)$ is given via a set of *inter-time-slice edges* between the nodes of V' and V, and via *intra-time-slice edges* between nodes of V. A *trajectory* starting with the distribution over V^0 consists of a series of distributions over $V^0, V^1, V^2, ...$ The distributions are calculated by recursive application of the transition rule,

$$\mu(V^{t+1} = v) = \sum_{v'} \mu(V = v|V' = v')\mu(V^t = v')$$

for $t \in \mathbb{N}_0$.

By repeated application of the transition model we can therefore evolve the initial distribution over V^0 with a potentially stochastic forward model. For $T \in \mathbb{N}_0$ we call the static Bayesian network of the whole time-evolution the *unrolled DBN*. It consists of the graph over random variables $(V^0, ..., V^T)$, where all nodes are equipped with the according intra- and inter-time-slice edges of the transition model and the initial distribution. The conditional distributions of the static network are given by the conditional distributions associated with all intra- and inter-time-slice edges.

Note that in DBNs the probability distribution after some time steps does not have to be in the same product structure as the initial distribution. Correlations between variables in one time slice beyond the product structure can be induced by the transition model, and this effect is called *entanglement*.

6.1.2 Cellular Probabilistic Automata

Cellular probabilistic automata (CPA) are time-, space- and state-discrete systems for density based uncertainty propagation under distinct PDE dynamics. These automata have a similar structure as DBNs and approximate the PDE. We review the construction of Chap. 4 in the language of probabilistic graphical models to settle a unified notation for the following.

Consider the PDE

$$\partial_t v = \tilde{h}(\partial_{xx} v, \partial_x v, v), \quad v(x,t) \in \Omega \subseteq \mathbb{R}^n$$

defined on a one-dimensional space-domain $x \in [a, b]$ for $a, b \in \mathbb{R}$, where Ω is bounded and positively invariant. CPA are derived in three steps: time and space discretization, state discretization and use of locality.

First, we apply a finite difference method with space discretization $\Delta x = \frac{b-a}{m-1}$, where $m \in \mathbb{N}$, $m \geq 2$, and time step τ to the PDE. Under some assumptions the PDE can be translated into a flow Φ^τ with only local interaction this way; see Def. 4.1. The dynamics is described by a graph consisting of several time slices, where each has nodes $I = \{1, ..., m\}$. There are no edges in each time slice, but directed ones between subsequent time slices: for each node in a time slice the finite difference stencil defines a set of parent nodes in the previous time slice, and edges are drawn to the appropriate children; see Fig. 6.1. Exceptions are of course the initial time slice and the boundary nodes. To account for uncertainties in the initial and boundary values we describe the system's state in each time slice by a probability distribution over random variables $V = (V_1, ..., V_m)$ that are associated with the nodes. The time evolution of this distribution is given by the Frobenius-Perron operator (FPO) [106].

Second we partition the state space Ω with a coding map $T : \Omega \to E$, where E is finite. Because of the spatial shift-invariance this naturally induces a partition E^I of Ω^m with coding map $\hat{T}$:

$$\hat{T} : \Omega^m \to E^I, v \mapsto \hat{T}(v) \text{ with } (\hat{T}(v))_i = T(v_i) \text{ for } i \in I.$$

The distribution over $\{V_1, ..., V_m\}$ on Ω^m can be approximated by a piecewise constant distribution by averaging in every coding domain. The latter in turn naturally corresponds to a distribution on E^I. Formally this leads to new random variables with discrete values in E, which we also denote $V_1, ..., V_m$ to keep the notation simple. Like in set oriented

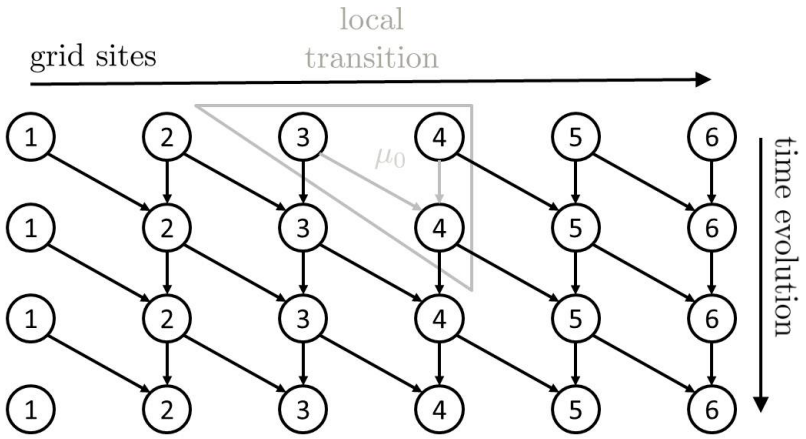


Figure 6.1: An example of a graph that is constructed from a PDE to describe its time evolution; see the text. Here $I = \{1, ..., 6\}$, and the finite difference stencil contains the left neighbor and the node itself in the previous time slice. This gives rise to a local transition rule μ_0.

numerics we approximate the FPO by a time-discrete Markov chain working with the approximate distributions over E^I. The Markov chain is specified by transition probabilities between different states of E^I, i.e., between different coding domains.

Third we take into account that the interaction is local and identical at every node in I away from the boundaries. We can hence approximate the global transition probabilities by a product of the same local transition probabilities at every node. Such a locality concept is at the core of cellular automata theory. The local transition probabilities from the variables at a node's parents to the variable at node i itself can be specified as a conditional probability distribution. It is called the *local transition rule* and describes how much of a local preimage coding domain is mapped to a local image coding domain in one time step τ,

$$\mu_0(V_i = v_i | \mathrm{Pa}(V_i) = v') = \frac{\lambda(\Omega_{v'} \cap \Phi^{-\tau}(\Omega_{v_i}))}{\lambda(\Omega_{v'})}.$$

Here λ is the Lebesgue measure, and $\Omega_v \subseteq \Omega^m$ is the subset of all continuous states which are encoded by $v \in E^J$ on $J \subseteq I$. The definition is independent of i because of the shift-invariance of the transitions as long as i is a site away from the boundaries. Note that at the boundaries other suitable local transition rules have to be defined; see Chap. 4 for details. In practice the transition probabilities can be calculated by mapping test points from each local preimage domain with the stencil rule and determining the percentage of points that are mapped into the several local image domains.

We also respect the locality in the representation of the distributions over V as factorized distributions. A *factorized distribution* over $V = (V_1, ..., V_m)$ is given as a product of marginals over subsets $W_k \subseteq \{V_1, ..., V_m\}$ for $k = 1, ..., K$ with $\cup_{k=1}^K W_k = \{V_1, ..., V_m\}$,

$$\mu(V_1, ..., V_m) \propto \prod_{k=1}^{K} \mu(W_k).$$

The subsets need not be disjoint, and the proportionality constant is given by normalization from conditioning on the overlap; see Def. 4.12. The distribution over all variables

or any subset thereof is then specified uniquely by the $\mu(W_k)$. We call a distribution fully factorized if it is factorized with $W_1 = \{V_1\}, ..., W_m = \{V_m\}$.

Definition 6.3. A *Cellular Probabilistic Automaton (CPA)* is a pair (B_0, μ_0). B_0 is a state-discrete Bayesian network over the graph I with random variables $V^0 = (V_1^0, ..., V_m^0)$ associated with the nodes and no edges. It represents the fully factorized initial distribution. μ_0 is a local transition rule. A *trajectory* starting with the fully factorized distribution over V^0 consists of a series of fully factorized distributions over $V^1, V^2, ...$ The marginals are calculated by recursive application of the local transition rule at every node,

$$\mu(V_i^{t+1}) = \sum_{v'} \mu_0(V_i^{t+1}|\mathrm{Pa}(V_i^{t+1}) = v')\mu(\mathrm{Pa}(V_i^{t+1}) = v').$$

The structure is similar to DBNs, but there is one major difference. With CPA one maintains a factorized distribution as an approximation of the global distribution during time evolution, whereas nodes can get entangled in DBNs. In the language of Chap. 4 a DBN is a discretized FPO. Furthermore in CPA the transition rule is identical at every node.

If we choose a more coarse-grained factorization, i.e., if we map whole patterns of local states at once in the local transition rule and factorize our distribution accordingly, we can preserve more correlations. Note that such patterns would involve the definition of suitable intra-time-slice edges. In the limit of large patterns and a high state space resolution we get back the exact time evolution with the FPO. The technical details for these considerations can be found in Chap. 4. In practice it is often sufficient to work with very small pattern lengths to already capture the most important correlations in a time evolution.

6.1.3 Bayesian Inference

In this section we review the Bayesian setting to infer model parameters from measurement data in an inverse problem on the basis of [61]. Bayesian inference is based on the *Bayesian interpretation* of a probability distribution as a *degree of belief* about the according model parameter [13, 156]. This interpretation opposes the *frequentist viewpoint*, in which it is the limit distribution of outcomes of a large number of identical experiments. In the Bayesian setting model parameters and measurement data are hence described by random variables, and the distribution associated with the parameters expresses the state of knowledge about possible values. Bayes' fundamental theorem describes how this degree of belief is updated by measurements. We only need the version for state-discrete random variables in the following, although the statement holds in more generality.

Theorem 6.4. *(Bayes)* Let $(X, \mathcal{A}, \mu)$ be a probability space. Consider state-discrete random variables M and D for the model parameters and the model data with probability distributions $\mu(M)$ and $\mu(D)$, respectively. Then for data d with $\mu(D = d) \neq 0$

$$\mu(M = m|D = d) = \frac{\mu(D = d|M = m)\mu(M = m)}{\mu(D = d)}.$$

In Bayesian inference $\mu(M = m)$ is called the *prior density*. It summarizes all information about the system before the measurement process. $\mu(M = m|D = d)$ is the *posterior density* and represents the state of knowledge after having taken into account measurement

information. The forward model, the model discrepancy and a noise model for the measurements consitute the *likelihood* $\mu(D = d|M = m)$. $\mu(D = d)$, finally, is called the *evidence* or *marginal likelihood*.

The various methods described in Sec. 1.2.3 can be applied to evaluate the posterior distribution or suitable statistical properties as a solution to the inference problem. However, for expensive, e.g. PDE-based forward models it is still a big challenge to make Bayesian inference computationally tractable this way.

6.1.4 The Boyen-Koller Algorithm

In a dynamic problem we have to deal with a series of measurements and include it into our belief about the model parameters. In principle this can be done by multiple use of Bayes' theorem, Thm. 6.4. We review the inference problem in state-discrete DBNs and in particular the BK algorithm in the following on the basis of [21, 102].

Assume that our system is described by a state-discrete DBN (B_0, μ) over random variables $V = (V_1, ..., V_m, D)$, where D describes the measurement data. Our goal is to keep track of the so-called *belief state* $\sigma^t(V^t) = \mu(V^t|D^1 = d^1, ..., D^{t'} = d^{t'})$ over time $t \in \mathbb{N}$, where $\{d^1, ..., d^{t'}\}$ are measurement data and $t' \in \mathbb{N}$. We write $D^{1:t'} = d^{1:t'}$ as an abbreviation for $D^1 = d^1, ..., D^{t'} = d^{t'}$ in the following. There are three variants of the inference problem in such a dynamical system: For $t < t'$, we speak of *smoothing*, for $t = t'$ of *filtering* or *tracking*, and for $t > t'$ of *prediction*. Here the focus is on the filtering problem.

We start with $\sigma^0(V^0) = \mu(V^0)$. Assuming that $\sigma^t(V^t)$ is known, we want to determine $\sigma^{t+1}(V^{t+1})$. First, the state is propagated according to the transition model to the so-called *prior belief state* $\sigma^{\bullet t+1}(V^{t+1})$,

$$\begin{aligned}\sigma^{\bullet t+1}(V^{t+1}) &= \mu(V^{t+1}|D^{1:t} = d^{1:t}) \\ &= \sum_v \mu(V^{t+1}|V^t = v, D^{1:t} = d^{1:t})\mu(V^t = v|D^{1:t} = d^{1:t}) \\ &= \sum_v \mu(V^{t+1}|V^t = v)\sigma^t(V^t = v),\end{aligned}$$

and then the current observation is taken into account by Bayes' theorem, Thm. 6.4, to calculate the belief state

$$\begin{aligned}\sigma^{t+1}(V^{t+1}) &= \mu(V^{t+1}|D^{1:t+1} = d^{1:t+1}) \\ &= \frac{\mu(D^{t+1} = d^{t+1}|V^{t+1}, D^{1:t} = d^{1:t})\mu(V^{t+1}|D^{1:t} = d^{1:t})}{\mu(D^{t+1} = d^{t+1}|D^{1:t} = d^{1:t})} \\ &= \frac{\mu(D^{t+1} = d^{t+1}|V^{t+1})\sigma^{\bullet t+1}(V^{t+1})}{\mu(D^{t+1} = d^{t+1}|D^{1:t} = d^{1:t})}.\end{aligned}$$

Note that the denominator is just a normalization constant that is given by $\mu(D^{t+1} = d^{t+1}|D^{1:t} = d^{1:t}) = \sum_v \mu(D^{t+1} = d^{t+1}|V^{t+1} = v, D^{1:t} = d^{1:t})\sigma^{\bullet t+1}(V^{t+1} = v)$. Both the transition and the conditioning step can be interpreted and executed by a standard junction tree algorithm for inference in static Bayesian networks [102]. This leads to a convenient filtering method without approximations. However, handling the belief state over m

variables can be very expensive in practice. The idea of the BK algorithm is similar as in CPA to only maintain an approximate factorized belief state $\tilde{\sigma}^t$ instead of the exact one.

A factorized belief state $\tilde{\sigma}^t$ is a belief state which is represented as a product of marginals over so-called clusters $W_k \subseteq \{V_1, ..., V_m\}$ for $k = 1, ..., K$ with $\cup_{k=1}^K W_k = \{V_1, ..., V_m\}$,

$$\tilde{\sigma}^t(V^t) \propto \prod_{k=1}^{K} \mu(W_k^t | D^{1:t} = d^{1:t}).$$

Like in CPA the subsets need not be disjoint, and the proportionality constant is given by normalization from conditioning on the overlap. The BK algorithm adds a postprocessing to each time step in the exact filtering scheme: after a factorized belief state $\tilde{\sigma}^t$ has been updated with the above exact transition and conditioning rule, the resulting state $\hat{\sigma}^{t+1}$ is tranformed into a factorized state $\tilde{\sigma}^{t+1}$. This new approximate belief state is defined via $\tilde{\sigma}^{t+1}(W_k) = \hat{\sigma}^{t+1}(W_k)$ for all $k = 1, ..., K$, and the whole distribution is again the (normalized) product over these factors. Like in the exact case, both steps can be executed by a standard junction tree algorithm, when some clique of the tree contains W_k^t, and some contains W_k^{t+1} for all $k = 1, ..., K$.

Filtering only requires this *forward* or *upward pass*. However, we note that a *backward* or *downward pass* can be defined similarly and extend the scheme to the smoothing task.

6.2 Cellular Probabilistic Automata as Dynamic Bayesian Networks

We are interested in the exemplary problem of inferring the upstream boundary condition in a dynamical PDE problem with measurements of the downstream boundary values. In particular, we consider a stochastic upstream boundary condition in the sense of Sec. 5.1.4, i.e., a sequence of i.i.d. random variables, where each describes the boundary condition during one time step. The according distribution is assumed to be known up to the mean, and the goal is to infer the mean of the distribution from downstream measurement results.

We propose an inference method that uses the CPA concept for this purpose. The simplest CPA approach would be to replace the stochastic forward model by a CPA surrogate model similar to [118, 119] and sample the likelihood. Instead, we suggest to recast the parameter estimation as a filtering problem, where the parameter is a variable that does not change in time. The problem can then be solved by composing the methods described in the previous sections: First, we translate the PDE into a CPA which is time-, space- and state-discrete. Then we interpret the CPA as a state-discrete DBN and apply the BK algorithm for filtering to it. This approach is presented in more detail in the following.

Consider the PDE

$$\partial_t v = \tilde{h}(\partial_{xx} v, \partial_x v, v), \quad v(x,t) \in \Omega \subseteq \mathbb{R}^n,$$

defined on a one-dimensional spatial domain $x \in [a, b]$ for $a, b \in \mathbb{R}$ and on the bounded and positively invariant phase space Ω. As described in Sec. 6.1.2, by applying a finite difference scheme to it for time- and space-discretization we obtain a graph with a set of nodes I in every time step. The random variables $V^t = (V_1^t, ..., V_m^t)$ associated with

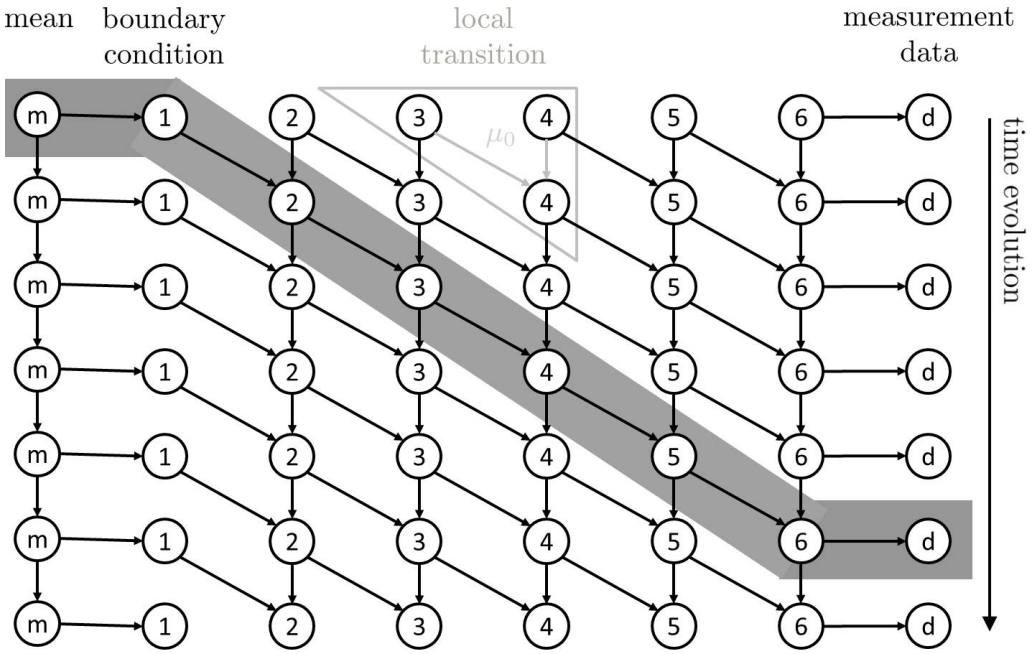


Figure 6.2: An example of a DBN that is constructed from a PDE to solve the filtering problem. It is the goal to infer the mean m from the measurement data d. In this example we take $I = \{1, ..., 6\}$ and assume that the finite difference stencil contains the left neighbor and the node itself in the previous time slice. Each row is a natural time slice, whereas the grey band highlights a characteristic line from the perspective of an advected species; see the text.

the nodes describe the time evolution of the system. Furthermore we discretize the state space and use the locality property of the interaction to construct a local transition rule for CPA. We choose the fully factorized version and introduce inter-time-slice edges between subsequent time slices according to the stencil; see Fig. 6.2. A global transition model is constructed by application of the local transition model at each node. Thus the resulting CPA can formally be interpreted as a state-discrete DBN. We note that the idea of state-discretization has already been applied to handle probabilistic graphical models with continuous states [102]. However, in the present work this idea is used for the translation of an inverse PDE problem, where in particular the discretization has to be performed only locally and just once.

To account for the mean of the stochastic boundary condition and the measurement data we extend I to the left and right by the nodes m and d, respectively. The mean is supposed to be constant over time with probability 1, i.e., the according node is *persistent*.

The mean is the time-independent model parameter that we want to infer, in contrast to the time-dependent model variables $V_1, ..., V_m$. Thus we introduce an inter-time-slice edge between the nodes m of subsequent time slices. An intra-time-slice edge is introduced between the node m of the mean and the node 1 of the stochastic boundary condition. The according conditional probability distribution of the boundary variable is fully determined by the mean and the given form. In addition, an intra-time-slice edge is introduced between the last node of I and the node d of the measurement data. We can potentially introduce a measurement model, e.g. incorporate measurement noise, this way. The whole

construction is shown in Fig. 6.2. We note that the structure of the graph is simplified a lot if we want to infer the mean from a steady state. The reason is that in this case nodes from different time slices can be identified and the time axis vanishes.

The BK algorithm has originally been introduced to approximate a system of weakly interacting subsystems. Our particular DBN shows such a substructure: there are no intra-time-slice edges, and the different nodes only influence each other indirectly via entanglement. If the variables at neighboring sites get strongly entangled over time, we will expect the approximation to get better with increasing overlap of the clusters. However, direct interaction between variables at neighboring sites via intra-time-slice edges imposes a stronger correlation than entanglement. If the PDE is an advection-reaction equation and if we can apply a one-sided stencil, we can even use such direct correlations naturally by considering the DBN along characteristic lines instead of along the standard time slices. We take the perspective of the advected species and define their moving system as a time slice; see Fig. 6.2. This way we naturally reinterpret some of the inter- as intra-time-slice edges. The measurement data then directly influence the belief about the mean. A similar approach would be to introduce bigger patterns in the CPA construction that also introduce intra-time-slice edges in the original time slices. Yet another alternative could be to use smoothing over enough time-slices instead of filtering such that direct influence can be established via inter-time-slice-edges.

It turns out in practice that with the straightforward time slices measurement results at site d almost only influence the neighboring variable. This happens for the fully factorized version of the BK algorithm as well as for larger clusters. The coupling via entanglement seems to be too weak to really transmit information in our setup. Smoothing and CPA with larger patterns are computationally quite expensive, and so we choose the approach with characteristic lines together with the fully-factorized BK algorithm. We note that with the standard BK algorithm usually one does not allow for a factorization, where not each intra-time-slice-edge is contained in some cluster. However, the full factorization does not prohibit the transmission of information here, as we condition on the measurement results before the projection on the factorized densities, when the sites are still coupled.

Our approach faces practical difficulties due to the size of the problem. They can be encountered by three requirements for an ideal filtering algorithm:

i) *locality* The size of every time slice grows linearly in the space dimension, which leads to an exponential scaling for the global probability distribution. Using locality ideas for factorization helps to overcome this problem. This requirement is perfectly met by our choice of the BK algorithm which involves factorization like in CPA. In fact, the fully factorized probability distribution scales linearly in the space dimension.

ii) *working online* The DBN that we are dealing with will also grow linearly in time, if we unroll it. Asking for the inference algorithm to work online means an implementation that only stores and processes information about the present and the most recent time slice instead of all information from the unrolled network.

iii) *sparsity* The number of discrete states at every node grows exponentially in the dimension of the state space Ω. The number is very high compared to usual applications of probabilistic graphical models for example in cellular networks [56]. However, most (local) transitions between domains of state space are not possible and hence happen with probability 0. So we have the possibility to only operate on the much smaller subset of actually occupied states, although the number of

states at each node scales exponentially bad. Numerical experiments in the context of uncertainty propagation with CPA indicate that this idea yields efficient algorithms; see Sec. 5.1.3. Also in the DBN community such approaches have been recognized [17, 102].

6.3 Application: Inference of Arsenate Source

We test our method at the arsenate adsorption problem in one pipe of Sec. 5.2.1; see Fig. 5.1a. The basis is the advection-reaction problem Eq. 5.1 with the parameters from Eq. 5.2 on the approximately positively invariant Ω given by $D \in [0,1]\frac{\text{mg}}{l}$ and $A \in [0,100]\frac{\text{mg}}{\text{m}^2}$. We use a finite difference method with $\Delta x = 100\text{m}$ and $\Delta t = 10\text{min}$ and partition the phase space equidistantly in 5×15 symbols, 5 in D- and 15 in A-direction. The domains are labeled by two indices (i,j) as in Fig. 5.1b, where $i \in \{0,...,4\}$ describes the domain in dissolved and $j \in \{0,...,14\}$ the domain in adsorbed direction. The local transition rule is constructed with the backward difference quotient with stencil $U = \{-1,0\}$ in the fully factorized version, i.e., patterns on $V = \tilde{V} = \{0\}$. We use the Trotter formula and the method of characteristics to map test points via intermediate steps with the smaller $\Delta x' = 1\text{m}$ and $\Delta t' = 0.1\text{min}$. The resulting transition probabilities undergo the postprocessing proposed in Sec. 5.3.

We consider a pipe consisting of 6 report locations, where node 1 is the stochastic boundary condition and node 6 the consumer's site; see Fig. 6.2. In the following we examine two test cases, which differ in the prior information and measurement data. In the first case we do not know anyting about the pipe's state and assume a uniform distribution over all possible states at all sites. The measurement data then come from a steady state, when the adsorbed arsenate is in equilibrium with the advecting arsenate. In the second case we take an empty pipe as a prior and assume that the contaminant is only in the boundary node. The measurement data comes from an according transient state when the arsenate starts advecting through the pipe. In both cases the goal is to infer the mean, about which we do not assume to know anything.

Let us first describe the (conditional) probability distributions that are assigned to all nodes of the DBN. We only have to specify them for the first characteristic line and for the transition model between subsequent characteristic lines. Together they define the whole DBN. This specification includes the priors, and we begin with the first test case. Note that in the first characteristic line only the mean node can be given a free initial distribution. All other nodes in this line have parents in the same characteristic line, and so their distributions are conditional distributions.

i) The mean node is only considered to describe the mean of the dissolved species in this application. As a prior we choose a uniform distribution over all 5 discrete states in the first time slice.

ii) The distribution of the stochastic boundary condition at node 1 is conditioned on the mean: If the mean is in state 0, we assume that there is no arsenate at all and hence the boundary condition is in state $(0,0)$ with probability 1. If the mean is in state $e \in \{1,2,3\}$, then the dissolved part of the boundary condition is with probability $\frac{1}{4}$ in the state $e-1$ and $e+1$, and with probability $\frac{1}{2}$ in the dissovled state e. For dissolved state 4 of the mean, i.e., the boundary domain of phase space, the dissolved

part of the boundary condition is with probability $\frac{1}{4}$ in state 3 and with probability $\frac{3}{4}$ in state 4. The adsorbed state of the boundary condition is always with probability $\frac{1}{3}$ in state $12, 13$ or 14, and the complete state in the 2-dimensional phase space is then given by all combinations of dissolved and adsorbed states. For example, if the mean is in state 2, the boundary condition will be in state $(2, 14)$ with probability $\frac{1}{6}$.

iii) For nodes $2, ..., 6$ we choose the non-informative uniform distribution over all states in the first characteristic line, i.e., for given left neighbor each state at the node under consideration has the same probability.

iv) We assume that just the state-discrete dissovled state is measured at node d. The conditional distribution is given by simply extracting the domain of the dissolved arsenate from node 6. So we use the simplest measurement model without adding any additional measurement noise.

v) The conditional distributions over m of following time slices are defined by copying the state from one time step to the next.

vi) The local transition rule defines the conditional probability distribution associated with the nodes $2, ..., 6$ in all time slices except for the first one. Note that we have to reinterpret the local transition rule as to connect two nodes in the image with one in the preimage characteristic line rather than one in the image with two in the preimage standard time slice.

In the dynamic case the transition model is the same, but we have to change the prior information for nodes $2, ..., 6$ in iii). Node 2 is the first site of an empty pipe in the first characteristic line, regardless of the state of the boundary node. So it needs to be decoupled from the boundary in the beginning. The according random variable is in state 0 with probability 1 for all states of the boundary node. Nodes $3, ..., 6$ are only connected to their respective left neighbors by an edge in the same characteristic line. They are assigned trivial rules to model the empty pipe: we copy simply the left neighbor's state. Because the state at node 2 is 0, then the whole pipe is empty.

Let us now specify the measurement data. In both cases we calculate the distribution of the measurements at node d with a forward simulation using the according CPA and pick a state with a random number generator at each time step. We set the mean to state 2 in the forward simulation in both cases and want to infer it again from the generated data. Note that the measurement distribution is time-independent in the steady state, but time-dependent in the transient state. The distributions can be found in Fig. 6.3 a,b), and the chosen states in Fig. 6.3 c,d). In the steady state we can start with the filtering process already from the beginning. The steady state is time-invariant, and all characteristic lines degenerate to one time slice. In the transient state, however, the first contamination information arrives at the measurement node at the 6. time step. Hence we include the evidence in the resulting belief state not before this time step, although they are in the same characteristic line.

Of the many existing toolboxes for inference in DBNs we choose the BNT software [132], because it offers immediate access to an implementation of the BK algorithm. Although it only fulfills the locality property from our list of requirements for an ideal inference algorithm, it already performs quite well. With a version of the BK algorithm that satisfies the online and sparsity requirement in addition, there will be hope for a very efficient algorithm.

Fig. 6.3 e,f) shows the resulting belief states about the mean for both cases. One can

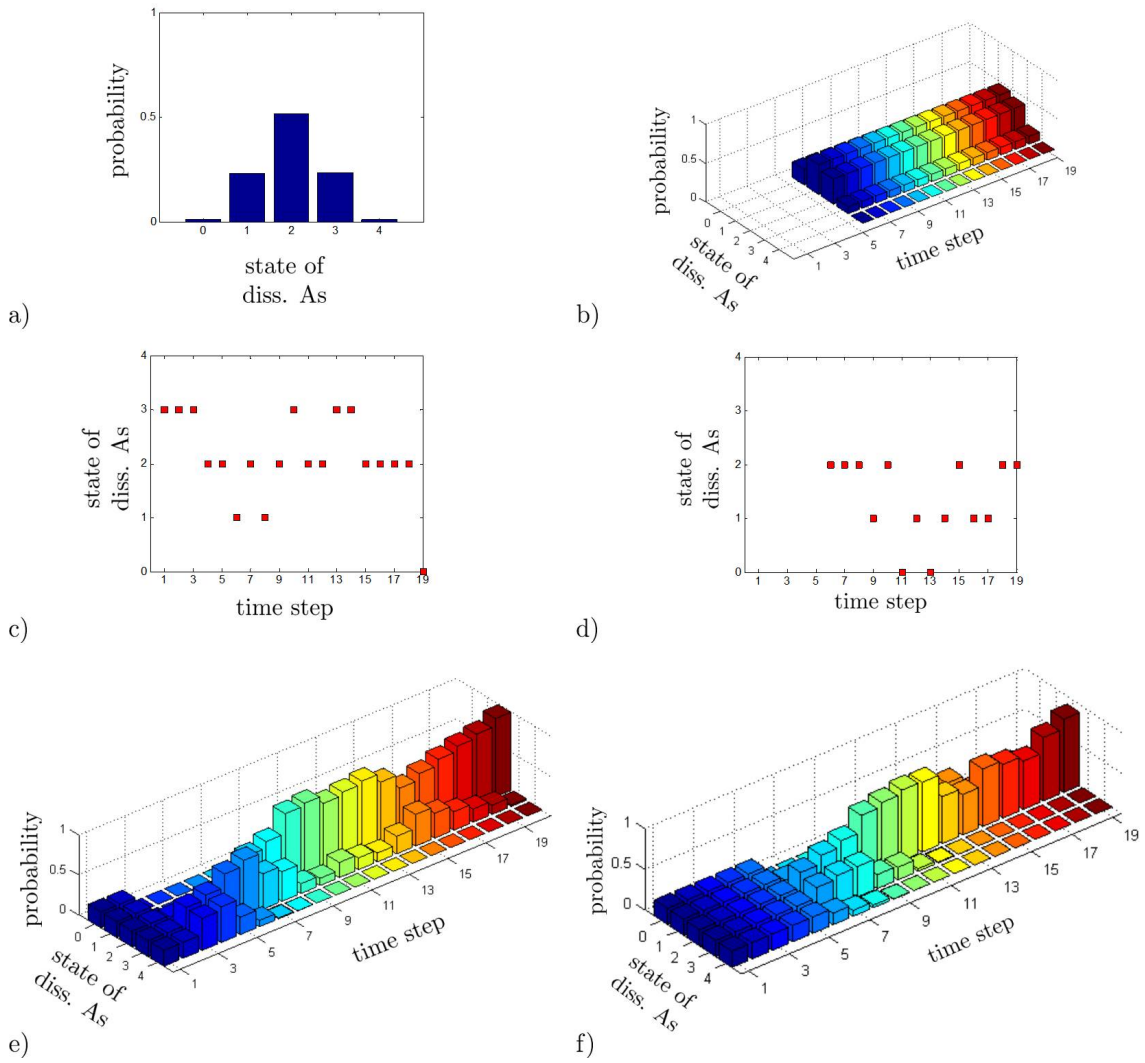


Figure 6.3: Filtering for the application of arsenate advection and adsorption in a pipe. a), c) and e) are for filtering in the steady, b), d) and f) in the transient state. a) and b) show the distribution of measurements at node d from a forward simulation with CPA, from which the data in c) and d) is drawn, respectively. e) and f) show the belief state for the mean at node m.

clearly see that already after a few measurements the mean is with very high probability in state 2 in both the steady and the dynamic case. This result matches the information that we put in the forward model when we generated the measurement data. Moreover, we can even observe, for example from Fig. 6.3 e), how the algorithm first assumes that the mean is most probably in state 3, because by chance the measurements in the first three time steps are in state 3. However, when additional measurement information is included, state 2 from time steps 4 and 5, and even more state 1 from step 6, the algorithm states that the mean is more probably in state 2. Another observation that we want to stress can be made in Fig. 6.3 f): although the measurements are all in state $0, 1$ or 2, the filtering result tells us that the mean is most probably in state 2. The reason is that in this dynamic case a lot of the inflowing arsenate is adsorbed at the empty pipe wall first, and that hence only

a lower concentration of dissolved arsenate arrives at the consumer. Also it can be seen in this figure how the algorithm reacts to the unlikely effect that two almost consecutive measurements are in state 0: The probability of the mean being in lower states rises for a while.

6.4 Conclusion and Outlook

In this chapter we have introduced a novel numerical approach to inverse problems in PDEs. We suggest to use the discretization of cellular probabilistic automata (CPA) to translate PDEs into dynamic Bayesian networks (DBNs), and then apply a graph-based inference algorithm. In the past the field of state-discrete probabilistic graphical models has been quite independent from the field of inverse problems in PDEs. With this work we aim to provide a connection between the two fields. In the context of the thesis this chapter shows how measurement data can be integrated in our completely discrete and probabilistic framework for modeling and simulation.

In particular, in the CPA discretization process first a finite difference scheme transfers the PDE into a time- and space-discrete dynamical system. Then the state space is discretized with tools of set oriented numerics, and the neighborhood concept of cellular automata theory is used to exploit the locality and shift-invariance of the interaction. A pattern parameter can be used to control the degree of locality in the CPA, but we only apply minimal patterns in this work. The resulting object is interpreted as a DBN.

Inference in DBNs or, more generally, in probabilistic graphical models has been extensively studied in the literature. We want to solve an inverse problem of parameter estimation and recast it as a filtering problem, for which a vast array of sophisticated algorithms exists. In principle many of them can be applied to our DBN that is constructed from a PDE. We have identified criteria with which algorithms and implementations can be evaluated with respect to our setup; our choice is the BK algorithm.

Our approach has been demonstrated at the example of arsenate transportation and adsorption in a drinking water pipe with a contamination source at the entrance. The mean of the arsenate concentration in the upstream boundary condition has been inferred by measuring the downstream boundary conditions.

It is desirable to extend the application of drinking water contamination to larger grids in order to handle real world problems. Then one is interested not only in the strength of a contamination source, but also in its location [39, 167]. For this purpose we want to combine the inference tools presented here with the ideas of Sec. 5.4 on how to connect pipes to a whole grid. From a theoretical point of view it will also be interesting to investigate further how larger patterns in CPA can be used to generalize our approach to more complex PDEs. An application could be leak detection in water or gas pipes [60, 148]. The experience with the Euler equation from Sec. 5.5 can help to tackle such fluid dynamical problems.

7 Conclusion and Outlook

We have introduced a modeling and simulation framework for technical spatio-temporal systems. It can be applied to existing models consisting of partial differential and algebraic coupling equations in a preprocessing and a simulation step. In the preprocessing a system model is translated into a cellular probabilistic automaton (CPA), and in the second step this object is used to propagate a probabilistic belief state about the system's state. If there are no measurement data, we simply evolve a probability distribution over an initial state with a potentially stochastic model. If data are available at some points in time, they can be integrated by Bayesian updates with one of the tools from the large pool of inference algorithms in dynamic Bayesian networks.

In particular, the framework has been developed in four steps in this thesis. First, we have introduced the precursor of CPA, cellular non-deterministic automata, in the context of cellular automata theory to analyze partial differential equations (PDEs) in a completely discrete setup. The insights have been used in the second step, the theoretical introduction of CPA for uncertainty propagation in PDEs. In the third step we have extended the CPA method to systems of coupled PDEs and have applied it to the simulation of contaminant fate in water grids. Furthermore, we have suggested a more sophisticated version to better deal with conservation laws and have tested it at the contaminant fate and the isothermal Euler equation of gas dynamics. In the last step, we have shown how measurement data can be integrated in CPA in a Bayesian setup.

The framework is designed to meet two common practical demands. As a probabilistic method, it allows to take into account uncertainties about model features or parameters. This leads to results that truly reflect the knowledge we have about reality. Furthermore, CPA compute on discrete states instead of on the continuous phase space of PDEs, and the discrete simulation step is fast compared to the preprocessing. The preprocessing takes approximately as long as a Monte Carlo computation, but has to be conducted only once. Since the discrete states can be interpreted as risk levels, fast calculations directly on this simplified state space suit industrial needs by avoiding an expensive numerical overhead.

The approach connects different research fields that have been only in loose contact before. For cellular automata theory the approach is interesting, because it provides a novel link to traditional PDE models and explores the potential of calculations with several trajectories at the same time. Similarly, for the field of probabilistic graphical models the method provides access to PDEs and cellular automata theory. Moreover, with the CPA approach, set-oriented numerics can broaden its range of applications to PDEs, while it has been focused mainly on systems of ordinary differential equations before. In the context of uncertainty quantification the approach might be interesting as a novel numerical tool for uncertainty propagation which is in principle capable of handling white noise. It also promises to overcome the curse of dimension in the random parameters in a sense that they only affect the preprocessing, but not the simulation. As it is scalable to large grids, it might be of specific interest in industrial numerical analysis. Finally, the method provides

access to a new and vast pool of algorithms and software for people who are interested in inverse problems and Bayesian inference for PDEs.

Although we have achieved promising results in our exemplary application, there are still some difficulties and open questions. We suggest two main directions for future research to meet these challenges. First, we are interested in really quantifying and improving the approximation error beyond our consistency result for CPA. This seems to be very demanding, since spatial correlations have to be quantified somehow. To increase precision we suggest to pursue the idea of non-rectangular state-space domains like introduced for CNDA in Sec. 3.3. Second, it would be interesting to think about ways for speeding up the CPA algorithm. One way is to adapt sophisticated ideas in the preprocessing that have been used in set-oriented numerics [73]. Although CPA have nice scaling properties in time, space and the dimension of the random space, they encounter complexity difficulties in the dimension of the phase space. Another way is therefore to formulate a surrogate model to work on a low-dimensional space [145].

We think that a major goal for modeling and simulation technolgoy in the future will be smoother integration in other disciplines and better access to non-experts. Our idea to directly work on a state space with practical interpretation is just one approach. Other approaches well-suited for product simulation are, for example, isogeometric analysis [30,86] which promises a seamless transition between computer-aided design (CAD) and simulation, or, more general, product lifecycle management software [146]. The latter aims at integrating simulation into a software for the whole product lifecycle from design over production to disposal. In the long run, however, the biggest challenge for modeling and simulation might be the completion of today's electronic hardware by conceptually very different but probably much more powerful devices. Quantum information [138] or, more specifically, quantum simulation [29,51] which provides a means to simulate complex processes naturally by other complex processes could have the potential to significantly increase the potential of simulation in practice.

Bibliography

[1] F. Augustin, A. Gilg, M. Paffrath, P. Rentrop, M. Villegas, and U. Wever. An accuracy comparison of polynomial chaos type methods for the propagation of uncertainties. *Journal of Mathematics in Industry*, 3, 2013.

[2] F. Augustin, A. Gilg, M. Paffrath, P. Rentrop, and U. Wever. A Survey in Mathematics for Industry - Polynomial chaos for the approximation of uncertainties: Chances and limits. *European Journal of Applied Mathematics*, 19:149–190, 2008.

[3] F. Augustin and P. Rentrop. Stochastic Galerkin techniques for random ordinary differential equations. *Numerische Mathematik*, 122:399–419, 2012.

[4] S. K. Bachman and M. Goodreau. Steady State - Is The Solution Realistic For The Piping Network?, PSIG report 0309. Pipeline Simulation Interest Group, 2003.

[5] A. Banaszuk, V. A. Fonoberov, T. A. Frewen, M. Kobilarov, G. Mathew, I. Mezić, A. Pinto, T. Sahai, H. Sane, A. Speranzon, and A. Surana. An accuracy comparison of polynomial chaos type methods for the propagation of uncertainties. *Annual Reviews in Control*, 35:77–98, 2011.

[6] M. K. Banda, M. Herty, and A. Klar. Gas Flow in Pipeline Networks. *Networks and Heterogeneous Media*, 1:41–56, 2006.

[7] O. Bandman. Simulating Spatial Dynamics by Probabilistic Cellular Automata. In S. Bandini, B. Chopard, and M. Tomassini, editors, *Cellular Automata*, volume 2493 of *Lecture Notes in Computer Science*, pages 10–19. Springer, 2002.

[8] A. Barth, C. Schwab, and N. Zollinger. Multi-level Monte Carlo Finite Element method for elliptic PDEs with stochastic coefficients. *Numerische Mathematik*, 119:123–161, 2011.

[9] R. G. Bartle. *The Elements of Integration and Lebesgue Measure*. Wiley, 1995.

[10] M. Bartlett. Contingency Table Interactions. *Journal of the Royal Statistical Society, Series B*, 2:248–252, 1935.

[11] M. Batty. *Cities and Complexity*. MIT Press, 2005.

[12] H. Beigi. *Fundamentals of Speaker Recognition*. Springer, 2011.

[13] J. Bernardo and A. Smith. *Bayesian theory*. Wiley, 2000.

[14] A. S. Besicovitch. *Almost Periodic Functions*. Cambridge University Press, 1954.

[15] L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, Y. M. Marzouk, L. Tenorio, B. van Bloemen Waanders, and K. Willcox, editors. *Large-Scale Inverse Problems and Quantification of Uncertainty*. Wiley, 2011.

[16] J. Bilmes. *The Graphical Models Toolkit (GMTK) [computer software]*. http://melodi.ee.washington.edu/ bilmes/gmtk/, 2009.

[17] J. Bilmes. *GMTK: The Graphical Models Toolkit, documentation*. http://melodi.ee.washington.edu/ bilmes/gmtk/, 2010.

[18] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[19] C. Bourdarias, S. Gerbi, and M. Gisclon. A kinetic formulation for a model coupling free surface and pressurised flows in closed pipes. *Journal of Computational and Applied Mathematics*, 218:522–531, 2008.

[20] X. Boyen and D. Koller. Approximate Learning of Dynamic Models. In *Proceedings of the 11th Conference on Advances in Neural Information Processing Systems*. MIT Press, 1998.

[21] X. Boyen and D. Koller. Tractable Inference for Complex Stochastic Processes. In *Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann, 1998.

[22] X. Boyen and D. Koller. Exploiting the Architecture of Dynamic Systems. In *Proceedings of the 16th National Conference on Artificial Intelligence*. American Association for Artificial Intelligence Press, 1999.

[23] J. Brouwer, I. Gasser, and M. Herty. Gas Pipeline Models Revisited: Model Hierarchies, Nonisothermal Models, and Simulations of Networks. *Multiscale Modeling and Simulation*, 9:601–623, 2011.

[24] Bundesministerium für Bildung und Forschung (BMBF). *Bericht der Bundesregierung: Zukunftsprojekte der Hightech-Strategie (HTS-Aktionsplan)*. Bundesministerium für Bildung und Forschung (BMBF), 2012.

[25] P. Campbell (editor). Special Issue: Big data - science in the petabyte era. *Nature*, 455:1–136, 2008.

[26] T. Ceccherini-Silberstein and M. Coornaert. A generalization of the Curtis-Hedlund theorem. *Theoretical Computer Science*, 400:225–229, 2008.

[27] T. Ceccherini-Silberstein and M. Coornaert. *Cellular Automata and Groups*. Springer, 2010.

[28] S. Chib and E. Greenberg. Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, 49:327–335, 1995.

[29] J. I. Cirac and P. Zoller. Goals and opportunities in quantum simulation. *Nature Physics*, 8:264–266, 2012.

[30] J. A. Cottrell, T. J. R. Hughes, and Y. Bazilevs. *Isogeometric Analysis: Toward Integration of CAD and FEA*. Wiley, 2009.

[31] M. Davidich and S. Bornholdt. The transition from differential equations to Boolean networks: A case study in simplifying a regulatory network model. *Journal of Theoretical Biololgy*, 255:269–277, 2008.

[32] G. de Vries, T. Hillen, M. Lewis, J. Müller, and B. Schönfisch. *A course in Mathematical Biology: Quantitative Modeling with Mathematical and Computational Methods*. Society for Industrial and Applied Mathematics, 2006.

[33] T. Dean and K. Kanazawa. A model for reasoning about persistence and causation. *Computational Intelligence*, 5:142–150, 1989.

[34] M. Dellnitz and O. Junge. On the Approximation of Complicated Dynamical Behavior. *SIAM Journal on Numerical Analysis*, 36:491–515, 1999.

[35] M. Dellnitz and O. Junge. Set oriented numerical methods for dynamical systems. In B. Fiedler, editor, *Handbook of Dynamical Systems*, volume 2, pages 221–264. Elsevier Science, 2002.

[36] E. V. Denardo. *Dynamic Programming: Models and Applications*. Dover, 2003.

[37] P. Deuflhard and F. Bornemann. *Numerische Mathematik II - Gewöhnliche Differentialgleichungen*. Cambridge University Press, 2002.

[38] A. Deutsch and S. Dormann. *Cellular automaton modeling of biological pattern formation*. Birkhäuser, 2005.

[39] C. Di Christo and A. Leopardi. Uncertainty Effects on Pollution Source Location in Water Networks. In *Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium*. American Society of Civil Engineers, 2008.

[40] F. A. DiGiano and W. Zhang. Uncertainty Analysis in a Mechanistic Model of Bacterial Regrowth in Distribution Systems. *Environmental Science & Technology*, 38:5925–5931, 2004.

[41] A. Doostan and G. Iaccarino. A least-squares approximation of partial differential equations with high-dimensional random inputs. *Journal of Computational Physics*, 228:4332–4345, 2009.

[42] A. Doucet, N. De Freitas, and N. Gordon, editors. *Sequential Monte Carlo methods in practice*. Springer, 2001.

[43] M. Dunham and K. P. Murphy et al. *Probabilistic Modeling Toolkit for Matlab/Octave (version 3) (PMTK3) [computer software]*. http://code.google.com/p/pmtk3/, 2012.

[44] R. Durrett. Ten lectures on particle systems. In P. Bernard, editor, *Lectures on Probability Theory*, volume 1608 of *Lecture Notes in Mathematics*, pages 97–201. Springer, 1995.

[45] P. Dutta, A. Halder, and R. Bhattacharya. Uncertainty Quantification for Stochastic Nonlinear Systems using Perron-Frobenius Operator and Karhunen-Loève Expansion. In *Proceedings of CCA - IEEE International Conference on Control and Applications*. Institute of Electrical and Electronics Engineers, 2012.

[46] T. A. El Moselhy and Y. M. Marzouk. Bayesian inference with optimal maps. *Journal of Computational Physics*, 231:7815–7850, 2012.

[47] M. Evans and T. Swartz. Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science*, 10:254–272, 1995.

[48] P. C. Evans and M. Annunziata. *Industrial Internet: Pushing the Boundaries of Minds and Machines*. General Electric, 2012.

[49] S. N. Evans and P. B. Stark. Inverse problems as statistics. *Inverse Problems*, 18:R55–R97, 2002.

[50] H. G. Evertz. *Computer Simulations, lecture notes*. Technical University Graz, 2009.

[51] R. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21:467–488, 1982.

[52] G. S. Fishman. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer, 1996.

[53] J. Foo, X. Wan, and G. E. Karniadakis. The multi-element probabilistic collocation method (ME-PCM): Error analysis and applications. *Journal of Computational Physics*, 227:9572–9595, 2008.

[54] B. L. Fox. *Strategies for Quasi-Monte Carlo*. Kluwer Academic, 1999.

[55] J. Franke. *Untersuchung der hydrodynamischen Stabilität von Verdampferheizflächen mit Hilfe nichtlinearer Simulation, PhD thesis*. Technical University of Darmstadt, 1986.

[56] N. Friedman. Inferring Cellular Networks Using Probabilistic Graphical Models. *Science*, 303:799–805, 2004.

[57] B. E. Fristedt and L. F. Gray. *A Modern Approach to Probability Theory*. Birkhäuser, 1996.

[58] M. Gardner. Mathematical Games - The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 223:120–123, 1970.

[59] R. J. Gaylord and L. J. D'Andria. *Simulating Society*. Springer, 1998.

[60] G. Geiger. *Principles of Leak Detection*. KROHNE Oil and Gas Brochure, 2008.

[61] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, 2004.

[62] R. Ghanem and P. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer, 1991.

[63] A. L. Gibbs and F. E. Su. On choosing and bounding probability metrics. *International Statistical Review*, 70:419–435, 2002.

[64] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov Chain Monte Carlo in practice*. Chapman and Hall/CRC, 1995.

[65] M. Golubitsky and I. Stewart. Nonlinear dynamics of networks: the groupoid formalism. *Bulletin of the American Mathematical Society*, 43:305–364, 2008.

[66] J. M. Greenberg and S. P. Hastings. Spatial Patterns for Discrete Models of Diffusion in Excitable Media. *SIAM Journal of Applied Mathematics*, 54:515–523, 1978.

[67] M. S. Grewal and A. P. Andrews. *Kalman Filtering: Theory and Practice Using MATLAB*. Wiley, 2008.

[68] Peter Grindrod. *Patterns and Waves: Theory and Applications of Reaction-Diffusion Equations*. Clarendon Press, 1991.

[69] T. Gross and B. Blasus. Adaptive Coevolutionary Networks: A Review. *Journal of the Royal Society – Interface*, 5:259–271, 2008.

[70] Pipeline Simulation Interest Group. www.psig.org.

[71] L. Grüne and O. Junge. Global optimal control of perturbed systems. *Journal of Optimization Theory and Applications*, 136:411–429, 2008.

[72] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, 1983.

[73] R. Guder, M. Dellnitz, and E. Kreuzer. An adaptive method for the approximation of the generalized cell mapping. *Chaos, Solitons & Fractals*, 8:525–534, 1997.

[74] W. Hackbusch. *Elliptic Differential Equations: Theory and Numerical Treatment*. Springer, 1992.

[75] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52, 1902.

[76] F. v. Haeseler, H.-O. Peitgen, and G. Skordev. Cellular automata, matrix substitutions and fractals. *Annals of Mathematics and Artificial Intelligence*, 8:345–362, 1993.

[77] H. Hanche-Olsen and H. Holden. The Kolmogorov-Riesz compactness theorem. *Expositiones Mathematicae*, 28:385–394, 2010.

[78] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.

[79] G. A. Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Theory of Computing Systems*, 3:320–375, 1969.

[80] Thomas Hillen. A Turing model with correlated random walk. *Journal of Mathematical Biology*, 35:49–72, 1996.

[81] C. Hirsch. *Numerical Computation of Internal and External Flows*. Wiley, 1990.

[82] J. E. Hopcroft, R. Motwani, and J. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison Wesley, 2000.

[83] H.-B. Horlacher and H.-J. Lüdecke. *Strömungsberechnung für Rohrsysteme*. ExpertVerlag, 2012.

[84] C. S. Hsu. *Cell-to-Cell Mapping: A Method of Global Analysis for Nonlinear Systems*. Springer, 1987.

[85] X. Huan and Y. M. Marzouk. Simulation-Based Optimal Bayesian Experimental Design for Nonlinear Systems. *Journal of Computational Physics*, 232:288–317, 2013.

[86] T. J. R. Hughes, J. A. Cottrell, and Y. Bazilevs. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer methods in applied mechanics and engineering*, 194:4135–4195, 2004.

[87] F. Y. Hunt. A Monte Carlo approach to the approximation of invariant measures. *Random & Computational Dynamics*, 2:111–133, 1994.

[88] Intel Corporation. *Probabilistic Network Library (PNL) [computer software]*. http://sourceforge.net/projects/openpnl/, 2003.

[89] F. James. Monte Carlo theory and practice. *Princeton University Bulletin*, 43:1145–1189, 1980.

[90] O. Junge, J. E. Marsden, and I. Mezić. Uncertainty in the dynamics of conservative maps. In *Proceedings of the 43rd IEEE Conference on Decision and Control*. Institute of Electrical and Electronics Engineers, 2004.

[91] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. Springer, 2005.

[92] J. Kari. Theory of cellular automata: A survey. *Theoretical Computer Science*, 334:3–33, 2005.

[93] G. E. Karniadakis, C. H. Su, D. Xiu, D. Lucor, C. Schwab, and R. A. Todor. Generalized polynomial chaos solution for differential equations with random inputs. *ETH Zürich, Research Report*, 2005-01, 2005.

[94] S. A. Kauffman. *The Origins of Order*. Oxford University Press, 1993.

[95] A. S. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.

[96] J. M. Keeling. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London, Series B*, 266:859–867, 1999.

[97] M. C. Kennedy and A. O'Hagan. Using Bayesian statistics in the estimation of heat source in radiation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63:425–464, 2001.

[98] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems*. Springer, 1996.

[99] A. Klenke. *Probability Theory*. Springer, 2007.

[100] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 2011.

[101] S. Klostermann, R. Murray, J. Szabo, J. Hall, and J. Uber. Modeling and simulation of arsenate fate and transport in a distribution system simulator. In *Proceedings of Water Distribution System Analysis*. American Society of Civil Engineers, 2010.

[102] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, 2009.

[103] L. K. Koopal and M. J. Avena. A simple model for adsorption kinetics at charged solid-liquid interfaces. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 192:93–107, 2001.

[104] H. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, 2001.

[105] H. P. Langtangen. Numerical solution of first passage problems in random vibrations. *SIAM Journal on Scientific Computing*, 15:977–996, 1994.

[106] A. Lasota and M. C. Mackey. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. Springer, 1993.

[107] S. Lauritzen. *Graphical Models*. Oxford University Press, 1996.

[108] C. L. Lawson and R. J. Hanson. *Solving least squares problems*. Prentice-Hall, 1974.

[109] R. J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser, 1999.

[110] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2004.

[111] T. Y. Li. Finite Approximation for the Frobenius-Perron Operator. A Solution to Ulam's Conjecture. *Journal of Approximation Theory*, 17:177–186, 1976.

[112] C. Lieberman, K. Willcox, and O. Ghattas. Parameter and State Model Reduction for Large-Scale Statistical Inverse Problems. *SIAM Journal on Scientific Computing*, 32:2523–2542, 2010.

[113] T. M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*. Springer, 1999.

[114] W. L. Loh. On Latin Hypercube Sampling. *Annals of Statistics*, 24:2058–2080, 1996.

[115] A. Machì and F. Mignosi. Garden of Eden Configurations for Cellular Automata on Cayley Graphs of Groups. *SIAM Journal on Discrete Mathematics*, 6:44–56, 1993.

[116] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Hung Byers. *Big data: The next frontier for innovation, competition, and productivity*. McKinsey & Company, 2011.

[117] J. Manyika, M. Chui, J. Bughin, R. Dobbs, P. Bisson, and A. Marrs. *Disruptive technologies: Advances that will transform life, business, and the global economy*. McKinsey & Company, 2013.

[118] Y. M. Marzouk and H. N. Najm. Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *Journal of Computational Physics*, 228:1862–1902, 2009.

[119] Y. M. Marzouk, H. N. Najm, and L. A. Rahn. Stochastic spectral methods for efficient Bayesian solution of inverse problems. *Journal of Computational Physics*, 224:560–586, 2006.

[120] E. S. Menon. *Pipeline Planning and Construction Field Manual*. Gulf Professional Publishing, 2011.

[121] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21:1087–1092, 1953.

[122] I. Mezić and T. Runolfsson. Uncertainty propagation in dynamical systems. *Automatica*, 44:3003–3013, 2008.

[123] Microsoft Corporation. *Infer.NET [computer software]*. http://research.microsoft.com/en-us/um/cambridge/projects/infernet/, 2013.

[124] A. Mielke and G. Schneider. Attractors for modulation equations on unbounded domains: existence and comparison. *Nonlinearity*, 8:1–26, 1995.

[125] A. Mohammad-Djafari. Bayesian inference for inverse problems. In *Proceedings of the 21st International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. American Institute of Physics, 2002.

[126] R. E. Moore, R. B. Kearfott, and M. J. Cloud. *Introduction to Interval Analysis*. Society for Industrial and Applied Mathematics, 2009.

[127] A. E. Motter and Y.-C. Lai. Cascade-based attacks on complex networks. *Physical Review E*, 66:065102 1–4, 2002.

[128] J. Müller and C. Spandl. A Curtis-Hedlund-Lyndon theorem for Besicovitch and Weyl spaces. *Theoretical Computer Science*, 410:3606–3615, 2009.

[129] J. Müller and C. Spandl. Embeddings of dynamical systems into cellular automata. *Ergodic Theory and Dynamical Systems*, 29:165–177, 2009.

[130] J. Müller and C. Spandl. Embeddings of dynamical systems into cellular automata - ERRATUM. *Ergodic Theory and Dynamical Systems*, 30:1271, 2010.

[131] K. P. Murphy. *Dynamic Bayesian Networks: Representation, Inference and Learning, Phd thesis*. University of California, Berkeley, 2002.

[132] K. P. Murphy. *Bayes Net Toolbox for Matlab (BNT) [computer software]*. https://code.google.com/p/bnt/, 2007.

[133] K. P. Murphy. Software for Graphical Models: a Review. *Newsletter of the International Society for Bayesian Analysis*, 14:13–15, 2007.

[134] K. P. Murphy. *Software Packages for Graphical Models*. http://www.cs.ubc.ca/~murphyk/Software/bnsoft.html, 2013.

[135] K. P. Murphy and Y. Weiss. The Factored Frontier Algorithm for Approximate Inference in DBNs. In *Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence*. Morgan Kaufmann, 2001.

[136] National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science. *Revolutionizing Engineering Science Through Simulation: A Report of the National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science*. National Science Foundation (USA), 2006.

[137] M. E. J. Newman. The Structure and Function of Complex Networks. *SIAM Review*, 45:167–256, 2003.

[138] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[139] K. Nishinari and D. Takahashi. Analytical properties of ultradiscrete Burgers equation and rule-184 cellular automaton. *Journal of Physics A: Mathematical and General*, 31:5439–5450, 1998.

[140] B. K. Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2002.

[141] A. J. Osiadacz. Different Transient Flow Models - Limitations, Advantages, And Disadvantages, PSIG report 9606. Pipeline Simulation Interest Group, 1996.

[142] M. L. Pierce and C. B. Moore. Adsorption of arsenite and arsenate on amorphous iron hydroxide. *Water Research*, 16:1247–1253, 1982.

[143] A. Reiss. *Zelluläre probabilistische Automaten für die Propagation von Unsicherheiten in einem Biofilm-Modell, Bachelor's thesis*. Technical University Munich, 2013.

[144] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2004.

[145] N. Roy, G. Gordon, and S. Thrun. Finding Approximate POMDP Solutions Through Belief Compression. *Journal of Artificial Intelligence Research*, 23:1–40, 2005.

[146] A. Saaksvuori and A. Immonen. *Product Lifecycle Management*. Springer, 2009.

[147] W. E. Schiesser. *The Numerical Method of Lines: Integration of Partial Differential Equations*. Academic Press, 1991.

[148] B. Schöne. Simulations that Localize Leaks. *Pictures of the Future: The Magazine for Research and Innovation, Spring 2012*, pages 64–65, 2012.

[149] F. Shang, J. G. Uber, and L. A. Rossman. Modeling Reaction and Transport of Multiple Species in Water Distribution Systems. *Environmental Science & Technology*, 42:808–814, 2008.

[150] F. Shang, J. G. Uber, L. A. Rossman, R. Janke, and R. Murray. *EPANET Multi-Species Extension User's Manual*. US Environmental Protection Agency, 2011.

[151] Siemens AG. *Perspectives: Connecting Productivity and Efficiency*. Siemens AG, Industry Sector, 2012.

[152] Siemens AG. *Pictures of the Future: The Magazine for Research and Innovation, Fall 2012*. Siemens AG, Corporate Communications and Corporate Technology, 2012.

[153] Siemens AG. *Water industry - Increasing efficiency with SIWA Network Management System*. Siemens AG, Industry Sector, 2012.

[154] Siemens AG. *Pictures of the Future: The Magazine for Research and Innovation, Spring 2013*. Siemens AG, Corporate Communications and Corporate Technology, 2013.

[155] I. Simonsen, L. Buzna, K. Peters, S. Bornholdt, and D. Helbing. Transient Dynamics Increasing Network Vulnerability to Cascading Failures. *Physical Review Letters*, 100:218701 1–4, 2008.

[156] D. Sivia and J. Skilling. *Data Analysis - A Bayesian Tutorial*. Oxford University Press, 2006.

[157] J. Smoller. *Shock Waves and Reaction-Diffusion Equations*. Springer, 1994.

[158] G. Steinebach, R. Rosen, and A. Sohr. Modeling and Numerical Simulation of Pipe Flow Problems in Water Supply Systems. In A. Martin, K. Klamroth, J. Lang, G. Leugering, A. Morsi, M. Oberlack, M. Ostrowski, and R. Rosen, editors, *Mathematical Optimization of Water Networks*, volume 162 of *International Series of Numerical Mathematics*, pages 3–15. Birkhäuser, 2012.

[159] R. Sternfels and C. J. Earls. Reduced-order model tracking and interpolation to solve PDE-based Bayesian inverse problems. *Inverse Problems*, 29:075014–075022, 2013.

[160] J. J. Stoker. *Water Waves: The Mathematical Theory with Applications*. Wiley, 1992.

[161] A. Surana, T. Sahai, and A. Banaszuk. Iterative Methods for Scalable Uncertainty Quantification in Complex Networks. *International Journal for Uncertainty Quantification*, 2:413–439, 2012.

[162] K. Sutner. De Bruijn Graphs and Linear Cellular Automata. *Complex Systems*, 5:19–30, 1991.

[163] W. Thirring. *Quantum Mechanics of Atoms and Molecules*. Springer, 1981.

[164] J. W. Thomas. *Numerical Partial Differential Equations: Finite Difference Methods*. Springer, 1995.

[165] T. Toffoli. Cellular Automata as an alternative to (rather than an approximation of) differential equations in modeling physics. *Physica 10D*, pages 117–127, 1984.

[166] T. Tokihiro, D. Takahashi, J. Matsukidaira, and J. Satsuma. From Soliton Equations to Integrable Cellular Automata through a Limiting Procedure. *Physical Review Letters*, 76:3247–3250, 1996.

[167] J. Uber. Identifiability of Contaminant Source Characteristics in Steady-State and Time-Varying Network Flows. In *Impacts of Global Climate Change, Proceedings of EWRI ASCE World Water and Environmental Resources Congress*. American Society of Civil Engineers, 2005.

[168] S. M. Ulam. *A Collection of Mathematical Problems*. Interscience, 1960.

[169] UN Millenium Project. *Investing in Development: A Practical Plan to Achieve the Millenium Development Goals. Overview.* United Nations Development Programme, 2005.

[170] U.S. Environmental Protection Agency. *EPAs Water Security Modeling and Simulation Research, EPA/600/S-11/007*. U.S. Environmental Protection Agency, 2011.

[171] U.S. National Academy of Engineering of the National Academies. *Grand Challenges for Engineering*. National Academy of Sciences, 2008.

[172] C. R. Vogel. *Computational Problems for Inverse Problems*. Society for Industrial and Applied Mathematics, 1996.

[173] R. Vollmar. *Algorithmen in Zellularautomaten*. Teubner, 1979.

[174] J. von Neumann. The general and logical theory of automata. In L. A. Jeffress, editor, *Cerebral Mechanisms in Behavior: The Hixon Symposium*, pages 1–41. Wiley, 1951.

[175] J. Wan and N. Zabaras. A probabilistic graphical model approach to stochastic multiscale partial differential equations. *Journal of Computational Physics*, 250:477–510, 2013.

[176] X. Wan and G. E. Karniadakis. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM Journal on Scientific Computing*, 28:901–928, 2006.

[177] P. P. Wang and C. Zheng. Contaminant transport models under random sources. *Ground water*, 43:423–433, 2005.

[178] S. Webel. Simulation Tools on Tap. *Pictures of the Future: The Magazine for Research and Innovation, Fall 2009*, pages 55–56, 2009.

[179] N. Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60:897–936, 1938.

[180] S. Willard. *General Topology*. Dover, 2004.

[181] S. G. Williams, editor. *Symbolic Dynamics and its Applications*, volume 60 of *Proceedings of Symposia in Applied Mathematics*. American Mathematical Society, 2004.

[182] J. Wloka. *Funktionalanalysis und Anwendungen*. Walter de Gruyter, 1971.

[183] N. Wohllaib. Predictive Vision. *Pictures of the Future: The Magazine for Research and Innovation, Fall 2009*, pages 101–102, 2009.

[184] D. A. Wolf-Gladrow. *Lattice-Gas Cellular Automata and Lattice Boltzmann Models*. Springer, 2000.

[185] S. Wolfram. *A new kind of science*. Wolfram Media, 2002.

[186] D. Xiu. Efficient collocation approach for parametric uncertainty analysis. *Communications in Computational Physics*, 2:293–309, 2007.

[187] L. Yang, D. L. Boccelli, and J. G. Uber. Assessing Uncertainty in Chlorine Residual Predictions in Drinking Water Distribution System. In *Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium*. American Society of Civil Engineers, 2008.

[188] W. Zhang, C. T. Miller, and F. A. DiGiano. Bacterial Regrowth Model for Water Distribution Systems Incorporating Alternating Split-Operator Solution Technique. *Journal of Environmental Engineering*, 130:932–941, 2004.