SIAM/ASA J. UNCERTAINTY QUANTIFICATION
Vol. 2, pp. 29–54

Cellular Probabilistic Automata—A Novel Method for Uncertainty Propagation*

Dominic Kohler†, Johannes Müller‡, and Utz Wever§

Abstract. We propose a novel density based numerical method for uncertainty propagation under distinct partial differential equation dynamics. The main idea is to translate them into objects that we call cellular probabilistic automata and to evolve the latter. The translation is achieved by state discretization as in set oriented numerics and the use of the locality concept from cellular automata theory. We develop the method using the example of initial value uncertainties under deterministic dynamics and show that it is consistent. As an application we discuss arsenate transportation and adsorption in drinking water pipes and compare our results to Monte Carlo computations.

Key words. uncertainty propagation for partial differential equations, set oriented numerics, cellular automata

AMS subject classifications. 65C20, 37M05, 68Q80, 80A32

DOI. 10.1137/120897183

1. Introduction.

The numerical treatment of differential equations that are subject to uncertain data has attracted a lot of interest lately. A prominent approach is to use polynomial chaos expansions [49, 8, 22]. It can be improved by decomposing the random space [47], and only recently numerical implementations of this improvement have been investigated [1]. Alternative approaches are based on the Monte Carlo idea, like the Markov chain Monte Carlo method [10], Latin hypercube sampling [33], the quasi Monte Carlo method [7], importance sampling [34], and the multilevel Monte Carlo method [2]. Further well-known approaches use the Itô calculus [36, 24, 25, 38] or the Fokker–Planck equation [29]. Although the approaches have proven to be successful for many tasks, they often encounter certain efficiency restrictions in higher dimensions of the random space. New methods are needed to meet these challenges.

Time-continuous dynamical systems on continuous state space can be approximated by time-discrete Markov chains on finite state space [18]. This technique of state space discretization has led to the powerful tools of set oriented numerics [4, 5]. It is especially useful to study ergodic theory, asymptotic dynamics, and optimal control [28, 11]. Recently, also contributions to uncertainty quantification have been made [20].

In this paper we introduce a novel numerical scheme for uncertainty propagation in distinct spatio-temporal processes. It is based on the concept of state space discretization and on ideas from cellular automata (CAs) theory [21, 6, 15]. We develop the method for the example of

*Received by the editors October 31, 2012; accepted for publication (in revised form) October 15, 2013; published electronically February 4, 2014.
http://www.siam.org/journals/juq/2/89718.html

†Technische Universität München, Center for Mathematical Sciences, 85748 Garching/Munich, Germany, and Siemens AG, Corporate Technology, 81730 Munich, Germany (dominic.kohler@mytum.de).
‡Technische Universität München, Center for Mathematical Sciences, 85748 Garching/Munich, Germany, and Helmholtz Zentrum München, Institute of Computational Biology, 85764 Neuherberg, Germany (johannes.mueller@mytum.de).
§Siemens AG, Corporate Technology, 81730 Munich, Germany (utz.wever@siemens.com).

the propagation of initial value uncertainties under deterministic partial differential equation (PDE) dynamics and pave the way towards an extension to more general stochastic influences on the system.

In particular we introduce a discretization of a PDE which does not depend explicitly on the independent variables. First, a finite difference scheme is applied 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. As this procedure emphasizes the interaction between neighboring sites, a property that strongly resembles the locality and shift-invariance in CAs, 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.

Cellular probabilistic automata (CPAs) basically consist of information about transition probabilities between discretized portions of phase space in a site's neighborhood. The transition probabilities are interpreted as approximating the evolution of the system's probability density in transfer operator theory [30]. Hence CPA can be used for uncertainty propagation [35]. While the translation from PDEs into CPAs may be rather time-consuming, the evolution of uncertainties with CPAs 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.

The paper is structured as follows. In section 2 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 CPAs in section 3. Here we also present a consistency result for our construction. In section 4 we show how CPAs can be extended to incorporate stochastic boundary conditions and apply the theory to the example of arsenate transportation and adsorption in water pipes. The results are compared to Monte Carlo computations. Finally, we give our conclusions in section 5.

2. Density based uncertainty propagation.

In this section we first formulate the problem and then develop the idea of density based uncertainty propagation. Finally, the CPA idea is derived in this context.

2.1. Problem formulation.

We are interested in the time evolution of uncertain initial data in a specific deterministic dynamical system. First, we introduce some notation from probability theory and the Frobenius–Perron operator as the suitable tool to describe this process. Second, we specify the deterministic dynamical system that we will work with, and third, we formulate the problem.

Let $(X, \mathcal{A}, \mu)$ be a probability space, $(X', \mathcal{A}', \mu')$ a measure space, and $V : X \to X'$ a random variable with distribution μ_V. We say that V has density g if there is $g \in \mathcal{L}^1(X', \mathcal{A}', \mu')$ such that

$$\mu_V[A'] = \int_{A'} g d\mu' \quad \forall\, A' \in \mathcal{A}'.$$

The set of densities on $(X', \mathcal{A}', \mu')$ is denoted by

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

Now let $(X', \mathcal{A}', \mu') = (\mathbb{R}^{mn}, \mathcal{B}(\mathbb{R}^{mn}), \lambda)$, where $m, n \in \mathbb{N}$, $\mathcal{B}(\mathbb{R}^{mn})$ is the Borel σ-algebra, and λ is 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 [30].

Definition 2.1. *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. We focus on a particular type of phase space evolution.

Definition 2.2. *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 $I = \{1, \dots, m\}$;*

(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 $U = \{-r, \dots, s\}$ with $r, s \in \mathbb{N}_0$, $r + s \leq m$, such that Φ has the locality property, i.e., that there is $h : T \times (\mathbb{R}^n)^{|U|} \to \mathbb{R}^n$ with*

$$\Phi(t, v)_i = h(t, v_{i-r}, \dots, v_{i+s})$$

for all $t \in T$, $v = (v_1, \dots, v_m) \in \times_I \mathbb{R}^n$, and $i \in \{1 + r, \dots, m - s\}$, and

(v) *that the system acts as the identity on $K = \{1, \dots, r\} \cup \{m - s + 1, \dots, m\}$, i.e., $\Phi(t, v)|_K = v|_K$ 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 and refer the reader to [12] for further information on dynamical systems. 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 $a, b \in \mathbb{R}$. 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 scheme; e.g., usually $U = \{-1, 0, 1\}$ is suitable to account for central second order difference quotients. Because of the PDE context we call I the set of sites. By considering only trajectories with $v^0|_K = k \in \times_K \mathbb{R}^n$, the system can be interpreted as obeying boundary conditions.

The time evolution of uncertain initial data in the deterministic dynamical system is described by real random variables $V^0, V^1, \dots : X \to \Omega^m$ on probability space $(X, \mathcal{A}, \mu)$, where $V^{n+1} = \Phi^\tau V^n$. We focus on deterministic boundary conditions: $V^0(x)|_K = k \in \times_K \mathbb{R}^n$ 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.

2.2. State space discretization. In this section, first, we introduce the concept of state space discretization. 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 [4, 5]. Here they are adapted to the special structure of the dynamical system.

Definition 2.3. *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 2.4.

(i) *$\mathcal{L}^1_{\hat{T}}(\mathbb{R}^{mn}) = \operatorname{span}(B)$ 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 $\kappa_B : \mathcal{L}^1_{\hat{T}}(\mathbb{R}^{mn}) \to \mathbb{R}^{E^I},\ g \mapsto \kappa_B(g)$ 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 $\rho \in E^K$ such that $\rho_i = T(k_i)$ for all $i \in K$. The densities that are compatible with the boundary conditions are given by*

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

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

Definition 2.5. *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 [42].

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.

2.3. 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 $V = \{-p, \dots, q\}$, $W = \{-t, \dots, u\}$ with $p, q \in \mathbb{N}_0$, $t, u \in \mathbb{Z}, -t \le u$, and $l \in \mathbb{Z}$ we use the conventions $l + V = \{-p + l, \dots, q + l\}$ and $V + W = \{-p - t, \dots, q + u\}$.

Definition 2.6. *Let $V = \{-p, \dots, q\}$ with $p, q \in \mathbb{N}_0$ and $p + q + r + s \le m$. 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 $i = 1 + p + r$, $\varphi \in E^{U+V}$, and $\psi \in E^V$.

Note that because of the locality property the definition is independent of the chosen site $i \in \{1 + p + r, \dots, m - q - s\}$. The set V controls the degree of locality, i.e., the number of

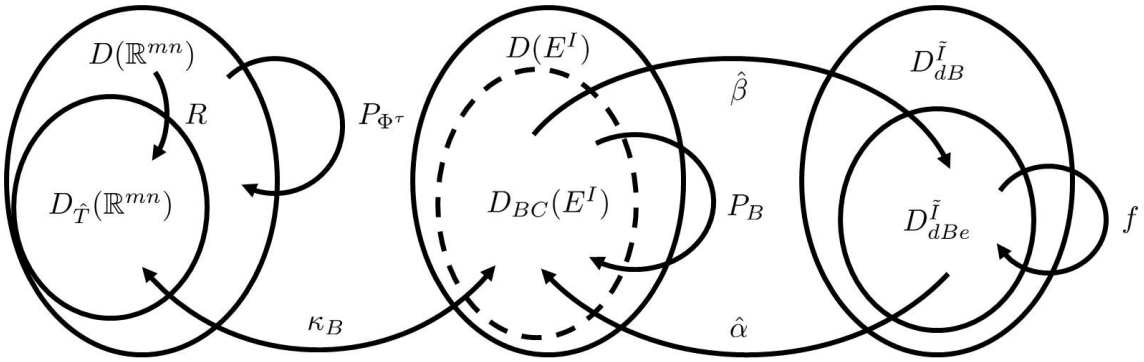

Figure 1. *The relations between the FPO P_{Φ^τ} and its approximations. By state discretization we obtain the discretized FPO P_B, which still works globally, and by exploiting locality we approximate P_B further by the CPA with global function f. The state space on which the CPA operates is a collection of local densities; see the text.*

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.

3. Cellular probabilistic automata.

CPAs are defined by extending the definition of deterministic CAs according to [6, 15]: In CPAs 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 CPAs to work on densities. This enables their utilization for uncertainty propagation.

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 $D(\mathbb{R}^{mn})$. CPAs further approximate the discretized FPO on a product space of local densities; see Figure 1 for a sketch. Uncertainty propagation with CPAs 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 section 3.1 we first introduce CPAs 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 section 3.2. This connection leads to the generalization of CPA to general V in section 3.3. Section 3.4 contains a consistency result for our algorithm and further investigation of the pattern approximation.

3.1. 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 $V = \{-p, \dots, q\}$ and $W = \{-t, \dots, u\}$ 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 $\alpha_W : (D(E^V))^W \to D(E^{V+W})$. β_W localizes the information to densities on states of size V and thus erases far-reaching cor-

relations. 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 3.1. *Let $W = \{-t, \dots, u\}$ 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) *We set $\beta_W : D(E^W) \to (D(E))^W$, $g \mapsto \hat{\beta}_W(g)$ 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 3.2. *A cellular probabilistic automaton (CPA) is a tuple (I, U, E, f_0), where for $m, r, s \in \mathbb{N}_0$ with $1 \le m$ and $1 + r + s \le m$,*

(i) *$I = \mathbb{Z}_m$ is a toroidal grid,*

(ii) *$U = \{-r, \dots, s\}$ 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 $n \in \mathbb{N}^+$.

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 Figure 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 deterministic CAs are special cases of CPAs: 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.

3.2. 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 on the basis of pattern ideas in CAs theory [14, 46], in the theory of de Bruijn graphs [40], and in pair approximation [23].

As before, we introduce a projection operator β_W that localizes the global information to densities on states of size V, this time $|V| \ge 1$. 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 3.3. *Let $V = \{-p, \dots, q\}$ and $W = \{-t, \dots, u\}$ 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 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

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

Example 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 [31], 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. Since a precise formulation is very technical, we only outline the construction and the according results at this point. The technical details and proofs of section 3.2 can be found in Appendix A.

We first introduce several definitions that are central to our approach; see also Figure 2(a).

Definition 3.4.

(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 size 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*

$$D_{dBe}^W = \{g \in D_{dB}^W \,|\, \times_{i \in W} \operatorname{supp} g(i) \in X_{dBe}^W\}.$$

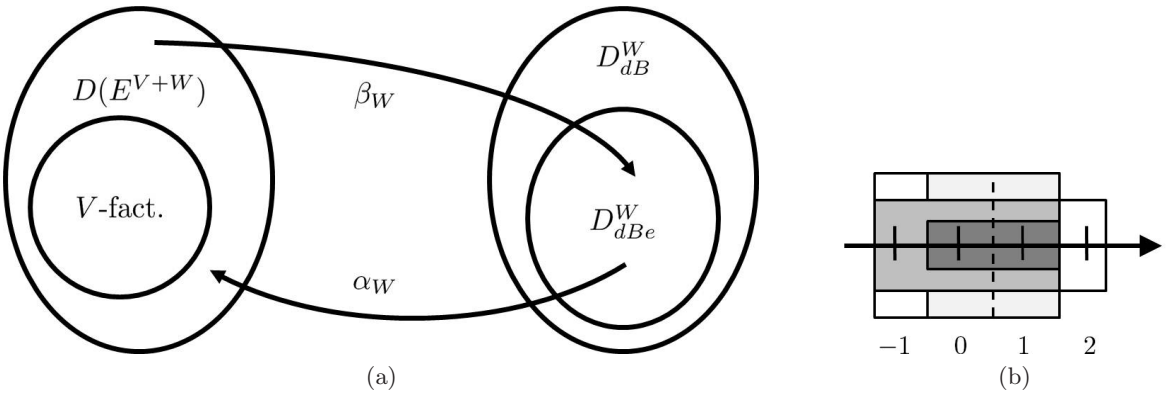

Figure 2. (a) *The relation between global densities and de Bruijn densities.* (b) *An example of how a global density is approximated by local densities via an approximation of Markov type for* $W = \{-1, 0, 1\}$, $V = \{0, 1\}$, *and* $i = 0$; *see Appendix* A *for details. The thin and dark grey box 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.*

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 Example 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 3.5. $\operatorname{im}(\beta_W) \subseteq D^W_{dBe}$.

Since only the extendable de Bruijn densities are addressed by β_W, we need only define α_W on D^W_{dBe}. 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 Figure 2(b). By repeating this procedure site by site to the left and right we cover the whole grid with the desired global state. However, we find that in general this construction depends on the site i.

Definition 3.6. *Let* $i \in W$. *Then* $\alpha_{W,i} : D^W_{dBe} \to D(E^{V+W})$ *is given by*

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

where $\mu_j[*|*]$ *denotes the conditional distribution associated with* $g(j) \in D(E^V)$ *for* $j \in W$, *and* $\psi \in E^{V+W}$. *Here* $\{\sigma_k(\psi)|_J\} := \{\varphi \in E^V \,|\, \varphi|_J = \sigma_k(\psi)|_J\}$, *where* $k \in W$ *and* $J \subseteq V$, *and* $V_+ = V \backslash \{-p\}$ *and* $V_- = V \backslash \{q\}$ *describe the overlap.*

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

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

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

Theorem 3.9. *Let $g \in$ im (β_W). Then $\beta_W \alpha_{W,i}(g) = g$ for all $i \in W$.*

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 3.10. *$g \in D(E^{V+W})$ 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 Example 1. g in the same example, however, has correlations over more than two 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 size $|\{0,1\}| = 2$, and therefore g is not $\{0,1\}$-factorizable.

Generally, correlations can be preserved if they are short-range with respect to V: $\alpha_{W,i}\beta_W$ will leave a global density invariant if a site's state is independent from the states at sites that are more than $|V| - 1$ sites apart. This can be formally understood in terms of a spatial Markov property of order $|V| - 1$; see Appendix A.

Obviously the set of V-factorizable states is the natural counterpart of im (β_W).

Lemma 3.11. *Let $i \in W$ and $g \in$ im (β_W). Then $\alpha_{W,i}(g)$ is V-factorizable.*

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

Theorem 3.12. *For all $g \in D(E^{V+W})$ there is a unique V-factorizable $\tilde{g} \in D(E^{V+W})$ such that $\beta_W(\tilde{g}) = \beta_W(g)$.*

β_W induces equivalence classes on $D(E^{V+W})$ by collecting all global densities with the same image in one class. According to Theorem 3.12 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 $i \in W$. Example 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 $g \in D^W_{dBe} \backslash$ im (β_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 $g \in$ im (β_W) the definition then coincides with any $\alpha_{W,i}$.

Definition 3.13. *$\alpha_W : D^W_{dBe} \to D(E^{V+W})$ 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 $\alpha_{W,i}(g)$.

3.3. General cellular probabilistic automata. With the de Bruijn calculus at hand we can now generalize the definition of CPAs to general V. To cope with boundary conditions it is necessary that the global function operate 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 3.14. *As before let $I = \{1, \dots, m\}$, $U = \{-r, \dots, s\}$, $V = \{-p, \dots, q\}$ for*

$p, q, r, s, m \in \mathbb{N}_0$ with $m \geq 1, 1 + p + q + r + s \leq m$, and $K = \{1, \dots, r\} \cup \{m - s + 1, \dots, m\}$. We now set $i_l = 1 + p + r$, $i_r = m - q - s$, and $\tilde{I} = \{i_l, \dots, i_r\}$.

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

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

(ii) *We set $\hat{\beta} : D_{BC}(E^I) \to D_{dBe}^{\tilde{I}}$, $g \mapsto \hat{\beta}(g)$ with*

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

In order to shorten notation at this point, we state the general definition for a CPA without specifying the terms "φ matches the boundary condition" and "$U(i)$ is the adaptation of U to the finite grid size." The precise meaning for these two statements is given in Appendix B.

Definition 3.15. *A cellular probabilistic automaton (CPA) is a tuple (I, U, V, E, f_0), where for $m, p, q, r, s \in \mathbb{N}_0$ with $m \geq 1$ and $1 + p + q + r + s \leq m$,*

(i) *$I = \{1, \dots, m\}$ is a finite grid,*
(ii) *$U = \{-r, \dots, s\}$ is the neighborhood,*
(iii) *$V = \{-p, \dots, q\}$ 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 definitions $i_l = 1 + p + r$, $i_r = m - q - s$, and $\tilde{I} = \{i_l, \dots, i_r\}$ the global function is given by

$$f : D_{dBe}^{\tilde{I}} \to D_{dBe}^{\tilde{I}}, \quad g \mapsto f(g),$$
$$f(g)(i)(\psi) = \sum \alpha_{\tilde{U}(i)}(\sigma_i(g)|_{\tilde{U}(i)})(\varphi|_{V+\tilde{U}(i)}) \cdot f_0(\varphi)(\psi),$$

where the sum is taken over all $\varphi \in E^{U+V}$ such that φ matches the boundary condition $\rho \in E^K$ in case of an overlap. $\tilde{U}(i)$ is the technical adaptation of U to the finite grid size m and almost coincides with U. 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 $n \in \mathbb{N}^+$.

See Figure 3 for a sketch of how the CPA works on general patterns. Note that f_0 is not arbitrary but connected to a dynamical system with the locality property. By exploiting this relation we can ensure that the global function is well defined; see Appendix B for the proof.

Lemma 3.16. *$f(D_{dBe}^{\tilde{I}}) \subseteq D_{dBe}^{\tilde{I}}$.*

We denote the case of $i_l = i_r$ with $V_{\max}$ and find that $\hat{\alpha}\hat{\beta}(g) = g$ for all $g \in D_{BC}(E^I)$ and $\hat{\beta}\hat{\alpha}(g) = g$ for all $g \in (D(E^V))^{\tilde{I}}$. Furthermore, $\tilde{U}(\tilde{I}) = \{0\}$ in this case, and it can be calculated for $g \in D_{dBe}^{\tilde{I}}$ and $\varphi \in E^{U+V}$ that

$$\alpha_{\{0\}}(\sigma_{i_l}(g))(\varphi|_V) = g(i_l)(\varphi|_V).$$

For $\psi \in E^V$ then

$$f(g)(i_l)(\psi) = \sum g(i_l)(\varphi|_V) \cdot f_0(\varphi)(\psi),$$

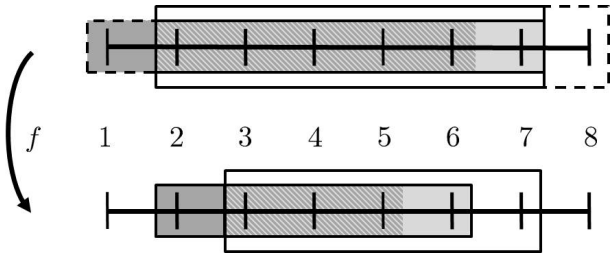

Figure 3. *An example of a CPA with $I = \{1, \dots, 8\}$, $U = \{-1, 0, 1\}$, and $V = \{-2, \dots, 2\}$. The global function considers patterns located at $\tilde{I} = \{4, 5\}$ and sketched in the image (lower part) by rectangles of small and large heights, respectively. The corresponding preimage patterns (upper part) are larger due to the neighborhood, and their probability of occurrence is influenced by boundary conditions (dashed parts). From an implementational point of view, V may be constructed from $\tilde{V} = \{-2, \dots, 1\}$ and $W = \{0, 1\}$; see section 4.2: 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.*

where the sum is taken over all $\varphi \in E^{U+V}$ such that the boundary conditions are fulfilled. 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 3.17. *For $V = V_{\max}$ we find that $\hat{\alpha} f \hat{\beta} = P_B$.*

3.4. 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 the two parameters' state space resolution and de Bruijn pattern size. In this section we first show that for maximal pattern size the approximation can be made arbitrarily close. A subsequent investigation of potential locality errors in the case of smaller pattern size complements this consistency result.

There are many ways to study distances between probability measures [9]. 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 [30]. A result by Li ensures that in this norm the discretized FPO converges pointwise to the original FPO for increasing state space resolution [32].

Theorem 3.18. *Let $g \in D(\mathbb{R}^{mn})$ with $\operatorname{supp}(g) \subseteq \Omega^m$, 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,*

$$\|R(g) - g\|_1 \to 0 \quad (\Delta\Omega \to 0).$$

Note that the one-dimensional proof by Li can immediately be extended to general dimension mn of the phase space.

With Proposition 3.17 we therefore find that for maximal pattern size $V_{\max}$ and a uniform partition with resolution $\Delta\Omega$ our algorithm is consistent: For $g \in \kappa_B^{-1}(D_{BC}(E^I))$,

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

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.

Lemma 3.19. *For all $n \in \mathbb{N}$, all $g \in D_{BC}(E^I)$, and all $i \in \tilde{I}$ it holds that*

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

Proof. Let $n \in \mathbb{N}$, $g \in D_{BC}(E^I)$, $i \in \tilde{I}$, and $\chi \in \operatorname{supp}\,(\hat{\beta}P_B^n(g)(i)) \in E^V$. 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 $\varphi_0, \dots, \varphi_{n-1} \in E^I$ 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 $k \in \{1, \dots, n\}$. 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, we conclude that $\hat{\beta}(g)(j)(\sigma_j(\varphi_0)) > 0$ 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 $\chi \in \operatorname{supp}\,(f^n\hat{\beta}(g)(i))$. ∎

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 size V. This is independent of the actual dynamics and a direct consequence of our approximation space; see section 3.2. 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.

Example 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 $I = \{1, 2, 3\}$ and the dynamical system that is given by the identity on $\Omega^m = [0, 1]^2$, i.e., $U = \{0\}$ and $K = \emptyset$. 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 $\tilde{I} = \{1, 2\}$. We find that $f_0(\varphi)(\psi) = \delta_{\varphi,\psi}$ for all $\varphi, \psi \in E^{\{0\}}$.

Consider $g \in D_{BC}(E^I) = D(E^{V+W})$ with $W = \tilde{I}$ from Example 1. $\hat{\beta}(g) = \beta_{\tilde{I}}(g)$, and also $f\hat{\beta}(g) = \beta_{\tilde{I}}(g)$. So $\hat{\alpha}\hat{\beta}(g) = \alpha_{\tilde{I}}\beta_{\tilde{I}}(g) = \tilde{g}$. However, $P_B(g) = g$, and so $\hat{\alpha}f\hat{\beta}(g) \neq P_B(g)$.

Example 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 locally. Consider $I = \{1, \dots, 4\}$ and the system on dynamically invariant state space $\Omega^m = [0, 1]^4$ given for all $n \in \mathbb{N}$ by $v_i^{n+1} = h(v_i^n, v_{i+1}^n) = \frac{v_i^n + v_{i+1}^n}{3.75}$ for $i \in \{1, \dots, m-1\}$. We have $U = \{0, 1\}$ and define five

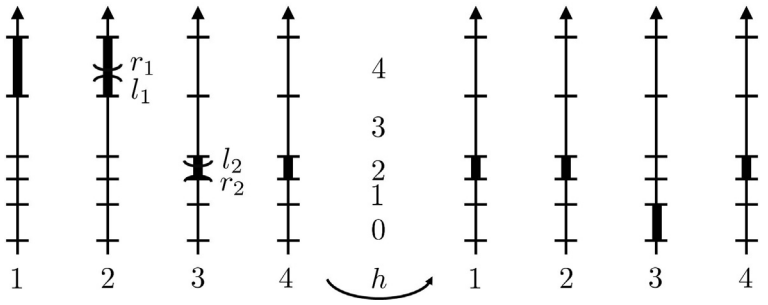

Figure 4. *Illustration of a transition from $\chi = (4,4,2,2)$ to $\psi = (2,2,0,2)$ in Example 3. The left side shows the preimage and the right the image state. The horizontal numbers correspond to the respective sites, while the vertical numbers display $E = \{0,\dots,4\}$. The states χ and ψ are marked by black rectangles at the corresponding sites.*

intervals

$$\Omega_j = [w_j, w_{j+1}), \quad \text{for } j \in \{0,\dots,3\}, \quad \Omega_4 = [w_4, w_5]$$

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 $E = \{0,\dots,4\}$; see Figure 4. The induced flow is denoted by Φ^1 for one time step. We consider the CPA with $V = U$ for deterministic input $g \in D_{BC}(E^I)$ given by $g(\varphi) = \delta_{\chi,\varphi}$, where $\chi = (4,4,2,2) \in E^I$. We focus on the image state $\psi = (2,2,0,2) \in E^I$ and determine

$$l_1 = 0.8, \quad l_2 = 0.3625, \quad r_1 = 0.83875, \quad r_2 = 0.32375$$

as the solution of the equations

$$h(w_4, l_1) = w_3, \quad h(l_1, l_2) = w_2, \quad h(r_1, r_2) = w_2, \quad h(r_2, w_2) = w_1.$$

It is possible to show that

$$\{\Omega_{\chi|_{1+U+V}} \cap \Phi^{-1}(\Omega_{\psi|_{1+V}})\} \subseteq \{v \in \Omega \,|\, w_4 \le v_2 \le l_1, l_2 \le v_3 \le w_3\},$$
$$\{\Omega_{\sigma_1(\chi|_{2+U+V})} \cap \Phi^{-1}(\Omega_{\sigma_1(\psi|_{2+V})})\} \subseteq \{v \in \Omega \,|\, r_1 \le v_2 \le 1, w_2 \le v_3 \le r_2\},$$

$f_0(\sigma_1(\chi|_{1+U+V}))(\sigma_1(\psi|_{1+V})) > 0$, and $f_0(\sigma_2(\chi|_{2+U+V}))(\sigma_2(\psi|_{2+V})) > 0$. Hence $\hat{\alpha} f \hat{\beta}(g)(\psi) > 0$, but

$$\begin{aligned} P_B(g)(\psi) &= \sum_{\varphi \in E^I} g(\varphi) P_{B,\varphi,\psi} = \frac{\lambda(\Omega_\chi \cap \Phi^{-1}(\Omega_\psi))}{\lambda(\Omega_\chi)} \\ &\le \frac{\lambda((\Omega_{\chi|_{1+U+V}} \cap \Phi^{-1}(\Omega_{\psi|_{1+V}})) \cap (\Omega_{\chi|_{2+U+V}} \cap \Phi^{-1}(\Omega_{\psi|_{2+V}})))}{\lambda(\Omega_\chi)} \\ &= \frac{\lambda(\emptyset)}{\lambda(\Omega_\chi)} = 0, \end{aligned}$$

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. Application. In this section we comment on how to implement an algorithm to evolve uncertainties with CPAs The algorithm is tested at the problem of arsenate transportation and adsorption in drinking water pipes, and the results are compared to a Monte Carlo calculation. Although the theory has been developed for uncertainties in initial conditions, in this applicational part we extend the concept slightly such that CPAs can cope with certain stochastic boundary conditions. This is important in contaminant transport modeling [48]. With this first generalization we want to provide evidence that with CPAs the treatment of more general stochastic spatio-temporal processes seems feasible.

4.1. Stochastic boundary conditions. We deal with stationary temporal white noise boundary conditions

$$\begin{aligned} g_l &\in D(E^{K_l}) \quad \text{for } K_l = \{1, \dots, r\}, \\ g_r &\in D(E^{K_r}) \quad \text{for } K_r = \{m - s + 1, \dots, m\} \end{aligned}$$

instead of deterministic $\rho \in E^K$. 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.

For this purpose the global function in Definition 3.15 is extended to

$$\begin{aligned} f : D_{dBe}^{\tilde{I}} \to D_{dBe}^{\tilde{I}}, \quad g \mapsto f(g), \\ f(g)(i)(\psi) = \sum_{\varphi \in E^{U+V}} \underline{g}(i)(\varphi) \cdot \alpha_{\tilde{U}(i)}(\sigma_i(g)|_{\tilde{U}(i)})(\varphi|_{V+\tilde{U}(i)}) \cdot \overline{g}(i)(\varphi) \cdot f_0(\varphi)(\psi), \end{aligned}$$

where $\underline{g}(i)$ and $\overline{g}(i)$ incorporate suitable parts of g_l and g_r, respectively, at the boundary sites and equal 1 otherwise. The precise definition is very technical and can be found in Appendix B. Note that $\hat{\alpha}$ and $\hat{\beta}$ also have to be generalized to match stochastic boundary conditions.

CPAs 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 remark that it is straightforward to use time-dependent stochastic boundary conditions instead of stationary ones.

4.2. Implementation. From an implementational point of view two steps of uncertainty propagation with CPAs 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. It turns out that step one is numerically more expensive than step two. For industrial applications like simulation based system monitoring the CPA method points towards real-time uncertainty quantification, because the slow step one

has to be performed only once before the actual simulation. We furthermore note that by construction the simulation is parallelizable in space.

Step one basically consists of the approximation of local transition probabilities. We propose a local version of the standard Monte Carlo quadrature approach [19] in set oriented numerics for this purpose. We remark that also advanced adaptive methods have been suggested; see, e.g., [13].

1. For $\varphi \in E^{U+V}$ choose W_φ test vectors $w_i = (w_{i,-r-p}, \dots, w_{i,s+q}) \in (\mathbb{R}^n)^{U+V}$, where $\{w_{i,j}\}_{i \le W_\varphi}$ is randomly distributed over coding domain $\Omega_{\varphi(j)} \subseteq \Omega$, respectively.
2. Compute for all $i \le W_\varphi$ the image points

$$\tilde{w}_i = (h(\tau, w_{i,-r-p}, \dots, w_{i,s-p}), h(\tau, w_{i,-r-p+1}, \dots, w_{i,s-p+1}), \dots, h(\tau, w_{i,-r+q}, \dots, w_{i,s+q})).$$

3. Determine $\psi_1, \dots, \psi_L \in E^V$ such that there is $l \le L$ 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

$$f_0(\varphi)(\psi) = \begin{cases} W_\psi / W_\varphi & \forall \psi \in \{\psi_1, \dots, \psi_L\}, \\ 0 & \text{else.} \end{cases}$$

With an increasing number of test points the approximation is expected to get better. However, the number of evaluations grows exponentially in $|V|$. So we suggest using de Bruijn calculus to determine transition probabilities for large V from transition probabilities for smaller $\tilde{V}$; see Figure 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 Example 3 that this is again just an approximation of the directly calculated f_0.

Regarding step two, the simulation with CPAs, we remark that it is important to follow and store only states with probability larger than a specified threshold whenever possible. Otherwise already for reasonably large E or V the calculations are not feasible. An example is the de Bruijn density $D^{\tilde{I}}_{dBe}$, where $|E|^{|V|}$ numbers would have to be handled at every site in $\tilde{I}$. The set of states with positive probability is much smaller, although it typically first grows and then shrinks again in the transient phase of dynamics. 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.

4.3. Arsenate fate in water pipe. Consider the advection and adsorption of arsenate in drinking water pipes, a topic that has attracted a lot of attention in the water supply community lately [39, 26]. 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 Figure 5(a). The physics is described by the Langmuir adsorption model [27]

$$\partial_t D + v\partial_x D = -\frac{1}{r_h}\frac{1}{\frac{1}{k_1}+\frac{1}{k_f}(S_{\max}-A)}(D(S_{\max}-A)-K_{eq}A),$$
$$\partial_t A = \frac{1}{\frac{1}{k_1}+\frac{1}{k_f}(S_{\max}-A)}(D(S_{\max}-A)-K_{eq}A),$$

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 [26, 37],

$$v = 10\frac{\mathrm{m}}{\mathrm{min}}, \qquad r_h = 50\frac{l}{\mathrm{m}^2},$$
$$k_1 = 0.2\frac{l}{\mathrm{mg\,min}}, \qquad S_{\max} = 100\frac{\mathrm{mg}}{\mathrm{m}^2},$$
$$K_{eq} = 0.0537\frac{\mathrm{mg}}{l}, \qquad k_f = 2.4\frac{l}{\mathrm{m}^2\,\mathrm{min}},$$

and consider the system on the approximately positively invariant Ω given by $D \in [0,1]\frac{\mathrm{mg}}{l}$ and $A \in [0,100]\frac{\mathrm{mg}}{\mathrm{m}^2}$. The backward difference with $U = \{-1,0\}$, $\Delta x = 100\mathrm{m}$, and $\Delta t = \Delta x/v = 10\mathrm{min}$ is used with the Trotter formula [41].

To obtain the local function of a CPA we map test points by using intermediate steps with the smaller $\Delta x' = 1\mathrm{m}$ and $\Delta t' = 0.1\mathrm{min}$. We use $V = \tilde{V} = \{0\}$ and partition the phase space equidistantly with five 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 Figure 5(b). 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. The observed dynamics is shown in Figures 5(c)–5(e): 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 Figures 6(a)–6(b). The latter has also been obtained on the basis of the Trotter formula 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 10min 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 state.

Furthermore, we plot the steady state results from CPAs for which the approximation parameters are altered. In Figure 6(c) the result is plotted for $\tilde{V} = V = \{0\}$ with an equidistant phase space partition of 5 domains in the D-direction and 15 domains in the A-direction,

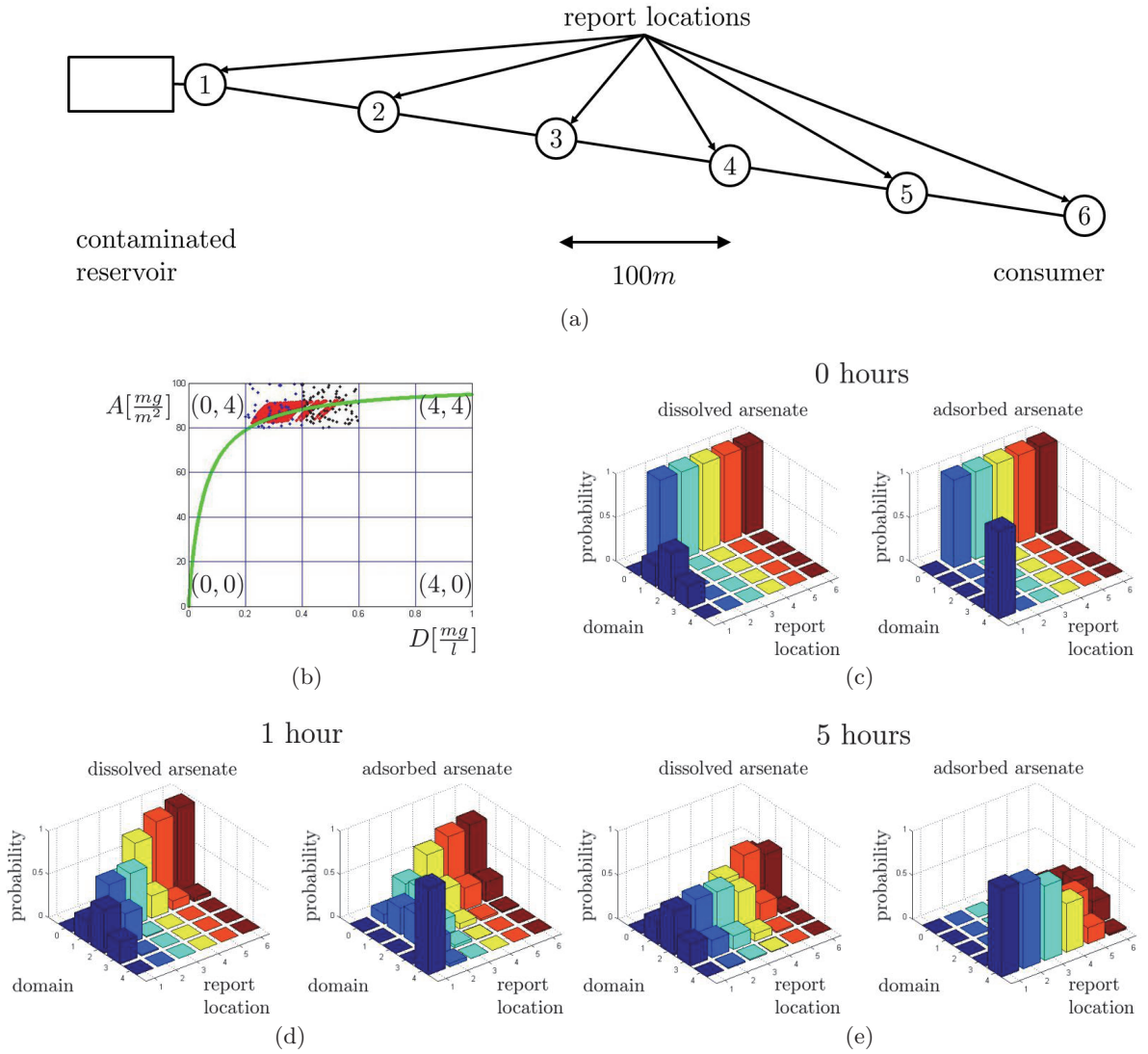

Figure 5. (a) *A reservoir on a hill is connected to a consumer in a valley through a pipe with six report locations.* (b) *The phase space at every report location is divided into* 5×5 *coding domains, and the steady states are drawn in green. For example,* $f_0(((1,4),(2,4)))$ *can be approximated 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 Figure* 6(b).

whereas in Figure 6(d) the pattern size is extended by $W = \{-1, 0\}$ to $V = \{-1, 0\}$. It is observed that in this example increasing the pattern size does not improve the result if compared to the Monte Carlo case, but increasing the state space resolution has a notable effect. In all cases we used 75 test points for the coding domain at site 0 and 37 at site -1 to approximate the local function and used a probability threshold of 0.00005 in the simulation.

We note that there is often no interest in global results and accordingly no need to trans-

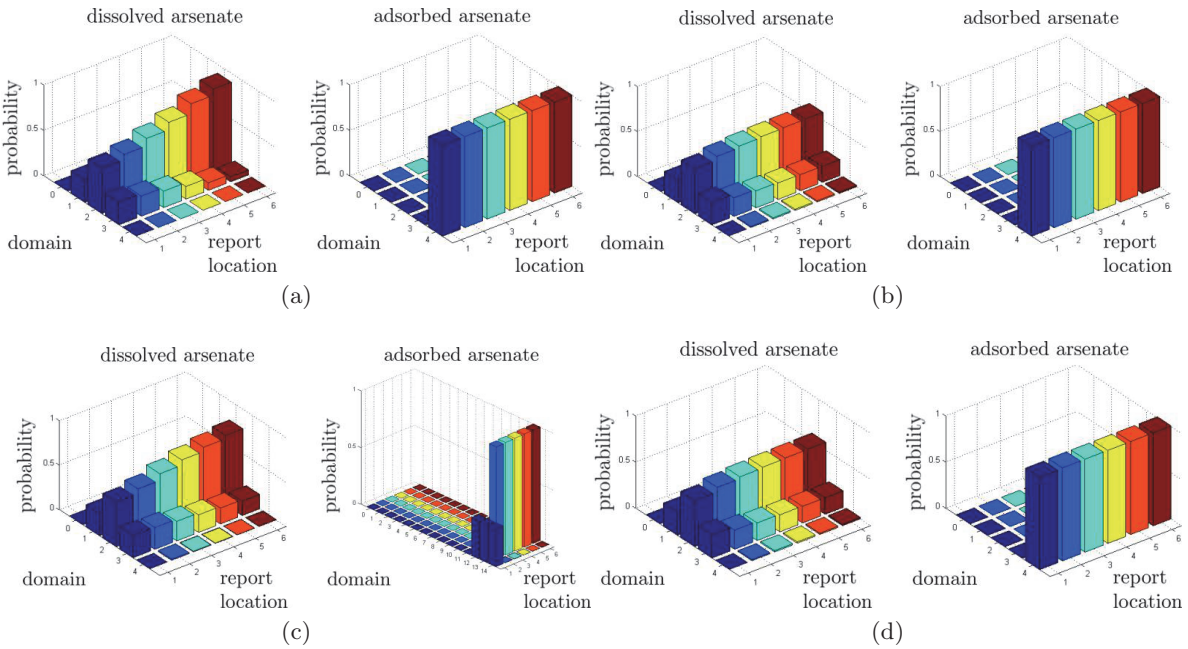

Figure 6. *Steady states after* 24 *hours.* (a) *shows the result of a Monte Carlo computation. In* (b) *one finds the result for a CPA with a state space resolution of* 5×5 *and pattern size* $\tilde{V} = V = \{0\}$*, in* (c) *the results for* 5×15 *and* $\tilde{V} = V = \{0\}$*, and in* (d) *those for* 5×5 *and* $\tilde{V} = \{0\}, V = \{-1, 0\}$.

form 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, we note that 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 [3].

5. Conclusion. We have introduced a numerical scheme for density based uncertainty propagation in distinct spatio-temporal systems. It consists of a preprocessing step, in which the underlying PDE system is translated into a CPA, and a simulation step, in which initial and boundary conditions are evolved. The simulation is parallelizable in the space extension and fast in relation to the preprocessing. Furthermore, it computes 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.

The algorithm is based on state space discretization like in set oriented numerics and on the de Bruijn state idea from CAs theory. There are two parameters that allow us to control the approximation of the exact density evolution: state space resolution and de Bruijn pattern size. We have shown consistency of the method for uncertain initial conditions under deterministic dynamics and 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.

We suggest three directions for future research. First, we are interested in really quantifying the approximation error for given de Bruijn pattern size beyond our consistency result. This seems to be a demanding problem, because spatial correlations have to be quantified somehow. One might use the rich theory on quantum (multipartite) entanglement measures [17, 16, 44, 45] and on the related matrix product states [43] for that purpose. Second, we will focus on random parameters. 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 preprocessing, whereas the simulation step is not changed. In this sense CPAs promise to overcome the curse of dimension in parameter space. And third, we want to investigate how our algorithm performs in more complex applications like the simulation of drinking or waste water grids.

Appendix A. De Bruijn calculus technique. In this appendix we introduce a more formal approach to the de Bruijn calculus of section 3.2 and give the proofs of the above results.

We start with a motivation of $\alpha_{W,i}$ in terms of a spatial Markov property, for which we first introduce some notation. For finite $J \subseteq \mathbb{Z}$ and $A, B \subseteq E^J$, $\mu[A] = \sum_{\varphi \in A} g(\varphi)$ denotes the distribution associated with $g \in D(E^J)$, and $\mu[A|B] = \frac{\mu[A \cap B]}{\mu[B]}$ denotes the conditional probability. Furthermore,

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

for finite $\tilde{J}, \hat{J} \subseteq \mathbb{Z}$, $\tilde{J} \subseteq J$, $\tilde{J} \subseteq \hat{J}$, and $\psi \in E^{\hat{J}}$, and we also write $\left\{\psi|_{\tilde{J}}\right\} = \{\psi|_{\tilde{J}}^{J}\}$ if J is clear from the context.

As before, let $V = \{-q, \dots, q\}$ and $W = \{-t, \dots, u\}$ with $p, q \in \mathbb{N}$ and $t, u \in \mathbb{Z}, -t \le u$. We calculate for $i \in W$ and $\psi \in E^{V+W}$ that

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

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

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

for $k \in \{-t, \dots, i-1\}$ and $l \in \{i+1, \dots, u\}$. They resemble a Markov property of order $|V| - 1$ in space; see also Figure 2(b): A site's state is independent from the states at sites that are more than $|V| - 1$ sites apart.

Lemma A.1. *Let μ_j be the distribution associated with $\beta_W(g)(j) \in D(E^V)$ for $j \in W$.*

Then

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

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

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

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

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

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

This leads directly to Definition 3.6 of $\alpha_{W,i}$.

In the remainder of this appendix we give the proofs of section 3.2.

Proof of Lemma 3.5. Let $g \in \operatorname{im}(\beta_W)$, $j \in W$, 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 $\times_{i\in W} \operatorname{supp} g(i) \in X^W_{dBe}$. $\blacksquare$

Proof of Lemma 3.7. 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

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

The steps indicated by $\cdots$ follow by induction in $|W|$. ∎

Proof of Lemma 3.8. We show that $\alpha_{W,i}(g) = \alpha_{W,i+1}(g)$ for all $i \in W\backslash\{-t\}$. 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 $k \in \{-t, \dots, i-1\}$ and $l \in \{i+1, \dots, u\}$,

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

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

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 $i \in W\backslash\{u\}$, and let $g = \beta_W(\tilde{g})$ for $\tilde{g} \in D(E^{V+W})$. Then

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

Proof of Theorem 3.9. 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 Lemma 3.8 and because a conditional distribution is a distribution as well, for $\psi \in E^V$,

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

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

Proof of Lemma 3.11. Let $g \in \operatorname{im}(\beta_W)$ and $i, j \in W$. Then by Lemma 3.8 and Theorem 3.9

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

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

Proof of Theorem 3.12. Let $g \in D(E^{V+W})$, and choose $\tilde{g} = \alpha_{W,u}\beta_W(g)$. Then $\tilde{g}$ is V-factorizable by Lemma 3.11, 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 Theorem 3.9. 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. ∎

Appendix B. Global functions. In this section we state the precise definition of the global function of general CPAs for deterministic boundary conditions and prove that it is well defined. Afterwards we extend the definition to stochastic boundary conditions.

Definition B.1 (extending Definition 3.15). *With the boundary conditions $\rho \in E^K$ for $K = \{1,\dots,r\}\cup\{m-s+1,\dots,m\}$ and the definitions $i_l = 1+p+r$, $i_r = m-q-s$, $\tilde{I} = \{i_l,\dots,i_r\}$, $\tilde{U}(i) = \{-\underline{u}(i),\dots,\overline{u}(i)\}$, and $\underline{u}, \overline{u} : \tilde{I} \to U$ with*

$$\underline{u}(i) = \begin{cases} i - i_l & \text{if } i \in \{i_l,\dots,i_l + r - 1\}, \\ r & \text{else,} \end{cases}$$

$$\overline{u}(i) = \begin{cases} i_r - i & \text{if } i \in \{i_r - s + 1,\dots,i_r\}, \\ s & \text{else,} \end{cases}$$

the global function is given by

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

$$f(g)(i)(\psi) = \sum_{\substack{\varphi \in E^{U+V} \text{ s.t. } \varphi(k-i)=\rho(k) \\ \text{for } k \in K \cap i+U+V}} \alpha_{\tilde{U}(i)}(\sigma_i(g)|_{\tilde{U}(i)})(\varphi|_{V+\tilde{U}(i)}) \cdot f_0(\varphi)(\psi).$$

Proof of Lemma 3.16. 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 $g \in D_{dBe}^{\tilde{I}}$, $i \in \{i_l,\dots,i_r - 1\}$, and $\psi \in \operatorname{supp} f(g)(i)$. By the construction of f and $\alpha_{\tilde{U}(i)}$ we know that there is $\varphi \in E^{U+V}$ such that $\sigma_j(\varphi)|_V \in \operatorname{supp}(g)(i+j)$ 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 $\sigma_j(\tilde{\varphi})|_V = \sigma_{j+1}(\varphi)|_V \in \operatorname{supp}(g)(i+1+j)$ for all $j \in \{-r,\dots,s-1\}$ and $\sigma_s(\tilde{\varphi})|_V \in \operatorname{supp}(g)(i+1+s)$.

This enables us to find $\tilde{\psi} \in \operatorname{supp} f(g)(i+1)$ 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 $i+1+q+s$ 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 $\tilde{\psi} \in \operatorname{supp} f(g)(i+1)$. ∎

For stochastic boundary conditions

$$\begin{aligned} g_l &\in D(E^{K_l}) \qquad \text{for } K_l = \{1, \dots, r\}, \\ g_r &\in D(E^{K_r}) \qquad \text{for } K_r = \{m-s+1, \dots, m\} \end{aligned}$$

the global function is extended to $f : D_{dBe}^{\tilde{I}} \to D_{dBe}^{\tilde{I}},\ g \mapsto f(g)$,

$$\underline{g}(i)(\varphi) = \sum_{\substack{\chi \in E^{K_l} \text{ s.t. } \sigma_{i-i_l}(\chi)|_{\{1,\dots,r+i_l-i\}} = \\ \sigma_{-i_l}(\varphi)|_{\{1,\dots,r+i_l-i\}}}} g_l(\chi),$$

$$\overline{g}(i)(\varphi) = \sum_{\substack{\chi \in E^{K_r} \text{ s.t. } \sigma_{i-i_r}(\chi)|_{\{m-s+1-i+i_r,\dots,m\}} = \\ \sigma_{-i_r}(\varphi)|_{\{m-s+1-i+i_r,\dots,m\}}}} g_r(\chi),$$

$$f(g)(i)(\psi) = \sum_{\varphi \in E^{U+V}} \underline{g}(i)(\varphi) \cdot \alpha_{\tilde{U}(i)}(\sigma_i(g)|_{\tilde{U}(i)})(\varphi|_{V+\tilde{U}(i)}) \cdot \overline{g}(i)(\varphi) \cdot f_0(\varphi)(\psi).$$

The relations between global and de Bruijn densities have to be generalized in the stochastic case to $\hat{\alpha} : D(E^{K_l}) \times D_{dBe}^{\tilde{I}} \times D(E^{K_r}) \to D(E^I)$, $(g_l \times g \times g_r) \mapsto \hat{\alpha}((g_l \times g \times g_r))$ with

$$\hat{\alpha}((g_l \times g \times g_r))(\psi) = g_l(\psi|_{K_l}) \alpha_{\tilde{I}}(g)(\psi|_{\{1+r,\dots,m-s\}}) g_r(\psi|_{K_r})$$

and $\hat{\beta} : D(E^I) \to D(E^{K_l}) \times D_{dBe}^{\tilde{I}} \times D(E^{K_r})$, $g \mapsto \hat{\beta}(g) = (g_l, g_{\tilde{I}}, g_r)$ with

$$g_l(\psi) = \sum_{\chi \in E^I \text{ s.t. } \chi|_{K_r} = \psi} g(\chi), \qquad g_r(\psi) = \sum_{\chi \in E^I \text{ s.t. } \chi|_{K_l} = \psi} g(\chi),$$

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

Acknowledgments. The authors would like to thank Birgit Obst from Siemens Corporate Technology for helpful discussions about the arsenate application. Advice from Oliver Junge from Technical University Munich on set oriented numerics is also gratefully acknowledged.

REFERENCES

[1] F. Augustin and P. Rentrop, *Stochastic Galerkin techniques for random ordinary differential equations*, Numer. Math., 122 (2012), pp. 399–419.
[2] A. Barth, C. Schwab, and N. Zollinger, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119 (2011), pp. 123–161.
[3] M. Davidich and S. Bornholdt, *The transition from differential equations to Boolean networks: A case study in simplifying a regulatory network model*, J. Theoret. Biol., 255 (2008), pp. 269–277.
[4] M. Dellnitz and O. Junge, *On the approximation of complicated dynamical behavior*, SIAM J. Numer. Anal., 36 (1999), pp. 491–515.
[5] M. Dellnitz and O. Junge, *Set oriented numerical methods for dynamical systems*, in Handbook of Dynamical Systems, Vol. 2, North–Holland, Amsterdam, 2002, pp. 221–264.
[6] A. Deutsch and S. Dormann, *Cellular Automaton Modeling of Biological Pattern Formation*, Birkhäuser, Boston, MA, 2005.
[7] B. L. Fox, *Strategies for Quasi-Monte Carlo*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.
[8] R. Ghanem and P. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.
[9] A. L. Gibbs and F. E. Su, *On choosing and bounding probability metrics*, Internat. Statist. Rev., 70 (2002), pp. 419–435.
[10] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall/CRC, Boca Raton, FL, 1995.
[11] L. Grüne and O. Junge, *Global optimal control of perturbed systems*, J. Optim. Theory Appl., 136 (2008), pp. 411–429.
[12] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1983.
[13] R. Guder, M. Dellnitz, and E. Kreuzer, *An adaptive method for the approximation of the generalized cell mapping*, Chaos Solitons Fractals, 8 (1997), pp. 525–534.
[14] F. v. Haeseler, H.-O. Peitgen, and G. Skordev, *Cellular automata, matrix substitutions and fractals*, Ann. Math. Artificial Intelligence, 8 (1993), pp. 345–362.
[15] J. Hawkins and D. Molinek, *One-dimensional stochastic cellular automata*, Topology Proc., 31 (2007), pp. 515–532.
[16] M. Horodecki, *Entanglement measures*, Quantum Inf. Comput., 1 (2001), pp. 3–26.
[17] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Quantum entanglement*, Rev. Modern Phys., 81 (2009), pp. 865–942.
[18] C. S. Hsu, *Cell-to-Cell Mapping: A Method of Global Analysis for Nonlinear Systems*, Springer, New York, 1987.
[19] F. Y. Hunt, *A Monte Carlo approach to the approximation of invariant measures*, Random Comput. Dynam., 2 (1994), pp. 111–133.
[20] 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, 2004.
[21] J. Kari, *Theory of cellular automata: A survey*, Theoret. Comput. Sci., 334 (2005), pp. 3–33.
[22] 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*, Research Report 2005-01, ETH Zürich, Zürich, Switzerland, 2005.
[23] J. M. Keeling, *The effects of local spatial structure on epidemiological invasions*, Proc. R. Soc. Lond. B, 266 (1999), pp. 859–867.
[24] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, New

York, 2011.
[25] P. E. Kloeden, E. Platen, and H. Schurz, *Numerical Solution of SDE through Computer Experiments*, Springer, Berlin, 1994.
[26] 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, Reston, VA, 2010.
[27] L. K. Koopal and M. J. Avena, *A simple model for adsorption kinetics at charged solid-liquid interfaces*, Colloid Surface Physicochem. Eng. Aspect, 192 (2001), pp. 93–107.
[28] H. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer, New York, 2001.
[29] H. P. Langtangen, *Numerical solution of first passage problems in random vibrations*, SIAM J. Sci. Comput., 15 (1994), pp. 977–996.
[30] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, Springer, New York, 1993.
[31] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, Prentice–Hall, Englewood Cliffs, NJ, 1974.
[32] T. Y. Li, *Finite approximation for the Frobenius-Perron operator. A solution to Ulam's conjecture*, J. Approximation Theory, 17 (1976), pp. 177–186.
[33] W. L. Loh, *On Latin hypercube sampling*, Ann. Statist., 24 (1996), pp. 2058–2080.
[34] R. E. Melchers, *Structural Reliability Analysis and Prediction*, John Wiley & Sons, Chichester, UK, 1999.
[35] I. Mezić and T. Runolfsson, *Uncertainty propagation in dynamical systems*, Automatica J. IFAC, 44 (2008), pp. 3003–3013.
[36] B. K. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, Springer, New York, 2002.
[37] M. L. Pierce and C. B. Moore, *Adsorption of arsenite and arsenate on amorphous iron hydroxide*, Water Res., 16 (1982), pp. 1247–1253.
[38] A. Rößler, *Runge–Kutta methods for the strong approximation of solutions of stochastic differential equations*, SIAM J. Numer. Anal., 48 (2010), pp. 922–952.
[39] F. Shang, J. G. Uber, L. A. Rossman, R. Janke, and R. Murray, *EPANET Multi-species Extension User's Manual*, US Environmental Protection Agency, Cincinnati, OH, 2011.
[40] K. Sutner, *De Bruijn graphs and linear cellular automata*, Complex Systems, 5 (1991), pp. 19–30.
[41] W. Thirring, *Quantum Mechanics of Atoms and Molecules*, Springer, New York, 1981.
[42] S. M. Ulam, *A Collection of Mathematical Problems*, Interscience, New York, 1960.
[43] F. Verstraete, J. I. Cirac, and V. Murg, *Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems*, Adv. Phys., 57 (2008), pp. 143–242.
[44] F. Verstraete, J. Dehaene, and B. De Moor, *Normal forms and entanglement measures for multipartite quantum states*, Phys. Rev. A (3), 68 (2003), 012103.
[45] F. Verstraete, M. Popp, and J. I. Cirac, *Entanglement versus correlations in spin systems*, Phys. Rev. Lett., 92 (2004), 027901.
[46] R. Vollmar, *Algorithmen in Zellularautomaten*, Teubner, Stuttgart, 1979.
[47] X. Wan and G. E. Karniadakis, *Multi-element generalized polynomial chaos for arbitrary probability measures*, SIAM J. Sci. Comput., 28 (2006), pp. 901–928.
[48] P. P. Wang and C. Zheng, *Contaminant transport models under random sources*, Ground Water, 43 (2005), pp. 423–433.
[49] N. Wiener, *The homogeneous chaos*, Amer. J. Math., 60 (1938), pp. 897–936.