

An adaptive Wavelet Method for The Chemical Master Equation

Tobias Jahnke

Preprint Nr. 10/02

INSTITUT FÜR WISSENSCHAFTLICHES RECHNEN
UND MATHEMATISCHE MODELLBILDUNG



Anschrift des Verfassers:

JProf. Dr. Tobias Jahnke
Institut für Angewandte und Numerische Mathematik
Karlsruher Institut für Technologie (KIT)
D-76128 Karlsruhe

AN ADAPTIVE WAVELET METHOD FOR THE CHEMICAL MASTER EQUATION*

TOBIAS JAHNKE†

Dedicated to Ernst Hairer on the occasion of his sixtieth birthday

Abstract. An adaptive wavelet method for the chemical master equation is constructed. The method is based on the representation of the solution in a sparse Haar wavelet basis, the time integration by Rothe’s method, and an iterative procedure which in each time-step selects those degrees of freedom which are essential for propagating the solution. The accuracy and efficiency of the approach is discussed, and the performance of the adaptive wavelet method is demonstrated by numerical examples.

Key words. stochastic reaction kinetics, chemical master equation, Haar wavelets, numerical integration, adaptive methods, high-dimensional problems, error bounds

AMS subject classifications. 65L05, 65L70, 65T60, 92C45

DOI. 10.1137/080742324

1. Introduction. Many biological processes are modeled as complex reaction systems in which different species interact via a number of reaction channels. Often the evolution of the entire system is crucially determined by one or two subpopulations which may consist of a very small number of individuals. This is the case, e.g., in gene regulatory networks where gene expression is regulated by a few activators or repressors, in viral kinetics where the fate of very few infectious individuals decides whether the infection spreads over large parts of the population, or in predator-prey systems where the presence of a few predators keeps the entire ecosystem in equilibrium. In all these examples, small changes in the population numbers of the critical species due to inherent stochastic noise can cause large-scale effects. Hence, a reasonable mathematical model of such processes must respect both the stochastic nature of the time evolution and the discreteness of the population numbers.

This insight has fostered an increasing interest in stochastic reaction kinetics. Here, the system is described by a time-dependent probability distribution $p(t, x)$ which, for each state $x = (x_1, \dots, x_d) \in \mathbb{N}^d$, indicates the probability that at time t exactly x_i individuals of the i th species exist. For given initial data $p(0, x) = \rho(x)$, the distribution $p(t, x)$ evolves according to the chemical master equation (CME), which is a particular type of the forward Kolmogorov equation. In most applications, solving the CME is a formidable task because many applications involve multiple scales in space and time, low regularity, and multimodal, metastable solutions. The main challenge, however, is the tremendous number of degrees of freedom, which originates

*Received by the editors December 1, 2008; accepted for publication (in revised form) September 14, 2009; published electronically January 8, 2010. Supported by the “Concept for the Future” of Karlsruhe Institute of Technology within the framework of the German Excellence Initiative, and the DFG priority programme SPP 1324 “Mathematische Methoden zur Extraktion quantifizierbarer Information aus komplexen Systemen.” A very short and preliminary form of this article was published in the Proceedings of the International Conference on Numerical Analysis and Applied Mathematics, Psalidi, Kos, Greece, 2008, AIP Conference Proceedings Vol. 1048, pp. 290–293.
<http://www.siam.org/journals/sisc/31-6/74232.html>

†Universität Karlsruhe (TH), Fakultät für Mathematik, Institut für Angewandte und Numerische Mathematik, Kaiserstr. 89-93, 76133 Karlsruhe, Germany (jahnke@math.uni-karlsruhe.de).

from the fact that the solution $p(t, x)$ has to be approximated in each state x of a large and multidimensional state space.

In the past few years several numerical methods for the CME have been proposed. In these approaches the number of degrees of freedom is reduced by, e.g., Krylov space techniques [3, 26], discrete orthogonal polynomials [7, 10, 11], dynamical low-rank approximations [23], sparse grids [18, 19], adaptive lumping of states [13], continuous or hybrid models [14, 19, 20, 29], or adaptive state space truncation [27]. Explicit solution formulas for monomolecular master equations have been proven in [22].

The new method presented here is based on an adaptive representation of the solution in a sparse Haar wavelet basis. Wavelets are a well-established tool to compress large data sets efficiently and to solve partial differential equations (cf. [5, 6, 30, 31]). An application to the chemical master equation is particularly promising because of the discreteness and simple geometry of the state space. In the wavelet representation the coefficients of smooth signals decay rapidly such that only a few terms are needed to approximate the data with high accuracy. Since the Haar basis is orthonormal, the truncation error can easily be controlled by computing the 2-norm of the truncated coefficients. Moreover, the fast wavelet transform allows one to switch between the canonical and the Haar basis at low computational costs. These and other well-known advantages of the Haar basis are used in the novel method to approximate the solution of the CME in a low-dimensional subspace which is chosen adaptively in each time-step. In order to select those degrees of freedom which are currently required to propagate the solution, Rothe's method is combined with an iterative procedure. In some respects our approach is loosely related to the method from [18], where a hierarchical aggregation-disaggregation strategy similar to nonadaptive Haar wavelets has been proposed. However, the resulting integrator is completely different from ours, and to the best of our knowledge, an attempt to solve the CME by adaptive wavelet techniques has not yet been made.¹

In the next two sections, we briefly summarize the most important facts about the CME (section 2) and the Haar basis (section 3). The purpose of these sections is to formulate the problem, to define our notation, and to compile the mathematical toolbox for later use. In section 4, the adaptive wavelet method is constructed step by step. Moreover, we prove an error bound and discuss the computational costs of the method. In section 5, the performance of the adaptive wavelet method is demonstrated by numerical examples. Possible refinements of the approach are sketched in the last section.

2. Stochastic reaction kinetics.

2.1. The chemical master equation. Consider a reaction system in which d species interact via K reaction channels ($d, K \in \mathbb{N}$).² Let $X_i(t) \in \mathbb{N}$ be the number of "particles" (e.g., molecules, animals, individuals, or other entities) of the i th species at time t . In stochastic reaction kinetics, the vector $X(t) = (X_1(t), \dots, X_d(t))$ is considered as the realization of a Markov jump process on the state space \mathbb{N}^d . If the k th reaction channel fires, X jumps from the current state x to the new state $x + \nu_k$. The vector $\nu_k \in \mathbb{Z}^d$ is called the stoichiometric vector and indicates how the particle numbers are changed by the reaction. Realizations of the Markov jump process can be generated with the stochastic simulation algorithm from [16]. The d -dimensional histogram of a sufficiently large number of realizations would yield, in principle, an

¹A sketch of the adaptive wavelet method presented here has already been published in [24].

²Here and below, \mathbb{N} means the set of all nonnegative integers including zero.

approximation to the probability distribution

$$p(t, x) = \mathbb{P}(X(t) = x), \quad x \in \mathbb{N}^d,$$

i.e., the probability that at time t there are exactly x_i particles of the i th species ($i = 1, \dots, d$). However, using stochastic simulations to approximate the probability distribution of a very reactive system turns out to be computationally inefficient because the random variable has to be updated every time one of the reaction channels fires (cf. the numerical examples in section 5).

Therefore, it is our goal to compute $p(t, x)$ directly by solving the CME

$$(2.1) \quad \partial_t p(t, \cdot) = \mathcal{A}p(t, \cdot),$$

where \mathcal{A} denotes the operator

$$(2.2) \quad (\mathcal{A}p(t, \cdot))(x) = \sum_{k=1}^K (\alpha_k(x - \nu_k)p(t, x - \nu_k) - \alpha_k(x)p(t, x))$$

and $\alpha_k(x) \geq 0$ is the propensity function of the k th reaction channel (cf. [17, 15]). For simplicity of notation, we let $p(t, x) = 0$ and $\alpha_k(x) = 0$ for all $x \notin \mathbb{N}^d$. Of course, (2.1) has to be complemented by an initial condition $p(0, \cdot) = \rho(\cdot)$ where ρ is a suitable probability distribution. The CME is a special form of the Kolmogorov forward equation; a derivation can be found in [17].

2.2. Truncation of the state space. Numerical computations are usually performed only on the truncated state space

$$(2.3) \quad \Omega_\xi = \{x \in \mathbb{N}^d : x_1 < \xi_1, \dots, x_d < \xi_d\}$$

with some suitably chosen truncation vector $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{N}^d$. On the new boundaries, one can impose the Dirichlet boundary condition

$$(2.4) \quad p(t, x) = 0 \quad \text{for all } x \in \mathbb{N}^d \setminus \Omega_\xi$$

or the discrete Neumann boundary condition

$$(2.5) \quad \alpha_k(x) = 0 \quad \text{for all } x \in \Omega_\xi \quad \text{with } x + \nu_k \notin \Omega_\xi,$$

which suppresses all reactions leading from x to a state outside the truncated state space. The error caused by truncation with Dirichlet boundary condition has been investigated in [27]. From now on, we simply assume that Ω_ξ is so large that $p(t, \cdot)$ almost vanishes outside the artificial boundary. Then, the truncation error can be neglected, and the choice of boundary conditions hardly makes any difference. Of course, the truncation vector ξ could be chosen adaptively, but for simplicity it will be kept fixed in what follows.

2.3. Properties of the truncated chemical master equation. If the Neumann boundary condition is imposed, this implies

$$\sum_{x \in \Omega_\xi} \dot{p}(t, x) = \sum_{x \in \Omega_\xi} \sum_{k=1}^K (\alpha_k(x - \nu_k)p(t, x - \nu_k) - \alpha_k(x)p(t, x)) = 0$$

and hence $\sum_{x \in \Omega_\xi} p(t, x) = \sum_{x \in \Omega_\xi} p(0, x) = 1$. All solutions with nonnegative initial data remain nonnegative for all times, because $p(t, x_0) = 0$ for some $x_0 \in \Omega_\xi$ and

$p(t, x) \geq 0$ for all $x \neq x_0$ implies $\dot{p}(t, x_0) = \sum_{k=1}^K \alpha_k(x_0 - \nu_k)p(t, x_0 - \nu_k) \geq 0$. Hence, the solution of the CME on the truncated state space is a probability distribution if $p(0, \cdot)$ is a probability distribution and (2.5) is used. If (2.5) is replaced by (2.4), then $\sum_{x \in \Omega_\xi} \dot{p}(t, x) \leq 0$ and hence $\sum_{x \in \Omega_\xi} p(t, x) \leq 1$. In this case, $1 - \sum_{x \in \Omega_\xi} p(t, x)$ is the probability that the system has left the truncated state space, but this probability remains small on bounded time intervals if ξ is sufficiently large.

By a slight abuse of notation, the operator acting on functions on the truncated state space is again denoted by \mathcal{A} . Since Ω_ξ contains $N = \xi_1 \cdots \xi_d$ states, \mathcal{A} is now bounded and could be represented as a $(N \times N)$ -matrix A with nonpositive diagonal entries $a_{ii} \leq 0$, nonnegative off-diagonal entries $a_{ij} \geq 0$, and column sum $\sum_{i=1}^N a_{ij} = 0$ (Neumann) or $\sum_{i=1}^N a_{ij} \leq 0$ (Dirichlet), respectively. Applying the Gershgorin circle theorem shows that every eigenvalue either is zero or has a strictly negative real part. If the Neumann boundary condition is used, then zero is indeed an eigenvalue with trivial left eigenvector $(1, \dots, 1)$. Thus, $(rI - \mathcal{A})$ is invertible for all $r > 0$, and all entries of the inverse are nonnegative, which follows from the fact that $-A$ is an M -matrix.

2.4. The numerical challenge. The truncated (finite) state space contains $N = \xi_1 \cdots \xi_d$ states, and each state corresponds to one degree of freedom in $p(t, \cdot)$. In typical applications the dimension d and/or the upper limits ξ_i are so large that any attempt to solve the CME with traditional methods (e.g., by applying a Runge–Kutta scheme, by computing the matrix exponential in a straightforward manner, or by computing its eigenbasis) is absolutely hopeless. Any numerical method for the CME must reduce the tremendous number of degrees of freedom considerably without destroying the essential information. Such a method is constructed in section 4.

2.5. Notation. Let $\mathcal{H}(\Omega_\xi)$ be the linear space of all discrete functions $f : \Omega_\xi \rightarrow \mathbb{R}$. Any element $f \in \mathcal{H}(\Omega_\xi)$ can be represented as a vector (if $d = 1$), a matrix (if $d = 2$), or a tensor (if $d > 2$) because f maps any (multi-)index x to a real number. Nevertheless, we rather consider f as a function (usually a probability distribution) on a discrete space and write $f(x)$ instead of f_x . $\mathcal{H}(\Omega_\xi)$ is a Hilbert space with respect to the inner product

$$(2.6) \quad \langle f, g \rangle = \sum_{\Omega_\xi} f(x)g(x), \quad f, g \in \mathcal{H}(\Omega_\xi),$$

and the induced norm $\|f\|_2 = \sqrt{\langle f, f \rangle}$. However, we also consider the norm

$$\|f\|_1 = \sum_{\Omega_\xi} |f(x)|, \quad f \in \mathcal{H}(\Omega_\xi),$$

which is sometimes more appropriate because the solution of the CME is a probability distribution.

3. Haar wavelets. In this section we provide a brief summary of the most important facts about the Haar basis; for a more general introduction to wavelets, we refer to, e.g., [4]. The Haar wavelet is usually defined as a function on a real interval, but since the state space of the CME is discrete, we adapt all definitions to the discrete setting. For simplicity, assume that $\xi_1 = \cdots = \xi_d = 2^r$ for some $r \in \mathbb{N}$ such that $N = 2^{rd}$ is the total number of states.

3.1. Haar basis in one dimension. First, we consider the one-dimensional case, i.e., $d = 1$ and $\Omega_\xi = \{0, \dots, 2^r - 1\}$. Let $0 \leq l \leq r$, put $L = 2^{r-l}$, and, for $k = 1, \dots, 2^l$, define ϕ_k^l by

$$(3.1) \quad \phi_k^l(x) = \begin{cases} 1/\sqrt{L} & \text{if } (k-1)L \leq x < kL, \\ 0 & \text{else.} \end{cases}$$

The set $\{\phi_1^l, \dots, \phi_{2^l}^l\}$ is an orthogonal basis of the subspace $\mathcal{S}^l = \text{span}\{\phi_1^l, \dots, \phi_{2^l}^l\} \subset \mathcal{H}(\Omega_\xi)$. The spaces $\mathcal{S}^0 \subset \mathcal{S}^1 \subset \dots \subset \mathcal{S}^r = \Omega_\xi$ are nested because $\phi_k^l = (\phi_{2k-1}^{l+1} + \phi_{2k}^{l+1})/\sqrt{2}$. Note that $\mathcal{S}^r = \Omega_\xi$ because Ω_ξ contains only 2^r states; this is *not* true for wavelets on a real interval where infinitely many refinements are possible. Let \mathcal{P}^l denote the orthogonal projection from $\mathcal{H}(\Omega_\xi)$ to \mathcal{S}^l . Then, every $f \in \mathcal{H}(\Omega_\xi)$ can be represented by the multiscale decomposition

$$(3.2) \quad f = \mathcal{P}^r f = \mathcal{P}^0 f + \sum_{l=0}^{r-1} (\mathcal{P}^{l+1} - \mathcal{P}^l) f,$$

and the term $(\mathcal{P}^{l+1} - \mathcal{P}^l) f$ is called the *detail* added to the approximation of f on the $(l + 1)$ th level. This term can be represented by the functions

$$(3.3) \quad \psi_k^l(x) = \frac{\phi_{2k-1}^{l+1}(x) - \phi_{2k}^{l+1}(x)}{\sqrt{2}} = \begin{cases} \frac{1}{\sqrt{L}} & \text{if } (k-1)L \leq x < (k-\frac{1}{2})L, \\ -\frac{1}{\sqrt{L}} & \text{if } (k-\frac{1}{2})L \leq x < kL, \\ 0 & \text{else,} \end{cases}$$

where $L = 2^{r-l}$ is the number of nonzero entries of ψ_k^l , $l \in \{0, \dots, r-1\}$ is the refinement level, and $k \in \{1, \dots, 2^l\}$ is the number of the element on that level. All ψ_k^l can be obtained by shifting and scaling of the mother wavelet $\text{const.} \cdot (-1, \dots, -1, 1, \dots, 1)$ and padding the remaining entries with zeros. The coarsest subspace $\mathcal{P}^0 \mathcal{H}(\Omega_\xi)$ is spanned by the base line $\bar{\phi} = \phi_1^0 = (1, \dots, 1)/\sqrt{2^r}$. The set

$$(3.4) \quad \{\bar{\phi}\} \cup \{\psi_k^l \mid 0 \leq l \leq r-1, 1 \leq k \leq 2^l\}$$

is an orthonormal basis of $\mathcal{H}(\Omega_{2^r})$, called the *Haar basis*. Hence, every $p \in \mathcal{H}(\Omega_\xi)$ can be represented as

$$(3.5) \quad p = \bar{a} \bar{\phi} + \sum_{l=0}^{r-1} \sum_{k=1}^{2^l} a_k^l \psi_k^l, \quad a_k^l = \langle \psi_k^l, p \rangle, \quad \bar{a} = \langle \bar{\phi}, p \rangle.$$

Each coefficient and basis element depends on two parameters, but it is often convenient to order the elements in a linear way by a suitable enumeration ψ_1, \dots, ψ_N . In this notation, (3.5) reads $p = \sum_{j=1}^N a_j \psi_j$ with coefficients $a_j = \langle \psi_j, p \rangle$.

3.2. Properties of the Haar basis in one dimension. Using the Haar basis for solving the CME is motivated by several advantageous properties of (3.4). Besides the facts that the basis is orthonormal and that the elements allow fast evaluations due to their particularly simple form, a very useful feature is the *fast Haar wavelet transform* (cf. subsection 1.2 in [4]) which performs the mapping

$$(a_1, \dots, a_N) \longmapsto \sum_{j=1}^N a_j \psi_j$$

and its inverse with only $\mathcal{O}(N)$ instead of $\mathcal{O}(N^2)$ operations. Another important advantage is the observation that the wavelet coefficients of a smooth signal decay quickly because consecutive refinements add less and less detail information; cf. [4]. It is this property which allows one to reduce the number of degrees of freedom in the approximation of the CME, because the fast decay of the coefficients and the orthonormality of the basis mean that the signal can still be very well approximated if many of the terms in the expansion (3.5) are discarded. The following lemma relates smoothness — measured in terms of the difference between neighboring entries — to decay of the coefficients. The lemma is illustrated in Figure 3.1.

LEMMA 1. Let $d = 1$, $\xi = 2^r$, $0 \leq l \leq r$, $L = 2^{r-l} = \xi/2^l$, and $p \in \mathcal{H}(\Omega_\xi)$. Let a_k^l be the coefficient corresponding to the k th basis element on the l th level, i.e.,

$$a_k^l = \langle \psi_k^l, p \rangle = \frac{s_1 - s_2}{\sqrt{L}}, \quad s_1 = \sum_{x=(k-1)L}^{(k-\frac{1}{2})L-1} p(x), \quad s_2 = \sum_{x=(k-\frac{1}{2})L}^{kL-1} p(x).$$

If $\max_{0 \leq x < \xi-1} |p(x) - p(x+1)| < \varepsilon$ for some $\varepsilon > 0$, then

$$(3.6) \quad |a_k^l| \leq \frac{1}{4} \left(\frac{\xi}{2^l} \right)^{\frac{3}{2}} \varepsilon.$$

Remark 1. The estimate is sharp since equality is obtained for $p(x) = \varepsilon x$.

Remark 2. This lemma is a discrete counterpart of the well-known estimates in the continuous case; see page 25 in [4] or Equation (1.2.9) in [5].

Proof. For simplicity, we define $p^{(i)} = p(kL - L + i - 1)$ such that

$$(p^{(1)}, \dots, p^{(L)}) = (p(kL - L), \dots, p(kL - 1)), \quad s_1 = \sum_{i=1}^{L/2} p^{(i)}, \quad s_2 = \sum_{i=L/2+1}^L p^{(i)}.$$

By definition we have

$$s_1 - s_2 = \sum_{i=1}^{L/2} (p^{(i)} - p^{(i+L/2)}) = \sum_{i=1}^{L/2} \sum_{j=1}^{L/2} (p^{(i+j-1)} - p^{(i+j)}),$$

and since $|p^{(i+j-1)} - p^{(i+j)}| \leq \varepsilon$ by assumption, this yields

$$|a_k^l| = \frac{|s_1 - s_2|}{\sqrt{L}} \leq \frac{1}{\sqrt{L}} \left(\frac{L}{2} \right)^2 \varepsilon = \frac{1}{4} \left(\frac{\xi}{2^l} \right)^{\frac{3}{2}} \varepsilon. \quad \square$$

The Haar basis shares the above properties (orthonormality, fast transforms, fast decay of the coefficients) with the Fourier basis. What makes the Haar basis more suitable, however, is the fact that the support of most basis elements is small. This means, in particular, that a local change of the signal does not induce changes in all coefficients, and that the unpleasant Gibbs phenomenon is avoided when nonsmooth signals have to be approximated. Again, this is important in case of the CME, because often the solution varies only within a small subset of the state space, and nonsmooth solution profiles appear in certain applications (see, e.g., the numerical example in subsection 5.2).

In the context of partial differential equations, it has often been pointed out that the basis functions (when defined as functions rather than vectors) are discontinuous,

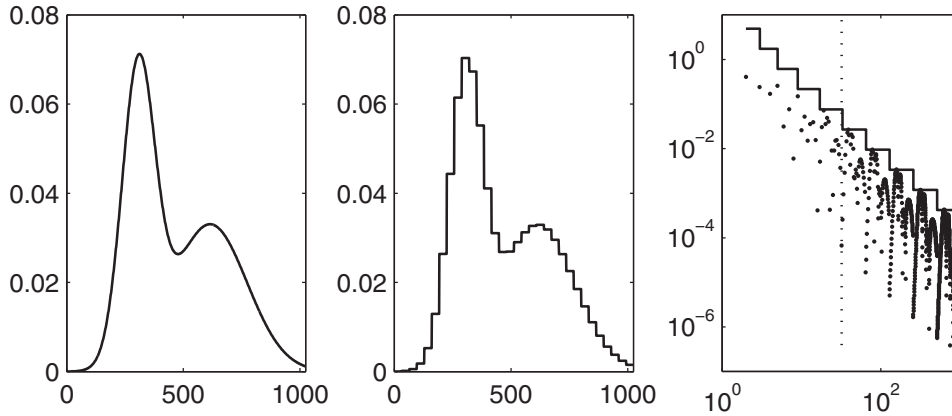


FIG. 3.1. Left panel: A smooth signal $p \in \Omega_\xi$ with $\xi = 2^{10}$. Middle panel: Approximation in the space spanned by the base line $\bar{\phi}$ and the basis elements ψ_k^l corresponding to the first five levels ($l = 0, \dots, 4$). This approximation reduces the number of degrees of freedom from 1024 to 32. Right panel: Absolute value of the Haar coefficients a_k^l (dots) and the bound from Lemma 1 (solid line) in logarithmic scaling. Each step of the “staircase” corresponds to one refinement level. Higher levels contain more and more coefficients. Only the coefficients on the left-hand side of the dashed line were used to compute the approximation shown in the middle panel.

that using this basis on domains with complicated geometry requires considerable technical efforts, and that posing boundary conditions can become difficult. All these arguments, however, do not apply in case of the CME. Here, the truncated state space Ω_ξ is simply a d -dimensional cuboid, the boundary conditions can be incorporated into the definition of the operator, and the discontinuity of the Haar basis does not matter since the solution is a spatially discrete object anyway. For all these reasons, a wavelet approximation of the CME is particularly promising.

3.3. Haar basis in many dimensions. There are several ways to extend the wavelet concept to the multivariate case. Probably the most popular alternative is to replace the one-dimensional basis elements defined in (3.1) by tensor products

$$\phi_K^l = \phi_{k_1}^l \otimes \dots \otimes \phi_{k_d}^l, \quad \phi_K^l(x) = \phi_{k_1}^l(x_1) \dots \phi_{k_d}^l(x_d),$$

where $K \in \mathbb{N}^d$ is a multi-index with entries $k_i \in \{1, \dots, 2^l\}$. As before, the subspaces $\mathcal{S}^l = \text{span}\{\phi_K^l \mid K = (k_1, \dots, k_d), k_i \in \{1, \dots, 2^l\}\} \subset \mathcal{H}(\Omega_\xi)$ are nested, and $f = \mathcal{P}^r f = \mathcal{P}^0 f + \sum_{l=0}^{r-1} (\mathcal{P}^{l+1} - \mathcal{P}^l) f$ for every $f \in \mathcal{H}(\Omega_\xi)$ with \mathcal{P}^l denoting the orthogonal projection from $\mathcal{H}(\Omega_\xi)$ to \mathcal{S}^l . The detail $(\mathcal{P}^{l+1} - \mathcal{P}^l) f$ can be represented by tensor products of univariate basis elements, but this time, mixed products composed of ϕ_k^l and ψ_j^l have to be used. As an example, let $d = 2$ and consider some $\phi_K^l = \phi_{k_1}^l \otimes \phi_{k_2}^l$ with $l < r$. On the next refinement level $l + 1$, there are four basis elements which are supported on a subset of the support of ϕ_K^l , namely,

$$\phi_{2k_1-1}^{l+1} \otimes \phi_{2k_2-1}^{l+1}, \quad \phi_{2k_1-1}^{l+1} \otimes \phi_{2k_2}^{l+1}, \quad \phi_{2k_1}^{l+1} \otimes \phi_{2k_2-1}^{l+1}, \quad \text{and} \quad \phi_{2k_1}^{l+1} \otimes \phi_{2k_2}^{l+1}.$$

In order to represent the detail added on the support of ϕ_K^l , three basis elements have to be added, namely,

$$\phi_{k_1}^l \otimes \psi_{k_2}^l, \quad \psi_{k_1}^l \otimes \phi_{k_2}^l, \quad \text{and} \quad \psi_{k_1}^l \otimes \psi_{k_2}^l$$

(see [28] for an illustration). In higher dimensions, the number of wavelets required to represent the detail added on each refinement level is $2^d - 1$. To be more precise, the Haar basis in d dimensions is the set

$$(3.7) \quad \left\{ \psi_{k_1}^{l,\sigma_1} \otimes \cdots \otimes \psi_{k_d}^{l,\sigma_d} \mid \psi_{k_i}^{l,0} = \phi_{k_i}^l, \psi_{k_i}^{l,1} = \psi_{k_i}^l, (\sigma_1, \dots, \sigma_d) \in \{0, 1\}^d \setminus 0 \right\} \cup \bar{\phi}$$

(cf. [4]) where $\bar{\phi} = \phi_1^0 \otimes \cdots \otimes \phi_1^0$ is the d -dimensional base line. As in the one-dimensional case, it is often more convenient to avoid multi-indices by a suitable enumeration of the basis elements and the corresponding coefficients.

The d -dimensional setting is notationally more complicated, but the composition via tensor products of univariate basis elements allows one to carry over all advantages of the one-dimensional Haar basis — orthonormality, fast transforms, localization, fast decay of the coefficients of smooth signals, conceptual simplicity — to the multivariate situation.

4. Adaptive Galerkin method with sparse wavelet basis.

4.1. Best approximation with m terms. The third panel of Figure 3.1 shows that the bound (3.6) can be much too pessimistic. Even on the coarsest approximation levels, some of the coefficients might almost vanish. Therefore, it is usually *not* an efficient strategy to approximate a signal by simply truncating the Haar representation (3.5) or its multidimensional counterpart after a certain refinement level l . This would be inefficient, e.g., if the signal almost vanishes on a large subset of the state space, which is typically the case for solutions of the CME.

Let $\{\psi_1, \dots, \psi_N\}$ be the Haar basis (3.7), with $N = \xi_1 \cdots \xi_d$ denoting the total number of states. Let $p = \sum_{j=1}^N a_j \psi_j$ and suppose that the basis elements are ordered in such a way that $|a_1| \geq \cdots \geq |a_N|$. Since the Haar basis is orthonormal, the best m -term approximation with respect to $\|\cdot\|_2$ (i.e., the best approximation which can be obtained by a linear combination of $m \in \mathbb{N}$ basis elements) is given by

$$\tilde{p} = \sum_{j=1}^m a_j \psi_j \quad \text{with error} \quad \|p - \tilde{p}\|_2 = \left(\sum_{j=m+1}^N |a_j|^2 \right)^{\frac{1}{2}}.$$

Hence, an ideal method would select the truncation index m in such a way that the error is smaller than a prescribed tolerance and approximate the solution by \tilde{p} . This would formally reduce the number of degrees of freedom from N down to $m \ll N$. However, there are at least two reasons why this simple strategy cannot be applied in practice:

- Although only a_1, \dots, a_m are required to compute the approximation, the truncation error can only be computed if the truncated coefficients are also available. If these coefficients were known, however, one could as well compute the exact solution.
- Since the solution $p(t)$ of the CME changes in time, the set of essential basis elements (that is, the set of basis elements which appear in the best m -term approximation) varies because some of the $a_{m+1}(t), \dots, a_N(t)$ might increase, whereas some of the $a_1(t), \dots, a_m(t)$ could vanish. This is clearly the case when transport effects dominate the evolution of $p(t)$, because each basis element has a local support. Again, selecting the essential basis elements is simple if *all* coefficients are known, but this is not a reasonable assumption.

In the next subsections we show how these problems can be avoided and how the essential basis elements of the *unknown* solution can be identified.

4.2. Nonadaptive Galerkin method. As a preparatory step we first consider the traditional, nonadaptive Galerkin method. Let $p(0, \cdot) = \rho(\cdot)$ be the initial distribution of the CME and assume that its best m -term approximation is given by³

$$\rho \approx y(0) = \sum_{j=1}^m a_j(0)\psi_j$$

(possibly after a suitable permutation of the basis elements). The CME can be projected to the low-dimensional Galerkin space $\text{span}\{\psi_1, \dots, \psi_m\}$ by imposing the Galerkin condition

$$\langle \partial_t y - \mathcal{A}y, \psi_i \rangle = 0 \quad \text{for all } i \in \{1, \dots, m\}.$$

The evolution of the coefficients $a = (a_1, \dots, a_m)$ corresponding to y is then given by the differential equation

$$(4.1) \quad \frac{da}{dt} = A_m a \quad \text{with} \quad A_m = \left(\langle \psi_i, \mathcal{A}\psi_j \rangle \right)_{i,j=1,\dots,m}.$$

After solving (4.1), an approximate solution of the CME is obtained from the linear combination $y(t) = \sum_{j=1}^m a_j(t)\psi_j$. The differential equation (4.1) contains only m degrees of freedom and can therefore be solved with significantly lower computational costs provided that $m \ll N$. However, the approximation error $\|y(t, \cdot) - p(t, \cdot)\|$ can become prohibitively large as t increases, even if $\|y(0, \cdot) - p(0, \cdot)\|$ is small. The reason is that $y(t)$ is confined to the small subspace $\text{span}\{\psi_1, \dots, \psi_m\}$, whereas the exact solution $p(t)$ propagates in the entire space \mathcal{H} . As pointed out in the previous subsection, the coefficients corresponding to some of the $\psi_{m+1}, \dots, \psi_N$ can increase so much that these terms can no longer be neglected. This does *not* mean that it is impossible to approximate the solution accurately with m terms — it only means that the Galerkin space has to be chosen adaptively.

4.3. Adaptive Galerkin method. The key idea of adaptive Galerkin methods is to update the Galerkin space such that the solution can always be sufficiently well approximated. Let $h > 0$ be the step size and suppose that

$$(4.2) \quad p(t_n) \approx p_n = \sum_{i=1}^{\eta} a_i \psi_{j_i}$$

where $t_n = t_0 + nh$ and $\{j_1, \dots, j_\eta\}$ is a subset of the index set $\{1, \dots, N\}$. Our task is to find a new selection $\{k_1, \dots, k_\mu\} \subset \{1, \dots, N\}$ and new coefficients b_i such that

$$p(t_{n+1}) \approx p_{n+1} = \sum_{i=1}^{\mu} b_i \psi_{k_i}.$$

In other words, in each time-step we have to identify the essential degrees of freedom and to propagate the solution in the corresponding basis. Of course, the number of coefficients μ is supposed to be as small as possible.

To this end we apply Rothe's method, which has been investigated in [1, 2] in the context of parabolic partial differential equations and has been applied to the

³Here and below, we will occasionally omit the spatial variable and write $p(t)$ instead of $p(t, x)$.

CME and similar problems in [7, 8, 32]. In contrast to the method of lines which first projects the problem into an approximation subspace and then performs the time discretization, Rothe's method first discretizes the problem in time and then adapts the spatial approximation in each time-step. Let $u_{n+1} \approx p(t_{n+1})$ denote the approximation obtained by the trapezoidal rule⁴ applied to the CME (2.1). For given u_n , the new approximation u_{n+1} is the solution of the linear system

$$(4.3) \quad \left(I - \frac{h}{2}\mathcal{A}\right) u_{n+1} = \left(I + \frac{h}{2}\mathcal{A}\right) u_n.$$

This system contains all N degrees of freedom and is therefore too large to be solved. However, it is neither necessary nor reasonable to solve (4.3) *exactly* because the time discretization causes a small approximation error anyway. Therefore, it is sufficient to approximate $u_{n+1} \approx p_{n+1}$ up to a certain tolerance \mathbf{tol} , i.e., to determine the k_i and b_i such that

$$(4.4) \quad \left\| \left(I - \frac{h}{2}\mathcal{A}\right) p_{n+1} - \left(I + \frac{h}{2}\mathcal{A}\right) p_n \right\|_1 \leq \mathbf{tol} \quad \text{holds for} \quad p_{n+1} = \sum_{i=1}^{\mu} b_i \psi_{k_i}.$$

When the new selection of basis elements is known, the coefficient vector $b = (b_i)_i$ can easily be computed by projecting the system (4.3) into the corresponding (low-dimensional) Galerkin space and solving

$$(4.5) \quad \Psi^* \left(I - \frac{h}{2}\mathcal{A}\right) \Psi b = \Psi^* \left(I + \frac{h}{2}\mathcal{A}\right) p_n,$$

where Ψ denotes the operator which maps the coefficient vector $b = (b_1, \dots, b_\mu)$ to the corresponding linear combination, i.e.,

$$(4.6) \quad \Psi : \mathbb{R}^\mu \longrightarrow \mathcal{H}(\Omega_\xi), \quad \Psi b = \sum_{i=1}^{\mu} b_i \psi_{k_i}.$$

Since the Haar basis is orthogonal, the inverse of Ψ is the adjoint operator

$$(4.7) \quad \Psi^* : \mathcal{H}(\Omega_\xi) \longrightarrow \mathbb{R}^\mu, \quad \Psi^* y = (b_1, \dots, b_\mu) \quad \text{with} \quad b_i = \langle \psi_{k_i}, y \rangle.$$

4.4. Selecting new basis elements. The main problem, however, remains: *How can we identify the essential basis elements* in each time-step; i.e., how do we select $\{k_1, \dots, k_\mu\}$? Clearly, it would be highly inefficient to change the basis randomly until (4.4) is met.

In this subsection, we show how to *enlarge* the present basis in a systematic way. The question which elements can be *removed* from the basis is easier and will be answered in subsection 4.5. Let $\{\psi_{k_1}, \dots, \psi_{k_\mu}\}$ be the basis from the previous step (i.e., $\mu = \nu$ and $k_i = j_i$), and let $p_n = \Psi a$ be the approximation at $t = t_n$. After solving (4.5), the residual of the new approximation $p_{n+1}^{(0)} = \Psi b$ is

$$r = \left(I - \frac{h}{2}\mathcal{A}\right) p_{n+1}^{(0)} - \left(I + \frac{h}{2}\mathcal{A}\right) p_n = p_{n+1}^{(0)} - p_n - \frac{h}{2}\mathcal{A} \left(p_{n+1}^{(0)} + p_n\right).$$

⁴The trapezoidal rule is chosen because it is A -stable and simple, but the same approach could be made with any other implicit method. A -stability is advantageous because all eigenvalues of the operator \mathcal{A} have a nonnegative and possibly large real part (cf. subsection 2.3).

Since (4.5) implies $\Psi^*r = 0$ and since $(I - \Psi\Psi^*)p_{n+1}^{(0)} = (I - \Psi\Psi^*)p_n = 0$, we obtain

$$(4.8) \quad r = (I - \Psi\Psi^*)r = -(I - \Psi\Psi^*)\frac{h}{2}\mathcal{A}(p_{n+1}^{(0)} + p_n).$$

(4.8) suggests that if the residual r is too large, then the current basis should be extended by so many new basis elements that the projection of the residual into the orthogonal complement of the enlarged Galerkin space is almost zero. However, enlarging the Galerkin space changes the approximation $p_{n+1}^{(0)}$ obtained from (4.5) and the residual r , too. Therefore, an iterative strategy is applied. First, we select only $\Delta\mu \in \mathbb{N}$ new basis elements $\psi_{k_{\mu+1}}, \dots, \psi_{k_{\mu+\Delta\mu}}$, where, for simplicity, the number $\Delta\mu \in \mathbb{N}$ is supposed to be defined a priori. The new elements are chosen in such a way that

$$\left\| (I - \Psi^{(1)}\Psi^{(1)*})r \right\|_2$$

is as small as possible for the current residual r , where $\Psi^{(1)}$ and $\Psi^{(1)*}$ are defined like Ψ and Ψ^* , but with μ replaced by $\mu + \Delta\mu$. In practice, this is achieved by computing a fast wavelet transform of r and finding the largest (in modulus) coefficients among those degrees of freedom which have not yet been selected. Then, a refined approximation $p_{n+1}^{(1)} = \Psi^{(1)}b^{(1)}$ is computed by solving

$$\Psi^{(1)*} \left(I - \frac{h}{2}\mathcal{A} \right) \Psi^{(1)}b^{(1)} = \Psi^{(1)*} \left(I + \frac{h}{2}\mathcal{A} \right) p_n$$

for $b^{(1)}$ and letting $p_{n+1}^{(1)} = \Psi^{(1)}b^{(1)}$. If the new residual

$$r^{(1)} = -(I - \Psi^{(1)}\Psi^{(1)*})\frac{h}{2}\mathcal{A}(p_{n+1}^{(1)} + p_n)$$

is small enough, the approximation is accepted. Otherwise, the procedure is iterated, which yields a series of approximations $p_{n+1}^{(0)}, p_{n+1}^{(1)}, p_{n+1}^{(2)}, \dots$, corresponding to an increasing family of approximation spaces. The full algorithm is listed in subsection 4.6.

4.5. Removing basis elements. In order to keep the computational workload as low as possible, basis elements which have been included in earlier steps must be discarded when they become negligible. Let $\hat{p}_{n+1} = p_{n+1}^{(l)} = \sum_{i=1}^{\hat{\mu}} b_i\psi_{k_i}$ be the approximation obtained after $l \in \mathbb{N}$ enlargements of the Galerkin basis, and assume that (after a suitable enumeration) $|b_1| \geq |b_2| \geq \dots \geq |b_{\hat{\mu}}|$. Let $p_{n+1} = \sum_{i=1}^{\mu} b_i\psi_{k_i}$ be the best μ -term approximation of \hat{p}_{n+1} (cf. subsection 4.1). Our goal is to choose μ as small as possible under the condition that the truncation error $\hat{p}_{n+1} - p_{n+1}$ does not exceed the tolerance \mathbf{tol} . If the error is measured with respect to $\|\cdot\|_2$, the truncation error is the 2-norm of the discarded coefficients (cf. subsection 4.1), and one can simply set

$$\mu = \min \left\{ m \in \mathbb{N} : \left(\sum_{i=m+1}^{\hat{\mu}} |b_i|^2 \right)^{\frac{1}{2}} \leq \mathbf{tol} \right\}.$$

In many cases, however, it is desirable to use the norm $\|\cdot\|_1$ because the solution of the CME is a probability distribution. In this case, one could use the equivalences of

norms and set

$$\mu = \min \left\{ m \in \mathbb{N} : \left(\sum_{i=m+1}^{\hat{\mu}} |b_i|^2 \right)^{\frac{1}{2}} \leq \sqrt{N} \text{tol} \right\}.$$

Because of the large factor \sqrt{N} this choice is often too pessimistic, which means that more basis elements are stored than actually needed. Better estimates for μ can be obtained iteratively by truncating a certain number of coefficients and checking how this affects the error $\|\hat{p}_{n+1} - p_{n+1}\|_1$. This procedure requires one fast wavelet transform in each iteration, but the additional numerical work is moderate because it is not necessary to determine the *optimal* value of μ .

In practice it is impossible to predict a priori how many degrees of freedom are required to reach the chosen accuracy. On the other hand, the limited memory of the computer imposes an upper bound for the maximal number of basis elements. Therefore, it is sometimes more convenient to prescribe the number of degrees of freedom instead of the accuracy of the approximation. In this case, the numbers μ and $\mu_{max} = \mu + \text{const.} \cdot \Delta\mu$ are chosen by the user. While μ is the number of basis elements which are kept *after* the time-step, μ_{max} denotes the maximal number of basis elements which are used *during* the time-step. This strategy was used in the numerical examples in section 5.

4.6. Algorithm. Summarizing, one time-step of the adaptive wavelet method for the CME proceeds as follows.

Parameter: step size $h > 0$, parameter $\Delta\mu \in \mathbb{N} \setminus \{0\}$, tolerance $\text{tol} > 0$
and/or maximal number of basis elements μ, μ_{max}
(the proper choice of tol is discussed in the next subsection)

Input: index subset $\{j_1, \dots, j_\eta\}$ and coefficients a_1, \dots, a_η of the
current approximation $p_n = \sum_{i=1}^{\eta} a_i \psi_{j_i} = \Psi a$
Galerkin matrix $\Psi^* \left(I - \frac{h}{2} \mathcal{A} \right) \Psi$

Output: index subset $\{k_1, \dots, k_\mu\}$ and coefficients b_1, \dots, b_μ of the new
approximation $p_{n+1} = \sum_{i=1}^{\mu} b_i \psi_{k_i}$
updated Galerkin matrix

1. Set $\hat{\mu} = \eta$ and $k_i = j_i$ for all $i = 1, \dots, \eta$.
2. Solve the linear system

$$\Psi^* \left(I - \frac{h}{2} \mathcal{A} \right) \Psi b = \Psi^* \left(I + \frac{h}{2} \mathcal{A} \right) p_n$$

and obtain the coefficients $b_1, \dots, b_{\hat{\mu}}$.

3. Compute the new approximation $\hat{p}_{n+1} = \sum_{i=1}^{\hat{\mu}} b_i \psi_{k_i} = \Psi b$.
4. Compute the residual $r = \left(I - \frac{h}{2} \mathcal{A} \right) \hat{p}_{n+1} - \left(I + \frac{h}{2} \mathcal{A} \right) p_n$.
5. If $\|r\|_1 > \text{tol}$ and/or $\hat{\mu} + \Delta\mu \leq m$:
 - (a) Compute $\chi_l = |\langle \psi_l, \mathcal{A}(p_n + \hat{p}_{n+1}) \rangle|$ for $l = 1, \dots, N$ by a fast wavelet transform.
 - (b) Find the indices $k_{\hat{\mu}+1}, \dots, k_{\hat{\mu}+\Delta\mu}$ of the $\Delta\mu$ largest entries of (χ_1, \dots, χ_N) among those degrees of freedom which are not contained in the current basis.

- (c) Enlarge the basis: add $\psi_{k_{\hat{\mu}+1}}, \dots, \psi_{k_{\hat{\mu}+\Delta\mu}}$ to the current set of basis elements and put $\hat{\mu} \mapsto \hat{\mu} + \Delta\mu$.
 - (d) Update the Galerkin matrix by adding new blocks corresponding to the new basis vectors.
 - (e) Go to step 2.
6. Compute p_{n+1} , μ , and $\{k_1, \dots, k_\mu\}$ by removing dispensable basis elements from the representation of \hat{p}_{n+1} as explained in subsection 4.5. The corresponding columns and lines are deleted from the Galerkin matrix.

Remark 1. In each step of the iteration a linear system has to be solved (in step 2). This can be done efficiently because in each iteration only $\Delta\mu$ rows and columns are appended to the matrix from the previous step. If a direct solver is used, the matrix decomposition available from the previous step can be updated. If an iterative method is applied, the old approximation \hat{p}_{n+1} provides an excellent starting value.

Remark 2. In order to start the algorithm, step 6 is applied to the initial distribution ρ . This yields a sparse approximation $p_0 = \sum_{i=1}^{\eta} a_i \psi_{j_i} \approx \rho$ which is the input for the very first time-step.

4.7. Accuracy. Next, we investigate how the choice of the tolerance `tol` affects the global error or, conversely, how this tolerance should be chosen in relation to the step size h . For convenience of the reader, the most relevant definitions are compiled in the following list:

- Let $p(t)$ denote the exact solution of the CME $\partial_t p(t, \cdot) = \mathcal{A}p(t, \cdot)$ with initial condition $p(0, \cdot) = \rho(\cdot)$ on the truncated state space Ω_ξ in the time interval $[0, t_{\text{end}}]$.
- Let u_n be the approximation obtained by formally applying the trapezoidal rule

$$(4.9) \quad \left(I - \frac{h}{2}\mathcal{A}\right) u_{n+1} = \left(I + \frac{h}{2}\mathcal{A}\right) u_n, \quad u_0 = \rho$$

with step size $h > 0$ for $n = 0, \dots, n_{\text{max}}$ where $n_{\text{max}} \leq t_{\text{end}}/h$. (4.9) includes the full CME operator \mathcal{A} without any spatial approximation. Thus, the error $\|p(t_n) - u_n\|$ is the error of the time integration only. As we have explained earlier, the u_n cannot really be computed in practice due to the large size of the linear system (4.9).

- Let p_n be the approximation computed by the adaptive wavelet method. Our method approximates u_n up to a small error which originates from the following two sources:
 - The linear system is only solved up to a small residual r_n ; i.e., for given p_n we determine \hat{p}_{n+1} such that

$$(4.10) \quad \left(I - \frac{h}{2}\mathcal{A}\right) \hat{p}_{n+1} = \left(I + \frac{h}{2}\mathcal{A}\right) p_n + r_n.$$

- At the end of each time-step, basis elements corresponding to small coefficients are removed from the representation of \hat{p}_{n+1} (cf. subsection 4.5 and step 6 in subsection 4.6). The result of this truncation is p_{n+1} , and the corresponding difference is denoted by $\delta_{n+1} = p_{n+1} - \hat{p}_{n+1}$.

The error $\|p(t_n) - u_n\|$ of the temporal approximation can be estimated in a straightforward way.

LEMMA 2. *The global error of the trapezoidal rule (4.9) is bounded by*

$$(4.11) \quad \|p(t_n) - u_n\|_1 \leq Ch^2 t_{\text{end}} \|\mathcal{A}^3 \rho\|_1$$

for all $n = 1, \dots, n_{\max}$. The proof of this lemma uses standard arguments and is therefore omitted.

Now, the error of the adaptive wavelet method is investigated. The following theorem states that if r_n and δ_{n+1} are small enough, then the accuracy of the adaptive wavelet method is essentially the same as the accuracy of the trapezoidal rule (4.9) without spatial approximation. “Small enough” means that \mathbf{tol} should be in the same order of magnitude as the local error of the time integration.

THEOREM 1. *Let ρ be the initial distribution, and assume that the sparse wavelet approximation p_0 used as input for the first time-step (cf. remark 2 in subsection 4.6) satisfies $\|\rho - p_0\|_1 \leq Ch^2$. Let $\mathbf{tol} = Ch^3$ and suppose that $\|r_n\|_1 \leq \mathbf{tol}$ and $\|\delta_{n+1}\|_1 \leq \mathbf{tol}$ for all n . Then the error of the method is bounded by*

$$(4.12) \quad \|p(t_n) - p_n\|_1 \leq Ch^2 t_{\text{end}} \|\mathcal{A}^3 \rho\|_1$$

for all $n = 1, \dots, n_{\max}$. Here, C denotes constants which, at different occurrences, can take different values.

Remark. In most applications, the accuracy of the adaptive wavelet method is much better than the error bound (4.12) predicts. The rather large constant $\|\mathcal{A}^3 \rho\|_1$ is often too pessimistic. Moreover, the factor t_{end} suggests that the error increases linearly in time, which is usually not the case because the solution converges to a stationary distribution.

Proof. After applying the triangle inequality

$$\|p(t_n) - p_n\|_1 \leq \|p(t_n) - u_n\|_1 + \|u_n - p_n\|_1$$

and Lemma 2, we have to investigate only the defect $\|u_n - p_n\|_1$, that is, the error caused by the spatial approximation. Comparing (4.9) with (4.10) yields

$$(4.13) \quad p_n - u_n = \Phi_h(p_{n-1} - u_{n-1}) + \varepsilon_n$$

with $\varepsilon_n = (I - h\mathcal{A}/2)^{-1} r_{n-1} + \delta_n$ and with $\Phi_h = (I - h\mathcal{A}/2)^{-1} (I + h\mathcal{A}/2)$ denoting the propagator of the numerical flux. It can be shown by induction that applying (4.13) recursively gives

$$(4.14) \quad p_n - u_n = \Phi_h^n(p_0 - \rho) + \sum_{k=0}^{n-1} \Phi_h^k \varepsilon_{n-k-1} \quad \text{for all } n > 1.$$

Since the semigroup generated by \mathcal{A} is contractive (cf. [10]) the Hille–Yosida theorem implies that $\|(\lambda I - \mathcal{A})^{-1}\|_1 \leq 1/\lambda$ for all $\lambda > 0$ (see, e.g., [12]), and choosing $\lambda = 2/h$ yields $\|(I - \frac{h}{2}\mathcal{A})^{-1}\|_1 \leq 1$. Substituting

$$\|\varepsilon_j\|_1 \leq \left\| \left(I - \frac{h}{2}\mathcal{A} \right)^{-1} r_{j-1} \right\|_1 + \|\delta_j\|_1 \leq \|r_{j-1}\|_1 + \|\delta_j\|_1 \leq 2\mathbf{tol} \leq Ch^3$$

into (4.14) and using that $\|p_0 - \rho\|_1 \leq Ch^2$ by assumption gives

$$(4.15) \quad \|p_n - u_n\|_1 \leq \|\Phi_h^n\|_1 \cdot Ch^2 + n \max_k \|\Phi_h^k\|_1 \cdot Ch^3 \quad \text{for all } n = 1, \dots, n_{\max}.$$

Since

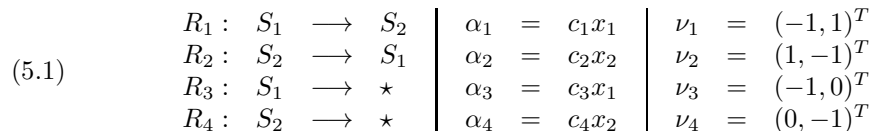
$$(4.16) \quad \|\Phi_h^k\|_1 \leq \|\exp(kh\mathcal{A})\|_1 + \|(\Phi_h^k - \exp(kh\mathcal{A}))\|_1 \leq 1 + Ch^2 t_{\text{end}} \|\mathcal{A}^3\|_1$$

is bounded, the assertion follows. \square

4.8. Computational cost. The efficiency of the method is due to the fact that the CME is projected to a small subspace which is adaptively changed along with the solution. This allows one to replace the huge linear system (4.3) by the much smaller problem (4.5) which involves only $\mu \ll N$ degrees of freedom. The only terms that are computed on the full state space are $\mathcal{A}\hat{p}_{n+1}$, $\mathcal{A}p_n$, the residual r , and the coefficients χ_l . Evaluations of \mathcal{A} can be performed at a cost of $\mathcal{O}(N)$ instead of $\mathcal{O}(N^2)$ because the operator \mathcal{A} is sparse in the sense that each entry is coupled to at most K other states if the system involves K reactions. The coefficients χ_l are obtained from a fast wavelet transform at the cost of $\mathcal{O}(N)$ operations. The Galerkin terms $\langle \psi_{k_i}, \mathcal{A}\psi_{k_l} \rangle$ required for assembling the Galerkin matrix do not require any operations on the large space in most cases, because the majority of Haar basis elements has only a small support. Moreover, the evaluation is greatly simplified by the fact that these elements are piecewise constant and all nonzero entries are ± 1 (up to a normalization constant). Thus, our method does not require any operation with a computational cost of $\mathcal{O}(N^2)$ or larger. Of course, the efficiency could still be improved by evaluating the above terms only on the essential support of the probability distribution. The essential support (i.e., all states where $p_n(x)$ is larger than some threshold) can be easily determined from the wavelet representation on a coarse scale.

5. Numerical examples. In this section the performance of the adaptive wavelet method is illustrated by several numerical examples. All computations were performed on a laptop equipped with Intel® Core™ 2 Duo 2.2 GHz processor and 2 GB of RAM. The adaptive wavelet method was implemented in MATLAB®; only the fast wavelet transform was coded in C.

5.1. Merging modes. Consider two species S_1 and S_2 which interact via the reaction channels



with rate constants

$$c_1 = 2, \quad c_2 = 1, \quad c_3 = 1, \quad c_4 = 0.5.$$

The purpose of this simple example is to verify the error bound proven in Theorem 1. This is possible because the *exact* solution of the CME corresponding to (5.1) is known: all reactions are *monomolecular*, and for such systems an explicit solution formula has been derived in [22].

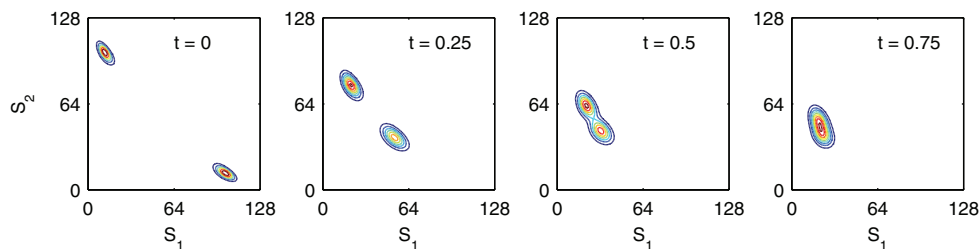


FIG. 5.1. Exact solution (5.3) of the reaction system (5.1) at $t = 0$, $t = 0.25$, $t = 0.5$, and $t = 0.75$ (from left to right).

For any $x \in \mathbb{N}^2$, $N \in \mathbb{N}$, and any $r = (r_1, r_2)$ with $r_1, r_2 \in [0, 1]$ and $r_1 + r_2 \leq 1$, the multinomial distribution $\mathcal{M}(x, N, r)$ is defined by

$$\mathcal{M}(x, N, r) = \begin{cases} N! \frac{r_1^{x_1} r_2^{x_2} (1 - r_1 - r_2)^{N - x_1 - x_2}}{x_1! x_2! (N - x_1 - x_2)!} & \text{if } x_1 + x_2 \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

\mathcal{M} is a two-dimensional extension of the well-known binomial distribution. For the initial distribution we choose

$$(5.2) \quad \rho(x) = 0.5 \cdot \mathcal{M}(x, N, r^{(1)}) + 0.5 \cdot \mathcal{M}(x, N, r^{(2)})$$

with $r^{(1)} = (0.8, 0.1)^T$, $r^{(2)} = (0.1, 0.8)^T$, and $N = 127$. Then, the exact solution of the corresponding CME is

$$(5.3) \quad p(t, x) = 0.5 \cdot \mathcal{M}(x, N, s^{(1)}(t)) + 0.5 \cdot \mathcal{M}(x, N, s^{(2)}(t)),$$

with

$$s^{(i)}(t) = \exp(t\mathcal{C})r^{(i)}, \quad \mathcal{C} = \begin{pmatrix} -(c_1 + c_3) & c_2 \\ c_1 & -(c_2 + c_4) \end{pmatrix}$$

(cf. [22]). Figure 5.1 shows that $p(t, x)$ consists of two modes which merge to one single peak as time evolves. Since the probability distribution is defined only at discrete points $x \in \mathbb{N}^d$, it is actually not correct to visualize it in a mesh or contour plot. Such plots, however, give a much clearer idea of the solution profile than other ways of visualization.

The adaptive wavelet method was applied to this problem with several different step sizes h and tolerances $\text{tol} = h^3$. Of course, our method did not take any advantage of the exact solution formula. The left panel in Figure 5.2 shows that the error with respect to $\|\cdot\|_1$ decays approximately like $\mathcal{O}(h^2)$ as predicted by Theorem 1. As expected, the constant $\|\mathcal{A}^3 \rho\|_1 \approx 24184$ appearing in the error bound (4.12) is too pessimistic; cf. the remark after Theorem 1.

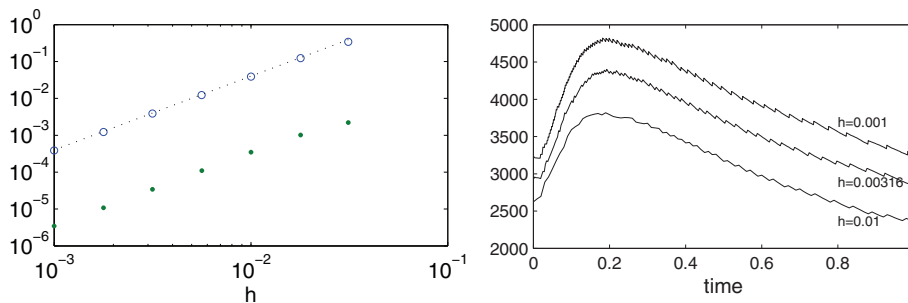


FIG. 5.2. Left: Errors of the adaptive wavelet method applied to the problem (5.1), (5.2) for different step sizes h and with tolerance $\text{tol} = h^3$. The errors were measured by the expressions $\max_{n=1, \dots, n_{\max}} \|p(t_n) - p_n\|_1$ (circles) and $\max_{n=1, \dots, n_{\max}} \|p(t_n) - p_n\|_\infty$ (dots). The function $h \mapsto 400 \cdot h^2$ (dotted line) is included for comparison. As predicted by Theorem 1, the global error of the adaptive wavelet method behaves like $\mathcal{O}(h^2)$ if $\text{tol} = \mathcal{O}(h^3)$. Right: Number of used basis elements as a function of time for step sizes $h = 10^{-2}, 10^{-2.5}, 10^{-3}$ and tolerances $\text{tol} = h^3$.

The adaptive wavelet method only uses as many degrees of freedom as necessary to represent and propagate the approximation. The right panel in Figure 5.2 illustrates how the number of degrees of freedom varies adaptively in time. The three curves correspond to the step sizes $h = 10^{-2}, 10^{-2.5}, 10^{-3}$ and tolerances $\tau_{ol} = h^3$. Smaller step sizes and tolerances imply higher accuracy, but require more basis elements. It is interesting to see that the number of basis elements first increases, but then *decreases* when the two modes approach each other and finally merge to one single mode.

5.2. Infectious diseases. The SEIS model (susceptible-exposed-infectious-susceptible) is widely used to describe the spread of an infectious disease within a population; cf. [21] and references therein. The population is divided into three classes ($d = 3$), namely, individuals susceptible to infection (S), infectious individuals (I), and exposed individuals (E) in the latent period who are infected but not yet infectious. The subpopulations interact via the following reactions:

$$(5.4) \quad \begin{array}{l} R_1 : \\ R_2 : \\ R_3 : \\ R_4 : \\ R_5 : \\ R_6 : \\ R_7 : \end{array} \begin{array}{l} S + I \longrightarrow E + I \\ E \longrightarrow I \\ I \longrightarrow S \\ S \longrightarrow \star \\ E \longrightarrow \star \\ I \longrightarrow \star \\ \star \longrightarrow S \end{array} \quad \left| \quad \begin{array}{l} \alpha_1 = c_1 x_1 x_3 \\ \alpha_2 = c_2 x_2 \\ \alpha_3 = c_3 x_3 \\ \alpha_4 = c_4 x_1 \\ \alpha_5 = c_5 x_2 \\ \alpha_6 = c_6 x_3 \\ \alpha_7 = c_7 \end{array} \quad \left| \quad \begin{array}{l} \nu_1 = (-1, 1, 0)^T \\ \nu_2 = (0, -1, 1)^T \\ \nu_3 = (1, 0, -1)^T \\ \nu_4 = (-1, 0, 0)^T \\ \nu_5 = (0, -1, 0)^T \\ \nu_6 = (0, 0, -1)^T \\ \nu_7 = (1, 0, 0)^T. \end{array} \right.$$

R_1 models the infection of susceptible individuals by infected ones. The infected person first enters the E class, but can become infectious after the latent period via reaction R_2 . Infectious individuals can recover from the disease via R_3 . $R_4, R_5,$ and R_6 describe the death of individuals, whereas R_7 represents the birth or the arrival of new susceptible individuals. It is assumed that this inflow is constant and does not depend on the current size of the population.

We consider a stochastic variant of the model, which accounts for the fact that an infection may start with only a few infected individuals. In this situation, the question whether the disease spreads over large parts of the population or disappears early depends on the fate of the first few infectious individuals. In our example, the parameters are

$$c_1 = 0.1, \quad c_2 = 0.5, \quad c_3 = 1, \quad c_4 = c_5 = c_6 = 0.01, \quad c_7 = 0.4,$$

and the initial distribution is a delta peak at $(50, 0, 2)$.

Figure 5.3 shows the time evolution of the probability distribution. Since the full distribution is a three-dimensional object, we depict mesh and contour plots of the two-dimensional marginal distribution in the S - E -plane at different times ($t = 1, 3, 4, 5, 7$ from top to bottom). The panels in the left and middle columns show the approximation given by the adaptive wavelet method. The results are compared with the distribution obtained with the stochastic simulation algorithm from [16] (right column).

Figure 5.3 shows that the solution splits up into two distinct peaks. The peak located at $(50, 0)$ indicates that with a certain probability the first few infectious individuals die or recover before a critical number of susceptibles has been infected such that the disease disappears after some time. However, if the infection spreads fast enough in the initial phase, then the system eventually reaches a state where most of the population is infected, as indicated by the second peak located at $(11, 27)$. Similar bimodalities appear in the marginal distributions of the S - I -plane and the E - I -plane (data not shown).

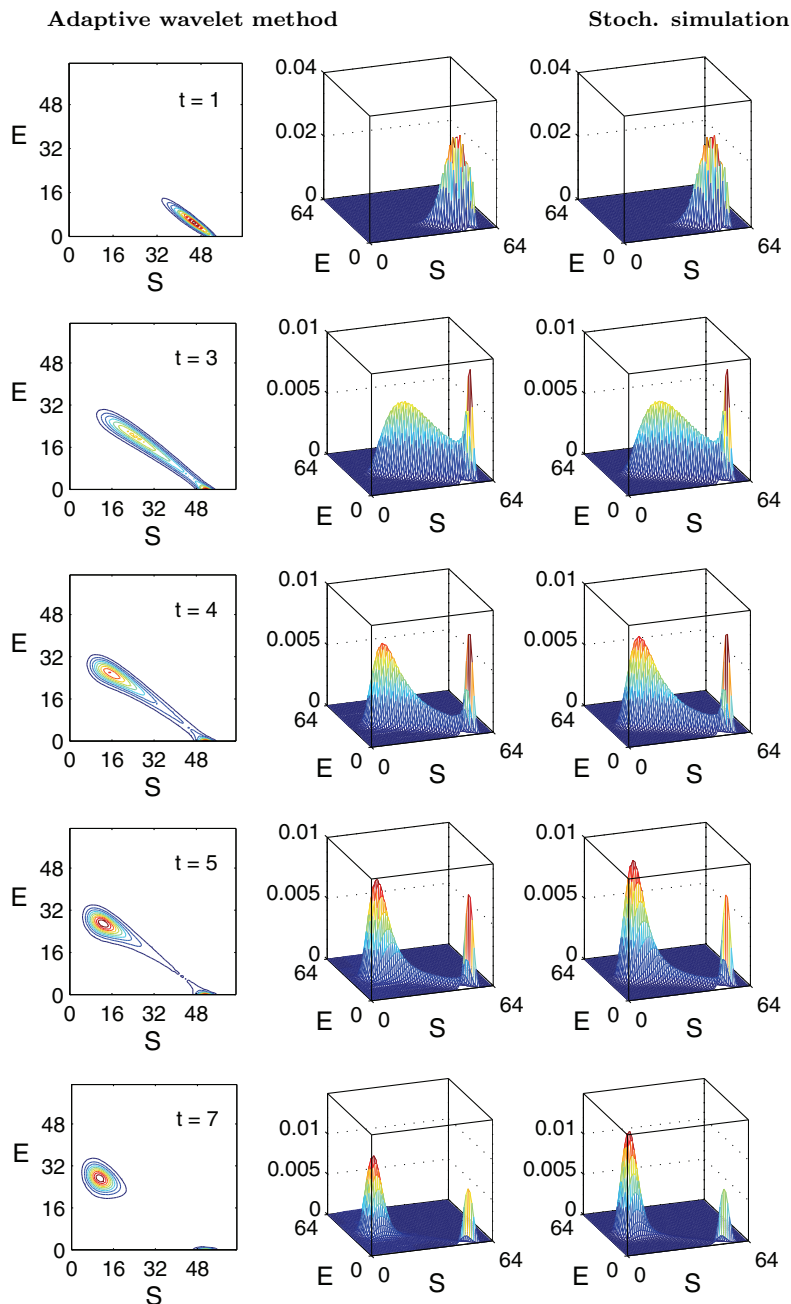


FIG. 5.3. Marginal distribution of the stochastic SEIS model (5.4) at times $t = 1, 3, 4, 5, 7$ (from top to bottom). Contour plot (first column) and mesh plot (second column) of the approximation obtained with the adaptive Haar wavelet method. Third column: Approximation computed with the stochastic simulation algorithm.

From the numerical perspective, the problem is challenging for the following reasons:

1. The size of the state space is $64 \times 64 \times 32$. Many of the $2^{17} = 131,072$ states are never populated throughout the time evolution such that the number of “essential” states is smaller. However, the information which states can be

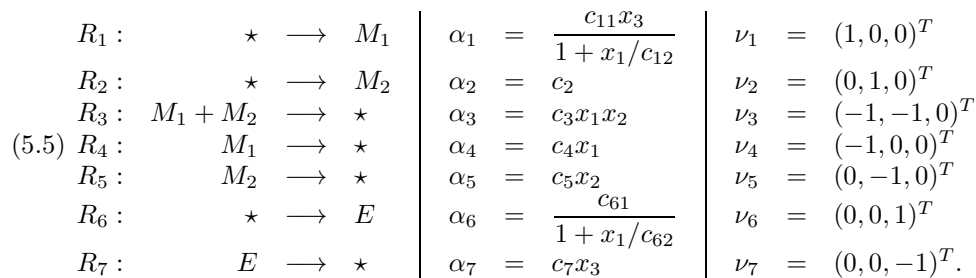
ignored is not known a priori in practical applications. The adaptive wavelet method finds out this information by itself.

2. At $t = 7$ the solution vanishes *close to* the S -axis, but does not vanish *on* that axis where one of the two peaks is located (see the panels in the last row of Figure 5.3). This local nonsmoothness poses difficulties to methods which assume a certain regularity of the solution. Some of these methods would even fail to represent the highly nonsmooth initial distribution (a delta peak) correctly.
3. In the initial phase, the problem is stiff.
4. The time evolution is dominated by transport rather than by diffusion. This is a disadvantage for our method because as the support of the probability distribution moves and grows, more and more new basis elements have to be selected.
5. Between $t = 3$ and $t = 4$ the solution profile is confined to a rather thin line which is not parallel to any of the axes. Methods which represent the solution in terms of global tensor products (e.g., the method from [22]) would need many degrees of freedom to achieve an acceptable accuracy. The adaptive wavelet method does not suffer from such problems, because the elements of the Haar basis are *local* tensor products with small support.
6. The fact that the distance between the two peaks of the bimodality is rather large poses problems to all methods which rely on the assumption that the solution remains more or less unimodal.

In this example, the adaptive wavelet method was used with step size $h = 0.05$. The iteration for solving the linear system (cf. subsection 4.6) was stopped if the tolerance $\text{tol} = 0.001$ was met or if the total number of basis elements exceeded 3600. At the end of each step, only the 3000 largest coefficients were kept, which corresponds to 2.3% of the total number of degrees of freedom. The total runtime was 26 min, 7 sec. The approximation in the third column of Figure 5.3 is based on 10,000,000 runs of the stochastic simulation algorithm, which took more than 47.5 h.

Figure 5.3 shows that the results of the adaptive wavelet method agrees well with the stochastic simulation. Only toward the end of the time interval do the two approximations differ with respect to the height of the second peak. The reason is that some part of the probability mass is spread over several states due to numerical errors in the coefficients of basis elements with large support. Of course, this effect vanishes if the approximation parameters (tolerances, step size, number of basis elements) are refined. At $t = 7$, the maximal difference between the two approximations was $5.52 \cdot 10^{-4}$. It should be pointed out, however, that the estimated accuracy of the stochastic simulation is only $4.23 \cdot 10^{-5}$. This value was obtained by computing the 90% confidence interval at $t = 7$ according to subsection 1.9 in [25].

5.3. Two metabolites and one enzyme. As a second example we consider the interaction of two metabolites and one enzyme in a model discussed in [9, 29]. The reactions are



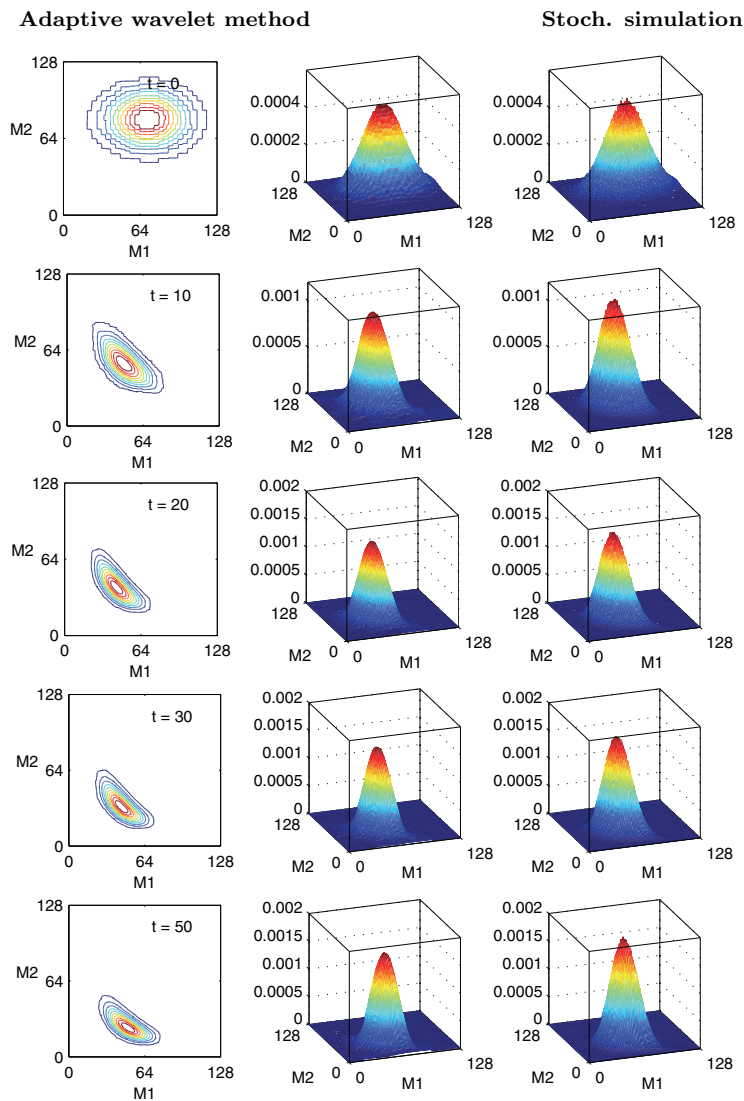


FIG. 5.4. Marginal distribution of the system (5.5) at times 0, 10, 20, 30, and 50 (from top to bottom). Contour plot (first column) and mesh plot (second column) of the approximation obtained with the adaptive Haar wavelet method. The third column shows the approximation computed with the stochastic simulation algorithm.

The parameters

$$c_{11} = 0.3, \quad c_{12} = 60, \quad c_2 = 1, \quad c_3 = 0.001, \quad c_4 = c_5 = c_7 = 0.002, \quad c_{61} = 0.02, \quad c_{62} = 30$$

were taken from [29]. In this example, the size of the state space is $128 \times 128 \times 32$ ($d = 3$). In order to demonstrate that the adaptive wavelet method can also be applied when the essential support of the solution is large, a discrete Gaussian with large variance was chosen as initial distribution. The adaptive wavelet method was configured in such a way that in each time-step 500 new basis elements were proposed and only the basis elements corresponding to the 4000 largest coefficients were stored. Hence, the solution was approximated with only 0.76% of the total

number of 524,288 degrees of freedom. Solutions were computed on the time interval $[0, 50]$ with step size $h = 0.5$, which took 20 min, 36 sec. In the stochastic simulation, 2,500,000 realizations were produced, which took more than 13 h.

The time evolution of the solution is shown in Figure 5.4. In spite of the rather sparse spatial representation, our method agrees very well with the solution obtained by stochastic simulation (third column). At $t = 50$ the maximal difference was $1.73 \cdot 10^{-4}$. The estimated accuracy of the stochastic simulation in terms of the 90% confidence interval is $2.43 \cdot 10^{-5}$.

6. Conclusion and possible extensions. In this work a new method for solving chemical master equations is proposed. The solution is represented in a small subset of the Haar basis, and the essential basis elements are determined adaptively in each time-step by combining Rothe's method with an iterative strategy.

The new method reduces the huge number of degrees of freedom considerably. However, we expect that the efficiency can still be largely improved by the following refinements. First, the Haar wavelet can be replaced by discrete wavelets of higher order. Such wavelets would allow even better compression rates thanks to their better approximation properties. Second, the trapezoidal rule for the time integration can be replaced by a higher order integrator or by a variable order strategy. Third, the time integration can be carried out with an adaptive step size. This is attractive because the stiffness in the initial phase imposes small time-steps, whereas much larger time-steps are possible when the fast modes of the solution have disappeared. Work on these refinements is currently in progress.

Acknowledgment. The Research Group "Numerical methods for high-dimensional systems" received financial support by the "Concept for the Future" of Karlsruhe Institute of Technology within the framework of the German Excellence Initiative. Moreover, support by the DFG priority programme SPP 1324 "Mathematische Methoden zur Extraktion quantifizierbarer Information aus komplexen Systemen" is acknowledged. The author thanks Steffen Galan for writing parts of the Matlab code, Christian Wieners for his valuable suggestions and comments, Tudor Udrescu for a thorough reading of the manuscript, and the two anonymous referees for their helpful remarks.

REFERENCES

- [1] F. BORNEMANN, *An adaptive multilevel approach to parabolic equations: I. General theory and 1d implementation*, IMPACT Comput. Sci. Eng., 2 (1990), pp. 279–317.
- [2] F. BORNEMANN, *An adaptive multilevel approach to parabolic equations: II. Variable-order time discretization based on a multiplicative error correction*, IMPACT Comput. Sci. Eng., 3 (1991), pp. 93–122.
- [3] K. BURRAGE, M. HEGLAND, S. MACNAMARA, AND R. B. SIDJE, *A Krylov-based finite state projection algorithm for solving the chemical master equation arising in the discrete modelling of biological systems*, in Markov Anniversary Meeting: An international conference to celebrate the 150th anniversary of the birth of A. A. Markov. A. N. Langville and W. J. Stewart, eds., Bosen Books, Raleigh, NC, 2006, pp. 21–38.
- [4] A. COHEN, *Numerical Analysis of wavelet methods*, Volume 32 of Studies in Mathematics and its applications, Elsevier, New York, 2003.
- [5] W. DAHMEN, *Wavelets and multiscale methods for operator equations*, Acta Numer., 6 (1997), pp. 55–228.
- [6] W. DAHMEN, *Wavelet methods for PDEs—some recent developments*, J. Comput. Appl. Math., 128 (2001), pp. 133–185.
- [7] P. DEUFLHARD, W. HUISINGA, T. JAHNKE, AND M. WULKOW, *Adaptive discrete Galerkin methods applied to the chemical master equation*, SIAM J. Sci. Comput., 30 (2008), pp. 2990–3011.

- [8] P. DEUFLHARD AND M. WULKOW, *Computational treatment of polyreaction kinetics by orthogonal polynomials of a discrete variable*, IMPACT Comput. Sci. Eng., 1 (1989), pp. 269–301.
- [9] J. ELF, O. BERG, AND M. EHRENBORG, *Comparison of repressor and transcriptional attenuator systems for control of amino acid biosynthetic operons*, J. Mol. Biol., 313 (2001), pp. 941–954.
- [10] S. ENGBLOM, *Spectral approximation of solutions to the chemical master equation*, J. Comput. Appl. Math., 229 (2009), pp. 208–221.
- [11] S. ENGBLOM, *Galerkin spectral method applied to the chemical master equation*, Commun. Comput. Phys., 5 (2009), pp. 871–896.
- [12] K.-J. ENGEL AND R. NAGEL, *One-parameter semigroups for linear evolution equations*, Number 194 in Grad. Texts in Math., Springer, Berlin, 2000.
- [13] L. FERM AND P. LÖTSTEDT, *Adaptive solution of the master equation in low dimensions*, Appl. Numer. Math., 59 (2009), pp. 187–204.
- [14] L. FERM, P. LÖTSTEDT, AND A. HELLANDER, *A hierarchy of approximations of the master equation scaled by a size parameter*, SIAM J. Sci. Comput., 34 (2008), pp. 127–151.
- [15] C. W. GARDINER, *Handbook of Stochastic Methods*, 3rd ed., Springer, New York, 2004.
- [16] D. T. GILLESPIE, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, J. Comput. Phys., 22 (1976), pp. 403–434.
- [17] D. T. GILLESPIE, *A rigorous derivation of the chemical master equation*, Physica A, 188 (1992), pp. 404–425.
- [18] M. HEGLAND, C. BURDEN, L. SANTOSO, S. MACNAMARA, AND H. BOOTH, *A solver for the stochastic master equation applied to gene regulatory networks*, J. Comput. Appl. Math., 205 (2007), pp. 708–724.
- [19] M. HEGLAND, A. HELLANDER, AND P. LÖTSTEDT, *Sparse grids and hybrid methods for the chemical master equation*, BIT, 48 (2008), pp. 265–284.
- [20] A. HELLANDER AND P. LÖTSTEDT, *Hybrid method for the chemical master equation*, J. Comput. Phys., 227 (2007), pp. 100–122.
- [21] H. W. HETHCOTE, *The mathematics of infectious diseases*, SIAM Rev., 42 (2000), pp. 599–653.
- [22] T. JAHNKE AND W. HUISINGA, *Solving the chemical master equation for monomolecular reaction systems analytically*, J. Math. Biol., 54 (2007), pp. 1–26.
- [23] T. JAHNKE AND W. HUISINGA, *A dynamical low-rank approach to the chemical master equation*, Bull. Math. Biol., 70 (2008), pp. 2283–2302.
- [24] T. JAHNKE AND S. GALAN, *Solving chemical master equations by an adaptive wavelet method*, in Numerical analysis and applied mathematics: International Conference on Numerical Analysis and Applied Mathematics 2008, Psalidi, Kos, Greece, 2008, volume 1048 of AIP Conference Proceedings, T. E. Simos, G. Psihoyios, and C. Tsitouras, eds., 2008, AIP, New York, pp. 290–293.
- [25] P. E. KLOEDEN AND E. PLATEN, *Numerical solution of stochastic differential equations*, Number 23 in Applications of Mathematics, Springer, New York, 1992.
- [26] S. MACNAMARA, K. BURRAGE, AND R. B. SIDJE, *Multiscale modeling of chemical kinetics via the master equation*, Multiscale Model. Simul., 6 (2008), pp. 1146–1168.
- [27] B. MUNSKY AND M. KHAMMASH, *The finite state projection algorithm for the solution of the chemical master equation*, J. Chem. Phys., 124 (2006), p. 044104.
- [28] Y. NIEVERGELT, *Wavelets made easy*, Birkhäuser, Boston, MA, 1999.
- [29] P. SJÖBERG, P. LÖTSTEDT, AND J. ELF, *Fokker-Planck approximation of the master equation in molecular biology*, Comput. Vis. Sci., 12 (2009), pp. 37–50.
- [30] CH. SCHWAB AND R. STEVENSON, *Space-time adaptive wavelet methods for parabolic evolution problems*, Math. Comp., 78 (2009), pp. 1293–1318.
- [31] T. von Petersdorff and Ch. Schwab, *Numerical solution of parabolic equations in high dimensions*, M2AN Math. Model. Numer. Anal., 38 (2004), pp. 93–127.
- [32] M. WULKOW, *The simulation of molecular weight distribution in polyreaction kinetics by discrete Galerkin methods*, Macromol. Theory Simul., 5 (1996), pp. 393–416.

IWRMM-Preprints seit 2008

- Nr. 08/01 Patrizio Neff, Antje Sydow, Christian Wieners: Numerical approximation of incremental infinitesimal gradient plasticity
- Nr. 08/02 Götz Alefeld, Zhengyu Wang: Error Estimation for Nonlinear Complementarity Problems via Linear Systems with Interval Data
- Nr. 08/03 Ulrich Kulisch : Complete Interval Arithmetic and its Implementation on the Computer
- Nr. 08/04 Armin Lechleiter, Andreas Rieder: Newton Regularizations for Impedance Tomography: Convergence by Local Injectivity
- Nr. 08/05 Vu Hoang, Michael Plum, Christian Wieners: A computer-assisted proof for photonic band gaps
- Nr. 08/06 Vincent Heuveline, Peter Wittwer: Adaptive boundary conditions for exterior stationary flows in three dimensions
- Nr. 08/07 Jan Mayer: Parallel Algorithms for Solving Linear Systems with Sparse Triangular Matrices
- Nr. 08/08 Ulrich Kulisch : Begegnungen eines Mathematikers mit Informatik
- Nr. 08/09 Tomas Dohnal, Michael Plum, Wolfgang Reichel: Localized Modes of the Linear Periodic Schrödinger Operator with a Nonlocal Perturbation
- Nr. 08/10 Götz Alefeld: Verified Numerical Computation for Nonlinear Equations
- Nr. 08/11 Götz Alefeld, Zhengyu Wang: Error bounds for complementarity problems with tri-diagonal nonlinear functions
- Nr. 08/12 Tomas Dohnal, Hannes Uecker: Coupled Mode Equations and Gap Solitons for the 2D Gross-Pitaevskii equation with a non-separable periodic potential
- Nr. 09/01 Armin Lechleiter, Andreas Rieder: Towards A General Convergence Theory For Inexact Newton Regularizations
- Nr. 09/02 Christian Wieners: A geometric data structure for parallel finite elements and the application to multigrid methods with block smoothing
- Nr. 09/03 Arne Schneck: Constrained Hardy Space Approximation
- Nr. 09/04 Arne Schneck: Constrained Hardy Space Approximation II: Numerics
- Nr. 10/01 Ulrich Kulisch, Van Snyder : The Exact Dot Product As Basic Tool For Long Interval Arithmetic
- Nr. 10/02 Tobias Jahnke : An Adaptive Wavelet Method for The Chemical Master Equation

Eine aktuelle Liste aller IWRMM-Preprints finden Sie auf:

www.mathematik.uni-karlsruhe.de/iwrmm/seite/preprints

Kontakt

Karlsruher Institut für Technologie (KIT)
Institut für Wissenschaftliches Rechnen
und Mathematische Modellbildung

Prof. Dr. Christian Wieners
Geschäftsführender Direktor

Campus Süd
Engesserstr. 6
76131 Karlsruhe

E-Mail: imm-sekretariat@math.kit.edu

[.math.kit.edu /imm/](http://math.kit.edu/imm/)

Herausgeber

Karlsruher Institut für Technologie (KIT)
Kaiserstraße 12 76131 Karlsruhe

Februar 2010

.kit.edu