HHS Public Access

Author manuscript

J Comput Phys. Author manuscript; available in PMC 2016 February 08.

Published in final edited form as:

J Comput Phys. 2014 June 1; 266: 89–100. doi:10.1016/j.jcp.2014.02.004.

Local error estimates for adaptive simulation of the Reaction–Diffusion Master Equation via operator splitting

Andreas Hellander[1,*], **Michael J Lawson**[2], **Brian Drawert**[2], and **Linda Petzold**[2]

[1]Department of Information Technology, Uppsala University, Box 337, SE-75105, Uppsala, Sweden

[2]Department of Computer Science, University of California Santa Barbara, Santa Barbara, CA 93106-5070, USA

Abstract

The efficiency of exact simulation methods for the reaction-diffusion master equation (RDME) is severely limited by the large number of diffusion events if the mesh is fine or if diffusion constants are large. Furthermore, inherent properties of exact kinetic-Monte Carlo simulation methods limit the efficiency of parallel implementations. Several approximate and hybrid methods have appeared that enable more efficient simulation of the RDME. A common feature to most of them is that they rely on splitting the system into its reaction and diffusion parts and updating them sequentially over a discrete timestep. This use of operator splitting enables more efficient simulation but it comes at the price of a temporal discretization error that depends on the size of the timestep. So far, existing methods have not attempted to estimate or control this error in a systematic manner. This makes the solvers hard to use for practitioners since they must guess an appropriate timestep. It also makes the solvers potentially less efficient than if the timesteps are adapted to control the error. Here, we derive estimates of the local error and propose a strategy to adaptively select the timestep when the RDME is simulated via a first order operator splitting. While the strategy is general and applicable to a wide range of approximate and hybrid methods, we exemplify it here by extending a previously published approximate method, the Diffusive Finite-State Projection (DFSP) method, to incorporate temporal adaptivity.

1 Introduction

To understand biological systems on the cellular level, it is often essential to account for the impact of noise due to small molecule count. For example, it has been demonstrated that stochasticity can have a profound effect on gene regulatory systems [23, 8]. Spatial distribution of molecules in a cell can result in locally small populations of key chemical species, such that noise drives essential behavior, as in the case of symmetry breaking across many eukaryotic cell types [30]. Spatial stochastic modeling has already begun to yield new

*To whom correspondence should be addressed. andreas.hellander@it.uu.se.

insights in systems such as spatiotemporal oscillators [10, 27, 26], MAPK signaling [28], self-organization of proteins into clusters [7] and polarization of proteins on the cell membrane in yeast [1]. Several modeling frameworks are used to model spatial stochastic system, the two most commonly used in systems biology being continuous space Brownian Dynamics (BD) methods exemplified by GFRD [31] and the mesoscopic Reaction-Diffusion Master Equation (RDME), the latter being the focus of this paper.

In the traditional RDME, space is subdivided into subvolumes that can individually be treated as well-mixed. Reactions within a subvolume are expressed in the form of the chemical master equation (CME) [13] and realizations of the process can be generated using Gillespie's stochastic simulation algorithm (SSA) [12]. Molecules can move freely between neighboring voxels via diffusive jumps, which are modeled as linear jump events in a Markov process. Optimized exact simulation methods such as the Next Subvolume Method (NSM) [7] can be used to generate statistically correct realizations of the RDME. As with all exact methods applied to RDME models, the NSM suffers from a potentially high computational cost due to having to explicitly simulate each diffusion event. The number of diffusive transfers between voxels grows rapidly as the mesh resolution is made finer, and as a result the majority of computation time tends to be spent on sampling diffusion events. Additionally, these methods are also inherently serial, which has thwarted attempts to increase efficiency via parallelization.

To speed up simulation of the RDME many methods rely on operator splitting. By splitting the operators, most often with a Lie-Trotter scheme [29], the reaction and diffusion steps can be solved independently. While diffusion carries the bulk of the computational burden in exact solvers, it is often possible to take advantage of the structure and linear nature of the discretized diffusion equation to speed up this step in an operator-split solver. Examples of approximate methods that have been proposed to speed up the simulation of the RDME by reducing the cost of the diffusive step include methods based on tau-leaping [25, 21], the multinomial simulation algorithm [19], spatially adaptive hybrid methods [11] and the diffusive finite state projection method (DFSP) [6]. While splitting in itself does not resolve the issue of the inherent stiffness of the diffusion operator, the continued introduction of methods for simulating the RDME via operator splitting highlights the potential of this approach. Another recent use of operator splitting in the RDME context is the use of Lie-Trotter splitting to simulate fractional diffusion [2]. Yet another advantage of splitting is that it converts a largely serial problem, which is known to be difficult to parallelize in an efficient manner, into a naturally parallelizable one. For the existing approximate and hybrid methods for the RDME, splitting the physics (reaction and diffusion) is necessary. For parallel implementations, another possibility is to split the computational grid into blocks, as proposed by [14], where the error introduced by operator splitting at block boundaries was analyzed. Our analysis here is different since it applies to the case of splitting the reaction and diffusion operators.

Splitting the operators introduces an error that depends on the size of the splitting time step, however previous algorithms that rely on operator splitting have not attempted an *a priori* error estimator. Without such an estimator these methods have no way to automatically control the splitting error. This limits their usefulness for practitioners. From an efficiency

point of view, not knowing and controlling the error might lead to the use of unnecessary small timesteps at the price of slower simulations. In this work we seek to address theses issues by presenting estimators of the local error in probability, mean and variance for a first order Lie-Trotter splitting of the RDME. The estimators allow control of the splitting error for spatial stochastic simulation and enables any method based on operator splitting to be implemented adaptively.

This paper is organized as follows: in Section 2 we introduce the RDME. In Section 3 we outline how to simulate the RDME using operator splitting. We derive our estimator for the local error in the PDF, mean and variance, and demonstrate the accuracy of the estimator for an example problem. Finally, in Section 4 we present how the local error estimates can be used to extend an approximate method for the RDME to incorporate temporal adaptivity.

2 Spatial Stochastic Simulation using the RDME

Given a system with N_s chemical species X_s reacting in a physical domain Ω, discretize Ω into N_v non-overlapping voxels $\mathscr{V}_i$, with volume $|\mathscr{V}_i|$, and let the state of the system be described by the $(N_v \times N_s)$ state matrix $\mathbf{x}$, where the element x_{is} is the copy number of species X_s in voxel $\mathscr{V}_i$. Let $\mathbf{x}_{i,\cdot}$ denote the i-th row of $\mathbf{x}$ and $\mathbf{x}_{\cdot,s}$ the s-th column. The reaction network consists of N_r chemical reactions $r = 1 \ldots N_r$. The $(1 \times N_s)$ stoichiometry vector $\mathbf{n}_{ir}$ describes the change in the state, $\tilde{\mathbf{x}}_{i,\cdot}=\mathbf{x}_i+\boldsymbol{n}_{ir}$, when reaction r occurs in voxel i and the propensity function for that reaction is $a_{ir}(\mathbf{x}_{i,\cdot})$.

Diffusion of species X_s along the edge (2D) or face (3D) connecting voxels $\mathscr{V}_i$ and $\mathscr{V}_j$ is modeled as a linear jump event, or diffusive transfer,

$$X_{is} \xrightarrow{d_{ijs}} X_{js}, \quad (1)$$

with propensity function $\mu(\mathbf{x}) = d_{ijs}x_{is}$. The change in state due to the diffusive transfer is described by the $(N_v \times 1)$ stoichiometry vector ν_{ijs}, such that the new state is $\tilde{\mathbf{x}}_{\cdot,s}=\mathbf{x}_{\cdot,s}+\nu_{ijs}$. All entries of ν_{ijk} are zero except $\nu_{ijs}(i) = -1$ and $\nu_{ijs}(j) = 1$. In the case of reactions only, the probability density function $p(\mathbf{x}, t) \equiv p(\mathbf{x}, t|\mathbf{x}_0, 0)$ obeys the master equation

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}p(\mathbf{x},t)=\mathscr{M}p(\mathbf{x},t) \equiv \\ \sum_{i=1}^{N_v}\sum_{r=1}^{N_r}a_{ir}(\mathbf{x}_{i\cdot}-\boldsymbol{n}_{ir})p(\mathbf{x}_{1\cdot},\ldots,\mathbf{x}_{i\cdot}-\boldsymbol{n}_{ir},\ldots,\mathbf{x}_{N\cdot},t) \\ -\sum_{i=1}^{N_v}\sum_{r=1}^{N_r}a_{ir}(\mathbf{x}_{i\cdot})p(\mathbf{x},t).\end{aligned} \quad (2)$$

For the case of one subvolume, $N_v = 1$, (2) reduces to the CME for a well stirred system. For a system with only diffusion, the master equation takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}p(\mathbf{x},t)=\mathscr{D}p(\mathbf{x},t)\equiv \sum_{s=1}^{N_s}\sum_{i=1}^{N_v}\sum_{j=1}^{N_v}\mu(\mathbf{x}_{\cdot j}-\nu_{ijs})p(\mathbf{x}_{\cdot 1},\ldots,\mathbf{x}_{\cdot j}-\nu_{ijs},\ldots,\mathbf{x}_{\cdot N},t) -\sum_{s=1}^{N_s}\sum_{i=1}^{N_v}\sum_{k=1}^{N_v}\mu(\mathbf{x}_{\cdot j})p(\mathbf{x},t). \quad (3)$$

For a system with both diffusion and reactions, the full RDME is then given by

$$\frac{d}{dt}p(\mathbf{x},t)=(\mathscr{M}+\mathscr{D})p(\mathbf{x},t). \quad (4)$$

The values of the diffusion rate constants d_{ijs} depend on the diffusion constant γ_s of species X_s and the shapes and sizes of voxels $\mathscr{V}_i$ and $\mathscr{V}_j$. Let $u_s(\zeta, t) = E[X_s/|\Omega|]$, i.e. the the expected value of the concentration process corresponding to (3). In the thermodynamic limit $X_s/|\Omega| \to \infty$, u_s is governed by the diffusion equation

$$\frac{\partial}{\partial t}u_s(\zeta,t)=\Delta u_s(\zeta,t). \quad (5)$$

Here, ζ is a spatial coordinate in a Cartesian coordinate system. Similarly, the Fokker-Planck equation that describes the time evolution of the probability density of a single particle undergoing Brownian motion is given by

$$\frac{\partial}{\partial t}p(\zeta,t)=\Delta p(\zeta,t). \quad (6)$$

Hence, by choosing d_{ijs} according to a consistent spatial discretization of the Laplace operator Δ, we obtain mesoscopic jump coefficients that are motivated both from a microscale and macroscale perspective. To comply with the description of the RDME, the discretization used needs to couple only nearest neighbors in the mesh. For a uniform, Cartesian mesh such as shown in Figure 1 (a), the most natural choice is a centered finite difference discretization, giving $d_{ijs} = \gamma_s/h^2$ where h is the side length of the voxels and γ_s is the diffusion constant of species s. For unstructured, tetrahedral meshes which will be used later in this paper, d_{ijs} can be obtained from a finite element (FE) discretization using linear Lagrange elements. For a detailed description of that methodology, see [9]. Figure 1 (b) shows a triangular, unstructured mesh. The local volume where the molecules are assumed to be well mixed are given by the dual mesh.

3 Operator splitting method for the RDME

In this section we describe how an operator splitting method can be applied to generate realizations from the RDME with a controlled temporal discretization error.

In the remainder of this paper we will tacitly assume that we can impose a bound on the state space. For every species in the system we assume that $P(X_{is} > x_{max}, t) = 0$, for all t. The state space is then finite (but very large). Technically, this will not necessarily hold true

for an open system, but from a biophysical perspective it is a reasonable assumption since a system would require infinite energy to blow up. In the finite case, the operators in the RDME can be represented by finite matrices. Hereafter we will use the notation $\mathscr{M}$ and $\mathscr{D}$ to mean representations of the bounded operators resulting from the above truncation of the state space. The solution of (4) can then simply be written

$$p(\mathbf{x},t)=e^{t(\mathscr{M}+\mathscr{D})}p(\mathbf{x},0). \quad (7)$$

We point out that even though this representation of the solution is simple, it is not feasible to solve (4) this way since the state state space is too large except for trivial models and discretizations. In the well mixed case, however, deterministic methods to solve the CME have been developed for small to medium-sized models, see for example [24, 3, 20, 18, 22].

Using a first order splitting method (Lie-Trotter splitting), $p(\mathbf{x}, t + \Delta t)$ can be approximated by

$$p_s(\mathbf{x},t+\Delta t)e^{\Delta t\mathscr{M}}e^{\Delta t\mathscr{D}}p(\mathbf{x},t). \quad (8)$$

Simulation based on (8) proceeds in two half-steps, with the diffusion operator and the reaction operator acting sequentially

$$\begin{aligned} &1.\ p_s^{n+1/2}(\mathbf{x},t)=e^{\Delta t\mathscr{D}}p_s^n(\mathbf{x},t)\\ &2.\ p_s^{n+1}(\mathbf{x},t+\Delta t)=e^{\Delta t\mathscr{M}}p_s^{n+1/2}. \end{aligned} \quad (9)$$

Provided that the time step Δt is sufficiently small, the local error in the scheme (9) is $\mathscr{O}(\Delta t^2)$ and from standard theory for numerical solution of differential equations, the global error in the PDF is then proportional to Δt, i.e.

$$\|p(\mathbf{x},t)-p_s(\mathbf{x},t)\|\leq C\Delta t. \quad (10)$$

We note that it is not in general necessary for the operators to be bounded for the splitting method to result in a global error proportional to Δt. Jahnke and Altintan show in [16] that the global error for Strang splitting applied to the CME is second order under certain assumptions on the chemical reactions. Unfortunately, as Jahnke and Altintan point out, while those conditions can be expected to hold for many systems, they are not easily verified in practice. Error bounds for exponential operator splitting have also been studied in other contexts in e.g. [17].

In the next subsection we derive estimates of the local error. In a subsequent section we will then illustrate how these estimates can be used to enable temporal adaptivity in an approximate method for the RDME.

Local error

Following standard theory for the analysis of scheme (9), the local error in the PDF (i.e. in probability) in the n+1-th timestep is given by

$$\in^{n+1}=(e^{\Delta t(\mathscr{M}+\mathscr{D})}-e^{\Delta t\mathscr{M}}e^{\Delta t\mathscr{D}})p^n=\frac{\Delta t^2}{2}[\mathscr{D},\mathscr{M}]p^n+\mathscr{O}(\Delta t^3). \quad (11)$$

where $[\mathscr{D},\mathscr{M}]=(\mathscr{D}\mathscr{M}-\mathscr{M}\mathscr{D})$ is the commutator of the operators. This follows directly from a series expansion of the left hand side in Δ*t*. Computing $\in^{n+1}$ from (11) is obviously not tractable in general since knowledge of p^n requires the solution of the full RDME. Even if an approximation of p^n were available, the state space is still very large and the cost of evaluating the commutator would be prohibitive. Also, our goal is a strategy to estimate the error during the course of the generation of individual trajectories, whereas an estimation of p^n would require very large ensembles of trajectories. By conditioning on the currently observed state $\mathbf{x}^n$ in timestep *n* of a specific realization, we obtain the following sequence of conditional errors,

$$\mathscr{E}^{n+1}=(\in^{n+1}|X(t^n)=\mathbf{x}^n)=\frac{\Delta t^2}{2}[\mathscr{D},\mathscr{M}]\delta(\mathbf{x}^n)+\mathscr{O}(\Delta t^3), \quad (12)$$

where δ is the Dirac delta function, hence $\delta(\mathbf{x}^n)$ corresponds to $P(X(t^n)=\mathbf{x}^n)=1$. Note that $\mathscr{E}^n$ is a random vector for each *n* with unconditional expectation $E[\mathscr{E}^{n+1}]=\in^{n+1}$.

While there are many possible states that can have a non-zero value after one application of the commutator, hereafter called reachable states, it is possible to obtain simple and computable expressions for $\mathscr{E}^{n+1}$ in (12), due to the sparsity of $\delta(\mathbf{x}^n)$

The $L=1+2N_eN_s+N_vN_r+2N_vN_rN_eN_s$ reachable state $\tilde{\mathbf{x}}_l$ in $[\mathscr{D},\mathscr{M}]\delta(\mathbf{x}^n)$ are $\mathbf{x}^n$, $\mathbf{x}^n+\nu_{ijs}$, $\mathbf{x}^n+\mathbf{n}_{kr}$ and $\mathbf{x}_n+\mathbf{n}_{kr}+\nu_{ijs}$, $s=1\ldots N_s$, $r=1\ldots N_r$, $i,j,k=1,\ldots,N_v$ and N_e is the number of connections in the mesh. Below, the term to the left of the colon is a reachable state and the term to the right is the value of that state after applying the indicated operator to $\delta(\mathbf{x}^n)$. We note that this value is the corresponding term from the matrix-matrix-vector multiplication that results from applying the reaction and then the diffusion operators (or vice versa) to the previous probability vector, it is not the propensity of the reachable state.

For example, the application of the diffusion operato $\mathscr{D}$ on $\delta(\mathbf{x}_n)$, $\mathscr{D}\delta(\mathbf{x}^n)$ results in

$$\mathbf{x}^n+\nu_{ijs}:d_{ijs}x^n_{is} \quad (13)$$

$$\mathbf{x}^n:-d_0(\mathbf{x}^n), \quad (14)$$

where *i*, *j* are connected subvolumes and $d_0(\mathbf{x}^n)\equiv\sum_{s=1}^{N_s}\sum_{i=1}^{N_v}\sum_{j\neq i}^{N_v}d_{ijs}x^n_{is}$. Equation (13) enumerates the new states with non-zero value after one application of $\mathscr{D}$ to $\delta(\mathbf{x}^n)$.

Applying $\mathscr{M}$ to $\mathscr{D}\delta(\mathbf{x}^n)$ then gives the following reachable states and associated values.

$$\mathbf{x}^n:a_0(\mathbf{x}^n)d_0(\mathbf{x}^n) \quad (15)$$

$$\mathbf{x}^n + \boldsymbol{n}_{kr} : -a_{kr}(\mathbf{x}^n) d_0(\mathbf{x}^n), \quad (16)$$

$$\mathbf{x}^n + \nu_{ijs} : -a_0(\mathbf{x}^n + \nu_{ijs}) d_{ijs} x_{is}^n, \quad (17)$$

$$\mathbf{x}^n + \nu_{ijs} + n_{kr} : a_{kr}(\mathbf{x}^n + \nu_{ijs}) d_{ijs} x_{is}^n, \quad (18)$$

where we have defined $a_0(\mathbf{x}^n) = \sum_{k=1}^{N_v} \sum_{r=1}^{N_r} a_{kr}(\mathbf{x}^n)$. By computing the analogous values for $\mathscr{D}\mathscr{M}\delta(\mathbf{x}^n)$ and taking the difference, we obtain the following expressions for the commutator error $\mathscr{E}^{n+1}$ for the reachable states $\tilde{\mathbf{x}}_l$

$$\mathbf{x}^n + \boldsymbol{n}_{kr} + \nu_{ijs} : \frac{\Delta t^2}{2} \begin{cases} d_{ijs} x_{is} (a_{ir}(x_{i,\cdot}^n) - a_{ir}(\mathbf{x}_{i,\cdot}^n + \nu_{ijs})) + d_{ijs} n_{irs} a_{ir}(\mathbf{x}_{i,\cdot}^n) & k = i \\ d_{ijs} x_{is} (a_{jr}(x_{j,\cdot}^n) - a_{jr}(\mathbf{x}_{j,\cdot}^n + \nu_{ijs})) & k = j \\ 0 & k \neq \{i, j\} \end{cases} \quad (19)$$

$$\mathbf{x}^n + \boldsymbol{n}_{ir} : -\frac{\Delta t^2}{2} \sum_{s=1}^{S} \sum_{j \neq i}^{N} d_{ijs} n_{irs} a_{ir}(\mathbf{x}_{i,\cdot}^n) \quad (20)$$

$$\mathbf{x}^n + \nu_{ijs} : \frac{\Delta t^2}{2} \sum_{r=1}^{R} d_{ijs} x_{is} (a_{ir}(\mathbf{x}_{i,\cdot}^n + \nu_{ijs}) - a_{ir}(\mathbf{x}_{i,\cdot}^n) + a_{jr}(\mathbf{x}_{j,\cdot}^n + \nu_{ijs}) - a_{jr}(\mathbf{x}_{j,\cdot}^n)) \quad (21)$$

$$\mathbf{x}^n : 0. \quad (22)$$

As seen in the expression for the cross-terms in (19), large parts of the operators commute. Only terms pertaining to degrees of freedom that are sharing an edge in the mesh will be non-zero. Furthermore, (19) will be non-zero only if the reaction described by $\mathbf{n}_{kr}$ depends on species s, i.e. n_{krs} is non-zero. A graphical representation of the reachable states is given in Fig. 2.

Local error in mean and variance

Equations (19) – (22) give expressions for the local error in the PDF in one timestep (at t^{n+1}), given that we observe state $\mathbf{x}^n$ at time t^n. The obvious advantage of computing these expressions directly is that they can be used to derive many estimates of the error such as l_1, l_∞ in PDF, Kolmogorov distance for marginal distributions, or moments. However, they are rather expensive to compute directly, even with appropriate optimizations. In many cases it is sufficient to control a weak error such as the errors in mean and variance. This is particularly true if the error in the individual methods for propagating the reaction operator and the diffusion operator is controlled only in a weak sense. For example, this would be the case if the diffusion operator were updated with τ-leaping such as in [11].

Based on (19) – (22) we can calculate the local error in the expected value of an arbitrary bounded function $g(X_{is})$

$$\Delta E[g(X_{is})] \equiv E[g(X_{is})] - E_{\tilde{f}}[g(X_{is})] = \sum_{k=1}^{S} g(x_{is}^n)\mathscr{E}_k, \quad (23)$$

where the subscript $\tilde{f}$ denotes expectation under the approximate PDF obtained from solving with the operator split method (9) and K is the length of the error vector. For brevity, we have here dropped the superscript $n + 1$ on x and $\mathscr{E}$.

We introduce the notation $\Delta a_{irs}^{-} = a_{ir}(x_{i,.} + \nu_{ijs}) - a_{ir}(x_{i,.})$ and $\Delta a_{irs}^{+} = a_{ir}(x_{i,.} + \nu_{ijs}) - a_{ir}(x_{i,.})$ and $\Delta g(y) = g(x_{is}^n + y) - g(x_{is}^n)$. By summing up outflow and inflow of differences in probability to state X_{is} and using the fact that $\Delta E[g(X_{is})] = \Delta E[g(X_{is}) - g(x_{is}^n)]$, we obtain

$$\begin{aligned}
(2\Delta t^{-2})\Delta E[g(X_{is})] = {} & \Delta g(-1)\sum_{j\neq i}^{N_v} d_{ijs} x_{is} \sum_{r=1}^{N_r} \Delta a_{irs}^{-} + \Delta g(+1)\sum_{j\neq i}^{N_v} d_{jis} \sum_{r=1}^{N_r} n_{jrs} a_{jr}(\mathbf{x}_{j,.}) \\
& + \Delta g(+1)\sum_{j\neq i}^{N_v} d_{jis} x_{js} \sum_{r=1}^{N_r} \Delta a_{ris}^{+} - \sum_{r=1}^{N_r} \Delta g(n_{irs}) \sum_{j\neq i}^{N_v} d_{ijs} n_{irs} a_{ir}(\mathbf{x}_{i,.}) \\
& - \sum_{r=1}^{N_r} \Delta g(n_{irs}) \sum_{j\neq i}^{N_v} \sum_{s'\neq s}^{N_s} d_{ijs'} x_{is'} \Delta a_{irs'}^{-} - \sum_{r=1}^{N_r} \Delta g(n_{irs}) \sum_{j\neq i}^{N_v} \sum_{s'\neq s}^{N_s} d_{jis'} x_{js'} \Delta a_{irs'}^{+} \\
& + \sum_{r=1}^{N_r} \Delta g(n_{irs} - 1) \sum_{j\neq i}^{N_v} (d_{ijs} n_{irs} a_{ir}(\mathbf{x}_{i,.}) - d_{ijs} x_{is} \Delta a_{irs}^{-} \\
& - \sum_{r=1}^{N_r} \Delta g(n_{irs} + 1) \sum_{j\neq i}^{N_v} d_{jis} x_{js} \Delta a_{irs}^{+}.
\end{aligned} \quad (24)$$

To obtain the error in the mean, set $g(x) = x$. After simplification, we obtain

$$\begin{aligned}
\Delta E^{n+1}[X_{is}|\mathbf{x}^n] = {} & 0.5\Delta t^2 (DR(\mathbf{x}^n))_{is} - 0.5\Delta t^2 \sum_{r=1}^{N_r}\sum_{s'=1}^{N_s}\sum_{j\neq i}^{N_v} n_{irs} d_{ijs'} x_{is'} \Delta a_{irs'}^{-} \\
& - 0.5\Delta^2 \sum_{r=1}^{N_r}\sum_{s'=1}^{N_s}\sum_{j\neq i}^{N_v} n_{irs} d_{jis'} x_{js'} \Delta a_{irs'}^{+} \\
= {} & 0.5\Delta t^2 (DR(\mathbf{x}))_{is} \\
& - 0.5\Delta t^2 \sum_{r=1}^{N_r}\sum_{s'=1}^{N_s} n_{irs} (\Delta a_{irs'}^{-} x_{is'} \sum_{j\neq i}^{N_v} d_{ijs'} + \Delta a_{irs'}^{+} \sum_{j\neq i}^{N_v} d_{ijs'} x_{js'}) \\
= {} & 0.5\Delta t^2 \left((DR(\mathbf{x}))_{is} - \sum_{r=1}^{N_r} n_{irs} \sum_{s'=1}^{N_s} \sigma_{irs'} \right).
\end{aligned} \quad (25)$$

where the values of the vector resulting from the operator R acting on $\mathbf{x}$ are defined by

$$R(\mathbf{x})_{is} = \sum_{r=1}^{N_r} n_{irs} a_{ir}(\mathbf{x}_{i,.}), \quad i = 1 \ldots N_v, \quad s = 1, \ldots, N_s. \quad (26)$$

and

$$\sigma_{irs'} = \Delta a^{-}_{irs'} x_{is'} \sum_{j \neq i}^{N_v} d_{ijs'} + \Delta a^{+}_{irs'} \sum_{j \neq i}^{N_v} d_{jis'} x_{js'}. \quad (27)$$

The ($N_vN_s \times N_vN_s$) matrix D is defined to have diagonal elements $-\sum_{j \neq i}^{N_v} d_{ijs}$ and off-diagonal elements d_{ijs}.

Equation (25) is an exact (up to $\mathcal{O}(\Delta t^3)$) formula for the estimate of the local error in mean for any functional form of the propensity functions $a_r(\mathbf{x})$. For zeroth-order mass action reactions we have $\sigma_{irs'} = 0$. For a first order reaction with a linear propensity (such as creation or monomolecular conversion) of the form $a_{ir}(\mathbf{x}_{i,\cdot}) = k_1 x_{is}$ we have

$$\sigma_{irs} = k_1(-x_{is}\sum_{j \neq i}^{N_v} d_{ijs} + \sum_{j \neq i}^{N_v} d_{jis} x_{js}) = k_1 (D\mathbf{x})_{is} = a_{ir}((D\mathbf{x})_{i,\cdot}) n_{irs} \sigma_{irs'} = 0,$$

and (25) simplifies to

$$\Delta E^{n+1}[X_{is}] = 0.5\Delta t^2 ((DR(\mathbf{x}^n))_{is} - R(D\mathbf{x}^n)_{is}), \quad (28)$$

in agreement with the analogous expression for the commutator error for the reaction-diffusion PDE.

To find the error in the second moment, set $g(x) = x^2$ in (24), to obtain

$$\begin{aligned} \Delta E[X_{is}^2] = {} & 0.5\Delta t^2 (|D|\, R(\mathbf{x}^n))_{is} + 2x_{is}\Delta E[X_{is}] \\ & - \Delta t^2 \sum_{r=1}^{N_r} n_{irs} \sum_{j \neq i}^{N_v} (d_{ijs} x_{is} \Delta a^{-}_{irs} - d_{jis} x_{js} \Delta a^{+}_{irs}) \\ & - 0.5\Delta t^2 \sum_{r=1}^{N_r} n_{irs}^2 \sum_{s'=1}^{N_s} \sigma_{irs'} - \Delta t^2 \sum_{j \neq i}^{N_v} d_{ijs} \sum_{r=1}^{N_r} n_{irs}^2 a_{ir}(\mathbf{x}_{i,\cdot}). \end{aligned} \quad (29)$$

Here, $|D|$ is defined as the element-wise absolute value of the matrix D. While this expression is more complicated than (25), the amount of additional work required to compute it is not that great since the complexity of the extra terms are all $\mathcal{O}(N)$ and most of the expensive propensity function evaluations can be overlapped with the computations involved in (25). A more commonly used second order statistic is the variance $V[X] = E[(X - E[X])^2]$. While this quantity is not readily obtained from (24), an approximate formula for $V[X]$ can be found from the observation

$$\begin{aligned} \Delta V[X] = {} & \Delta E[X^2] - E[X^2] + E_f[X]^2 = \\ & \Delta E[X^2] + (\Delta E[X])^2 - 2E[X] + \Delta E[X] = \\ & \Delta E[X^2] - 2x\Delta E[X] + \mathcal{O}(\Delta t^3), \end{aligned} \quad (30)$$

where $E_f[X]$ is the expected value under the operator split PDF. Hence

$$\Delta V[X_{is}] = 0.5\Delta t^2 (|D|\, R(\mathbf{x}^n))_{is} - \Delta t^2 \sum_{r=1}^{N_r} n_{irs} \sum_{j\neq i}^{N_v} \left(d_{ijs} x_{is} \Delta a^-_{irs} - d_{jis} x_{js} \Delta a^+_{irs} \right) \\ -0.5\Delta t^2 \sum_{r=1}^{N_r} n_{irs}^2 \sum_{s'=1}^{N_s} \sigma_{irs'} - \Delta t^2 \sum_{j\neq i}^{N_v} d_{ijs} \sum_{r=1}^{N_r} n_{irs}^2 a_{ir}(\mathbf{x}_{i,\cdot}) + \mathscr{O}(\Delta t^3). \tag{31}$$

In the asymptotic regime, $\mathscr{O}(\Delta t^3)$ terms are small by definition and $\Delta V[X_{is}]$ is a good approximation to the true local error in variance.

We point out that (11), and hence (25) and (29) are only good estimates of the local error if the $\mathscr{O}(\Delta t^3)$ terms are small. It is well known that an error estimate based on (11) will deteriorate in quality for large timesteps if $\mathscr{D}$ or $\mathscr{M}$ are stiff. As the norm of $\mathscr{D}$ increases rapidly with finer mesh resolution, this will lead to overly conservative estimates for large $\Delta t \gg h^2/\gamma_s$ (in the non-asymptotic regime). For this reason, error estimates that perform better in the non-asymptotic regime have been devised in the PDE case [4]. However, it is not clear how those approaches would apply to the RDME. In the following section we will demonstrate that (11) is simple enough to lead to a strategy that is both computable in practice during the coarse of a single realization of RDME and local in space so that it has potential to be efficiently implemented in parallel.

4 Results

To illustrate the accuracy and correctness of the estimate of the local splitting errors (25) and (29), we simulated a model of Min oscillations in *E. Coli* [7] in one spatial dimension for a single timestep. To isolate the splitting error, the reaction and diffusion steps were solved exactly by simulating them with SSA and NSM respectively. Fig. 3 shows the estimated conditional error in mean (25) and the true conditional error in mean as a function of the spatial coordinate (left) and for different timesteps (right). Fig. 4 shows the corresponding estimate of error in second moment based on (29). As can be seen, the estimated error accurately captures the true local error, and the quality of the estimate improves for smaller timesteps as expected from the $\mathscr{O}(\Delta t^3)$ error in the estimates. When computing the error in mean, the observed error is a combination of sampling error caused by a finite number of realizations, and the error caused by operator splitting. For small splitting errors, a large number of realizations are necessary to distinguish the sampling error from the splitting error, especially if we want to measure the error in e.g. the L_1 norm, in which case the variances in the different voxels add up. We note that to achieve tight confidence intervals on the true error for this simple system (see to the top row of Figs. 3 and 4) required 10^{11} realizations.

4.1 Example: Adaptive Diffusive Finite State Projection method

To illustrate the use of the local error estimate, we extended a previously published approximate method for the RDME, the Diffusive Finite State Projection (DFSP) method [6] with temporal adaptivity. This results in increased robustness of the solver.

DFSP relies on Lie-Trotter splitting to separate the reaction and diffusion updates over a discrete timestep Δt. The reactions are then updated in each voxel using Gille-spie's direct method [12]. Diffusion is updated by sampling from probability density functions that are precomputed by solving the diffusion master equation (3) locally up to the given timestep Δt. For an unstructured mesh, there is one such PDF for each vertex in the mesh and for each Δt. The (spatial) locality of the PDFs are enforced by applying an absorbing boundary condition at a certain distance away from the vertices. The timestep is assumed to be small enough that the majority of probability is located close to the vertex where the molecule is located at the beginning of the timestep so that the error due to the truncation of the statespace is small. It is shown in [6] that this strategy can speed up simulations by effectively aggregating the effects of many fast diffusive transfers over the splitting timestep, and in [5] it is discussed under what conditions one can expect DFSP to be more efficient than NSM for the MinCDE model [15].

DFSP has been implemented previously as an add-on solver in the URDME framework [5]. URDME is a modular framework that uses unstructured meshes for spatial stochastic simulations. URDME includes interfaces for handling of the geometry, mesh and computation of diffusion jump rates for the unstructured mesh. It has a modular design which facilitates the implementation of new algorithms as add-on solvers. In the current implementation of DFSP in URDME, Δt has to be chosen by trial and error by manually picking a timestep that results in satisfactory error in the computed solutions, judged by *a posteriori* checks.

We implemented our error estimation strategy in URDME for arbitrary processes (i.e. it uses the same input files as all the core solvers) for structured and unstructured meshes, and supplemented the DFSP solver with an adaptive selection of the timestep. To compute the next proposed timestep, we control the per species error in mean (25) in the L_1 norm such that

$$\frac{\sum_{i=1}^{N} |\mathscr{V}_i| \, |\Delta E[X_{is}]|}{\sum_{i=1}^{N} |\mathscr{V}_i| \, x_{is}} \leq \epsilon_s, \quad s=1,\ldots,S \qquad (32)$$

with ϵ_s being a user specified relative error tolerance. The L_1 norm is a natural choice in the context of DFSP since the error in the diffusion lookup-tables has a natural bound in the l_1 norm [6]. Figure 5 shows an overlay of the pole-to-pole oscillation pattern of membrane bound MinD in E. coli along with time steps selected by the adaptive step size selection algorithm. The timesteps are themselves stochastic variables and fluctuate during the course of the simulation. In the figure they have been smoothed by taking the average of the 10 last timesteps to more clearly visualize how they adapt to the dynamics of the solution.

We simulated the MinCDE model from [15] in 3D in three different ways: using the adaptive timestep control, using a fixed manually determined timestep and using the exact NSM solver, and compared the execution times. The geometry and diffusion constants were taken to be the same as in [10].The model files for URDME that contains all parameters of the model is included in the online Supplementary material of this paper. In DFSP, the

diffusion step is conducted by computing new positions for each molecule individually by sampling from a precomputed lookup-table. All the molecules' new positions can then be sampled in parallel. During the reaction update, all the voxels can be treated independently and in parallel. Note that the local error estimators used here have the same characteristics when it comes to the spatial access pattern as the DFSP diffusion operator. Thus the overall adaptive algorithm should parallelize quite well on a shared memory multicore architecture. To demonstrate this, we implemented a na¨ýve parallel version using openMP. Table 1 shows simulation speeds for the adaptive DFSP method, with error estimation conducted every tenth timestep, on a machine with an Intel(R) Core(TM) i7-2600 CPU @ 3.40GHz (Nehalem) processor and 6GB of RAM using one core and using 4 cores (8 threads). For comparison, we also include simulation times of the (serial) NSM solver in URDME. As can be seen, the DFSP algorithm shows an almost ideal speedup on 4 cores, and for this error tolerance, the adaptive code is roughly two times faster than NSM on a single core.

5 Discussion

In this work we have derived local error estimators for first order operator splitting in stochastic reaction-diffusion simulations based on the RDME. Operator splitting provides a way to decouple reactions and diffusion in spatial stochastic simulations, and the local error estimates enable this to be accomplished with a controlled error.

Decoupling the operators allows for approximate methods to be applied to the diffusion operator to reduce the cost of frequent diffusive transfers. The method we considered as an example in this paper, DFSP, falls into this category. Enabling temporal adaptivity for such approximate and hybrid methods has several benefits. Possibly the most important is the robustness it adds to the solver. While it is possible to prescribe an error tolerance that is likely to give a reasonable result across a range of models, the same is not possible by selecting a fixed timestep *a priori*. Adaptivity has long been available in state-of-the art software for e.g. numerical solution of ODEs and PDEs, but was previously lacking for this class of approximate methods for the RDME. Another major benefit is one of efficiency; using a globally conservative fixed timestep can lead to sub-optimal simulation speeds. We found that the adaptive version of the DFSP method presented here was roughly 1.5 times faster than the corresponding fixed timestep method using a conservative timestep for the oscillatory MinCDE model problem we considered. Even though this experiment is somewhat artificial since the conservative fixed timestep could not be known before running the adaptive code (which again illustrates the utility of adaptivity), it illustrates the potential efficiency benefits of an adaptive solver.

A detailed discussion and analysis about the possible speedups of a hybrid or approximate method has to consider the particular method that is used to update the diffusion step in the splitted scheme. In this paper we use the DFSP method as an example of an approximate method, and the DFSP method is described in greater detail in the reference [6]. There are certain tradeoffs specific to DFSP when it comes to computing and sampling from the lookup tables, and if the timesteps that are required based on the error estimate become too small, the speed benefit from DFSP will be modest since the number of effective diffusion events that can be aggregated are small. There is also an overhead from estimating the

splitting error. Again, this overhead will depend on the value chosen for the error tolerance, since this dictates the timestep. While DFSP as such can yield speedup for stiff systems (when the diffusion events are numerous), we expect the biggest performance gain compared to NSM to come from a parallel implementation since the splitted scheme enables more parallelism, while kMC methods like NSM are intrinsically very hard to parallelize efficiently. For the example in this paper we showed speedups of 7-10x over serial NSM for the approximate scheme on a 4 core machine, using a simplistic OpenMP implementation. A key property of the error estimator is that it can be efficiently applied in a parallel, shared memory implementation.

Finally, operator splitting allows the flexibility of coupling different types of models (possibly at different scales) and solvers. Thus this error estimator opens avenues of research in the development of hybrid algorithms with controlled errors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Per Lötsedt provided valuable comments on the manuscript. This work was funded by National Science Foundation (NSF) Award No. DMS-1001012, ICB Award No. W911NF-09-0001 from the U.S. Army Research Office, NIBIB of the NIH under Award No. R01-EB014877-01, and (U.S.) Department of Energy (DOE) Award No. DE-SC0008975. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of these agencies.

References

1. Altschuler SJ, Angenent SB, Wang Y, Wu LF. On the spontaneous emergence of cell polarity. Nature. 2008; 454:886–890. [PubMed: 18704086]
2. Bayati BS. Fractional diffusion-reaction stochastic simulations. J Chem Phys. 2013; 138(10): 104117. [PubMed: 23514475]
3. Burrage, K.; Hegland, M.; MacNamara, S.; Sidje, RB. A Krylov–based finite state projection algorithm for solving the chemical master equation arising in the discrete modelling of biological systems. In: Langville, A.; Stewardt, W., editors. 150th Markov Anniversary Meeting, Boson Books. Raleigh, NC: 2006. p. 21-38.
4. Descombes S, Dumont T, Louvet V, Massot M. On the local and global errors of splitting approximations of reaction-diffusion equations with high spatial gradients. Int J Comput Math. 2007; 84(6):749–765.
5. Drawert B, Engblom S, Hellander A. URDME: a modular framework for stochastic simulation of reaction-transport processes in complex geometries. BMC Systems Biology. 2012; 6(1):76. [PubMed: 22727185]
6. Drawert B, Lawson MJ, Petzold L, Khammash M. The diffusive finite state projection algorithm for efficient simulation of the stochastic reaction-diffusion master equation. J Chem Phys. 2010; 132(7):074101. [PubMed: 20170209]
7. Elf J, Ehrenberg M. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. Syst Biol. 2004; 1(2):230–236.
8. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. Science. 2002; 297(5584):1183–1186. [PubMed: 12183631]
9. Engblom S, Ferm L, Hellander A, Lötstedt P. Simulation of stochastic reaction–diffusion processes on unstructured meshes. SIAM J Sci Comput. 2009; 31(3):1774–1797.

10. Fange D, Elf J. Noise induced Min phenotypes in *E coli*. PLoS Comput Biol. 2006; 2(6):e80. [PubMed: 16846247]
11. Ferm L, Hellander A, Lötstedt P. An adaptive algorithm for simulation of stochastic reaction-diffusion processes. J Comput Phys. 2010; 229(2):343–360.
12. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reacting systems. J Comput Phys. 1976; 22:403–434.
13. Gillespie DT. A rigorous derivation of the chemical master equation. Physica A. Sep.1992 188:404–425.
14. Giorgos Arampatzis MAK, Plechac P. Parallelization, processor communication and error analysis in lattice based kinetic Monte Carlo. Arxiv preprint arXiv:1208.1049v1. 2012
15. Huang KC, Meir Y, Wingreen NS. Dynamic structures in *escherichia coli*: Spontaneous formation of MinE and MinD polar zones. Proc Natl Acad Sci USA. 2003; 100(22):12724–12728. [PubMed: 14569005]
16. Jahnke T, Altintan D. Efficient simulation of stochastic reaction systems with a splitting method. BIT Num Math. 2010; 50(4):797–822.
17. Jahnke T, Lubich C. Error bounds for exponential operator splittings. BIT Numerical Mathematics. 2000; 40:735–744.10.1023/A:1022396519656
18. Jahnke T, Udrescu T. Solving the chemical master equation by adaptive wavelet compression. J Comput Phys. 2010; 229(16):5724–5741.
19. Lampoudi S, Gillespie D, Petzold L. The multinomial simulation algorithm for discrete stochastic simulation of reaction-diffusion systems. J Chem Phys. 2009; 130(9):094104. [PubMed: 19275393]
20. MacNamara S, Burrage K, Sidje RB. Multiscale modeling of chemical kinetics via the master equation. Multiscale Model Simul. 2008; 6(4):1146–1168.
21. Marquez-Lago TT, Burrage K. Binomial tau-leap spatial stochastic simulation algorithm for applications in chemical kinetics. J Chem Phys. 2007; 127(10):104101. [PubMed: 17867731]
22. Mateescu M, Wolf V, Didier F, Henzinger T. Fast adaptive uniformisation of the chemical master equation. IET Systems Biology. 2010; 4(6):441–452. [PubMed: 21073242]
23. McAdams H, Arkin A. Stochastic mechanisms in gene expression. Proceedings of the National Academy of Sciences. 1997; 94(3):814–819.
24. Munsky B, Khammash M. The finite state projection algorithm for the solution of the chemical master equation. J Chem Phys. 2006; 124(4):044104. [PubMed: 16460146]
25. Rossinelli D, Bayati B, Koumoutsakos P. Accelerated stochastic and hybrid method for spatial simulations of reaction-diffusion systems. Chem Phys Lett. 2008; 451:136–140.
26. Sturrock M, Hellander A, Aldakheel S, Petzold L, Chaplain M. The role of dimerisation and nuclear transport in the hes1 gene regulatory network. Bulletin Math Biol. 2013 to appear.
27. Sturrock M, Hellander A, Marsavinos A, Chaplain M. Spatial stochastic modeling of the hes1 pathway: Intrinsic noise can explain heterogeneity in embryonic stem cell differentiation. J Roy Soc Interface. 2013; 10:20120988. [PubMed: 23325756]
28. Takahashi K, Tănase-Nicola S, ten Wolde PR. Spatio-temporal correlations can dratically change the response of a MAPK pathway. Proc Natl Acad Sci USA. 2010; 107(6):2473–2478. [PubMed: 20133748]
29. Trotter HF. On the product of semi-groups of operators. Proceedings of the American Mathematical Society. 1959; 10(4):545–551.
30. Wedlich-Soldner R, Li R. Spontaneous cell polarization: undermining determinism. Nat Cell Biol. Apr; 2003 5(4):267–270. [PubMed: 12669070]
31. van Zon JS, Rein ten Wolde P. Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. Phys Rev Lett. Apr.2005 94(12):128103. [PubMed: 15903966]

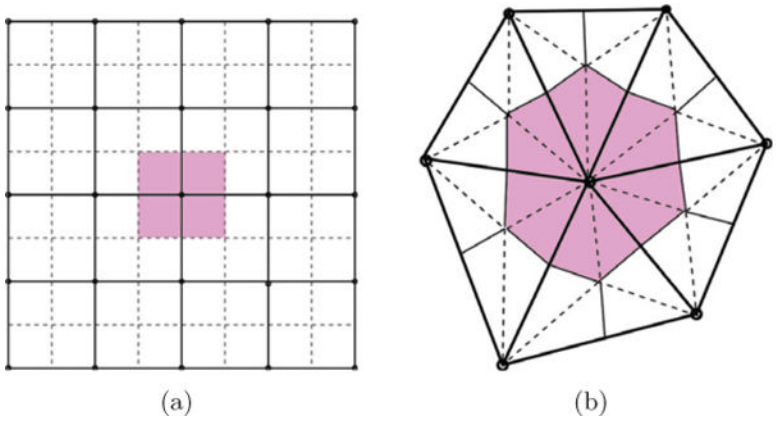

Figure 1.

The molecules are assumed to be well mixed in the dual elements (depicted in pink) of the primal mesh (solid thick lines). On a Cartesian grid (a), the duals are simply the volumes of the staggered grid. The dual of the triangular mesh in (b) is obtained by connecting the midpoints of the edges and the centroids of the triangles. In the conventional RDME, molecules are allowed to jump between immediate neighboring voxels.

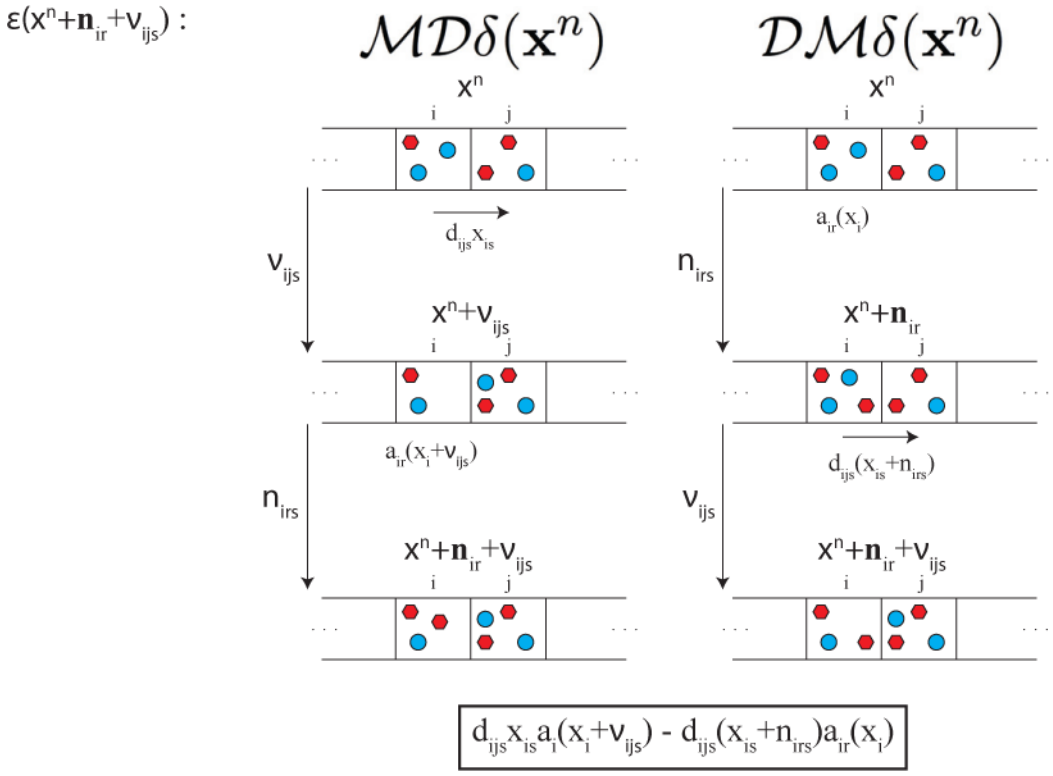

Figure 2.

Illustration of the reachable states in Eq. (19) for $k = i$. On the left side, the diffusion operator is applied first, as in Eq. (14), followed by the reaction operator, resulting in the value derived in Eq. (18). On the right side the order of the operators is reversed, resulting in a change in the final value for the reachable state $\mathbf{x}^n + \mathbf{n}_{ir} + \nu_{ijs}$ (note that the reachable state is, by definition, the same using both orders of the operators, but the value associated with that state is different). The terms in the box represent the difference between the two orders of applying the operators and show terms that do not commute. Note that if we had defined the reachable state as $k \neq i, j$ in Eq. (19) then the reaction would occur in neither the originating subvolume nor the destination subvolume for the diffusion event and the terms would commute (and the value in the box would be zero).

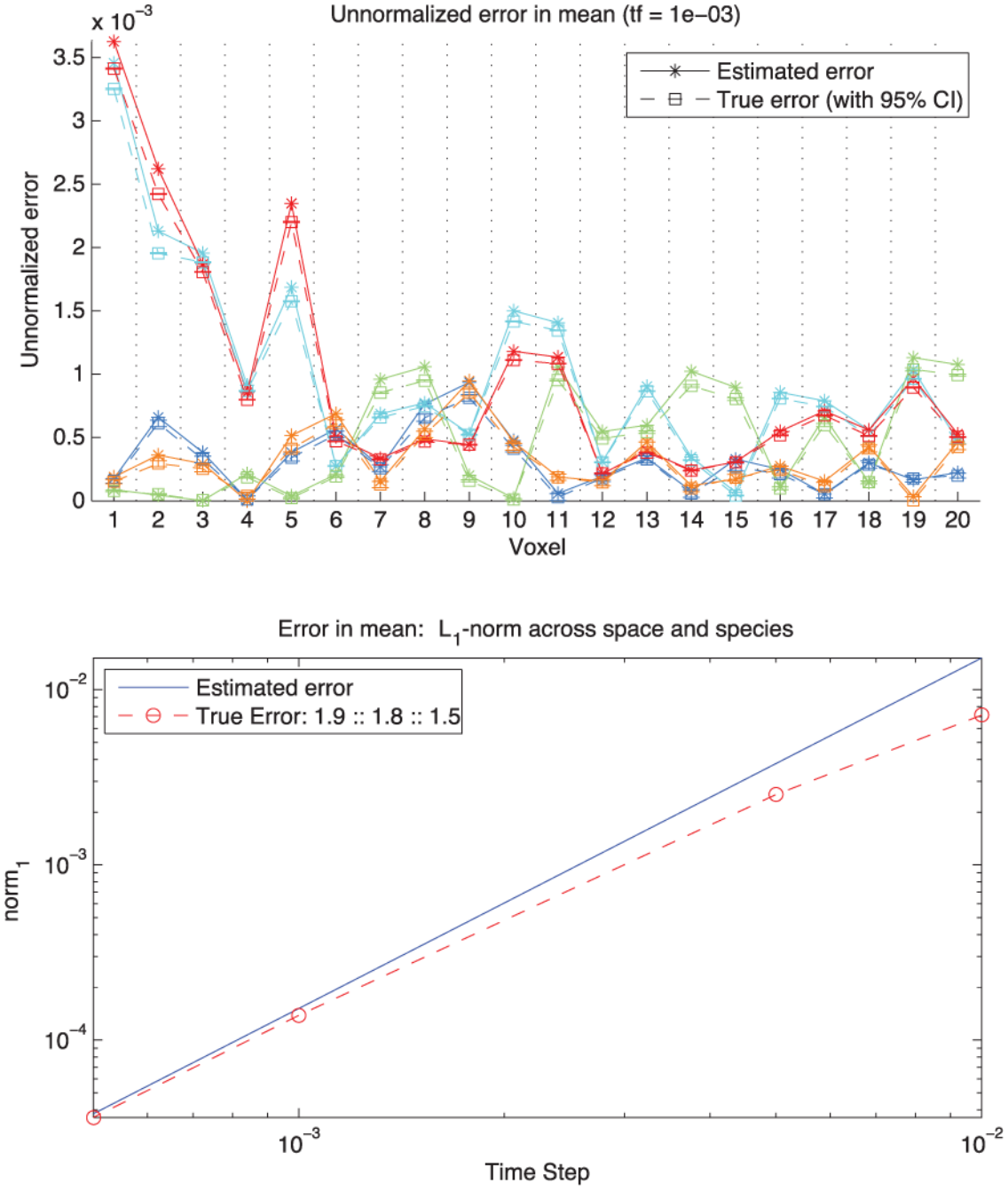

Figure 3.

The local error in mean for a single input point x0 vs the estimated error. **Top:** True (square) and estimated error (*) for each species and voxel for t = 0.001s (voxels are divided by black dotted lines). Blue: species 1, cyan: species 2, green: species 3, red: species 4, orange: species 5. **Bottom:** The L_1-norm across species and voxels of the value in Equation (32) for various time-steps. The solid blue is the estimated error from Equation (25). The dashed red line is the true local error. The values in the legend are the slope of the curves in the loglog

plot. Note that as the time step decreases the estimator converges toward the true error, which in turn approaches the expected $\mathcal{O}(\Delta t^2)$ convergence rate.

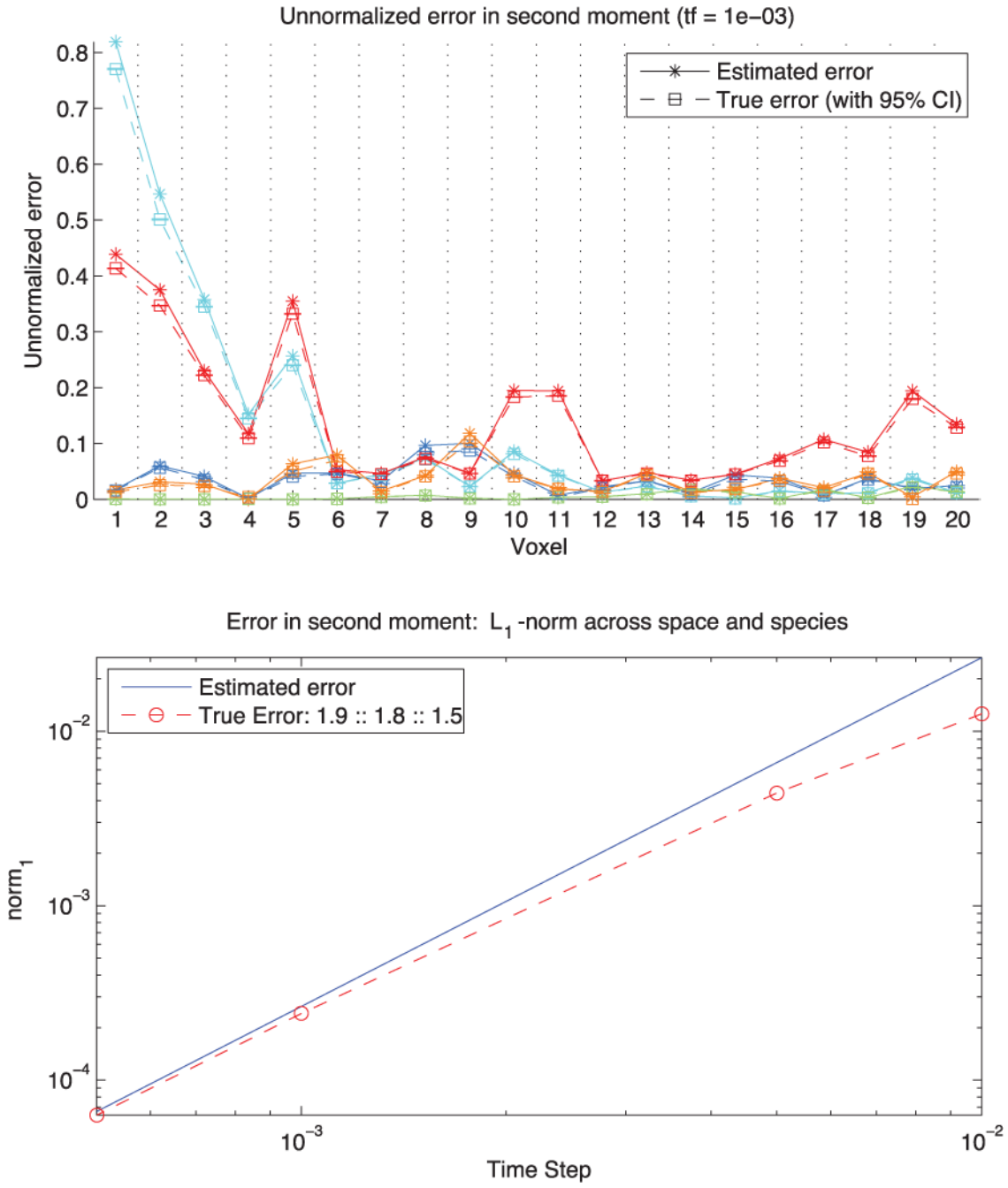

Figure 4.

The local error in second moment for a single input point x0 vs the estimated error. **Top:** True (squares) and estimated error (*) for each species and voxel for t = 0.001s (voxels are divided by black dotted lines). Blue: species 1, cyan: species 2, green: species 3, red: species 4, orange: species 5. **Bottom:** The L_1-norm across species and voxels of the value in Equation (32) for various time-steps. The solid blue is the estimated error from Equation (29). The dashed red line is the true error.

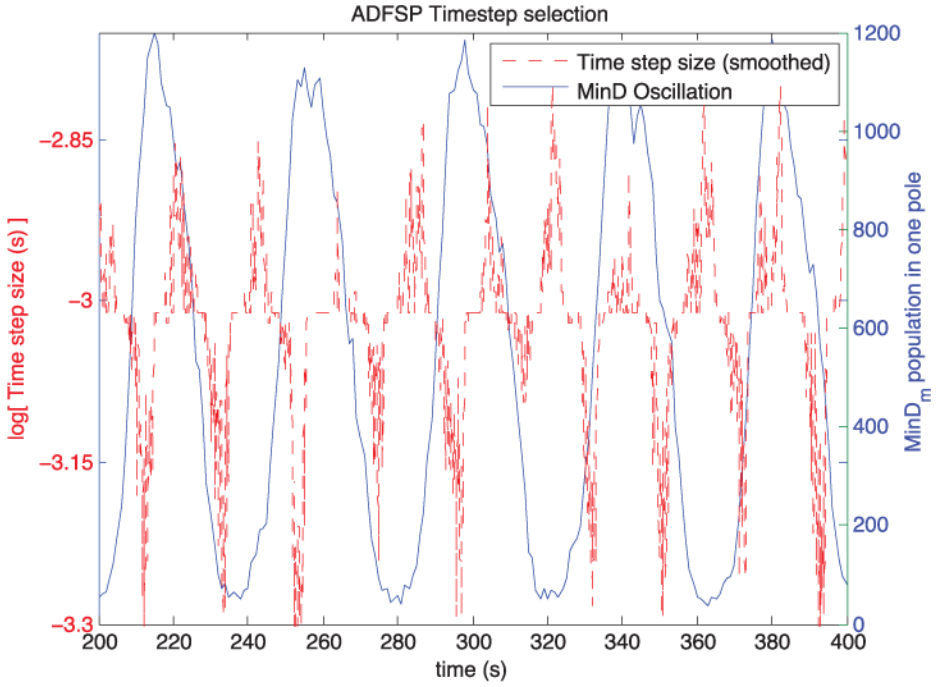

Figure 5.

The figure shows the time steps selected by the adaptive DFSP method, along with the oscillation pattern of the membrane bound MinD protein for a representative trajectory. Note that the timestep adapts to the dynamics of the MinD oscillation. Due to the oscillatory nature of this model, picking the most conservative timesteps based on the maximal errors would lead the solver to take unnecessarily many steps, resulting in suboptimal performance.

Table 1

Simulation time(s) for the adaptive version of the DFSP solver and for the exact NSM solver for the MinCDE problem with varying mesh resolution. The trajectories were simulated to a final time of 2000 s, and the state was sampled every second.

Adaptive DFSP (ϵ_s = 0.05)			
# Voxels	**1 core**	**4 cores**	**NSM (serial)**
588	124	35	337
1009	228	65	537
2818	757	190	1282