ADAPTIVE SOLUTION OF THE MASTER EQUATION IN LOW DIMENSIONS

LARS FERM[1], PER LÖTSTEDT[1] *

[1]*Division of Scientific Computing, Department of Information Technology*
Uppsala University, SE-75105 Uppsala, Sweden
emails: `ferm@it.uu.se, perl@it.uu.se`

Abstract

The master equation satisfied by a probability density function is solved on a grid with a cell size $h > 1$. A modified master equation is derived for the time development of the average of the density in the larger cells. The accuracy of the approximation is studied and the total probability is conserved. Based on an estimate of the discretization error, the cell size is dynamically adapted to the solution. The method is suitable for a few space dimensions and is tested on a model for the migration of people. Substantial savings in memory requirements and CPU times are reported in numerical experiments.

Keywords: master equation, adaptation, cell averages

AMS subject classification: 65M20, 65M50, 65M60

1 Introduction

The time evolution of many systems in molecular biology, population dynamics, epidemiology, and migration of animals and people are well modeled by macroscopic, deterministic systems of ordinary differential equations for the mean values. Quite often when the numbers of particles or animals involved are small, the randomness in the dynamic process cannot be ignored. A deterministic model based on the macroscopic differential equations is then inaccurate. A possibility is to add stochastic terms to the equations to account for the fluctuations but the properties of these terms are usually unknown. A better alternative is offered by the master equation (ME), where the modeling of the noise is inherent in the equation [6, 11]. Examples of stochastic, mesoscopic modeling of systems in biology, epidemiology, and sociology using the ME are found in [2, 3, 15, 16, 17, 20, 22, 23].

The ME for a system in e.g. biology, chemistry, or physics, is a scalar differential-difference equation for the dynamics of its probability density function (PDF) $p(\mathbf{x}, t)$. The PDF measures the probability of the system at time t to be in the discrete state $\mathbf{x}$ with a component $x_i \in \mathbb{Z}_+$ representing e.g. the number of molecules or animals of species i. The non-negative integer numbers are here denoted by $\mathbb{Z}_+$. With N number of species in the system, $\mathbf{x}$ has N components and the ME has to be solved in N dimensions and time. If $x_{\max}$ is the maximum copy number for each one of the species, then the solution domain will grow like $x_{\max}^N$ making direct computational solution of the ME possible only when $N = 2$ or 3 for large $x_{\max}$.

Gillespie's Stochastic Simulation Algortihm (SSA) [7] is a Monte Carlo method for simulation of one time trajectory of the system. The memory requirements for SSA grow linearly with N. By collecting the statistics from M trajectories or realizations, where M is large, the mean values, the higher moments, and the PDF can be determined. A disadvantage with SSA is the slow convergence in M.

For problems with low dimension, direct solution of the ME or a Fokker-Planck approximation is an alternative. The Fokker-Planck approximation is a partial differential equation and can be solved by standard methods. Solving the Fokker-Planck equation is about 2000 times faster than SSA in an example with $N = 2$ in [19] but SSA is often more efficient when $N = 4$. By splitting the variables into one set treated as stochastic variables and one set as deterministic variables in a hybrid scheme, the dimension of the stochastic part is reduced and the major part of the variables is treated macroscopically [10, 12]. The computational work in an example in [10] with $N = 22$ is an order of magnitude lower with the hybrid method compared to SSA.

An adaptive method to directly solve the ME for a few species is developed in this paper motivated by the efficiency of direct solution of the equation in low dimensions, the possibility of coupling it to the macroscopic model in a hybrid method, and that many models e.g. in [2, 16, 20, 23] are low-dimensional. The states are aggregated in cells containing more than one state so that the number of unknowns in the computational domain is much smaller than $x_{\max}^N$. The cell average of probability is updated at every time step with the inflow and the outflow of probability through the cell faces. Other methods for the low-dimensional ME are found in [4, 14].

The computational domain is partitioned into a number of blocks. All cells in the same block has the same number of discrete states. The cell size is allowed to change by a factor two at a block boundary. The cell size in a block is adapted to the time-dependent solution so that the estimated error due to the large cells in the ME is sufficiently small. A similar technique for space adaptivity is available for partial differential equations [5, 13] but then $\mathbf{x} \in \mathbb{R}^N$, where $\mathbb{R}$ denotes a real number. A static aggregation of the states is introduced in [9] combined with a sparse grid approximation [1] for higher dimensions. We merge the states into cells and split the cells dynamically based on the error in the ME. The algorithm

in [14] adapts the size of the representation of the solution of low-dimensional problems automatically by a Krylov method. Our method also reduces the size of the original problem but in a different manner.

The paper is organized as follows. The ME is stated and its properties are given in Section 2. Then the discretization of the equation over the cells in a block is presented in the next section. The partitioning of the computational domain and the interpolation at the block boundaries are discussed in Section 4. The error in the equation is estimated in Section 5 and the adaptive procedure is described. In Section 6, the algorithm is applied to a two-dimensional problem modeling the migration of people [23, 24]. Four different scenarios are obtained by changing the parameters in the model. Conclusions are drawn in the final section.

The i:th component of a vextor $\mathbf{x}$ is denoted by x_i. If $\mathbf{x} \geq 0$ or $\mathbf{x} < x_{\max}$ then all components of the vector satisfy the inequality relations. The maximum norm of $\mathbf{x}$ is written $|\mathbf{x}|_\infty$ and the ℓ_2-norm $\|\mathbf{x}\|_2$.

2   The master equation

Assume that we have N species X_i, $i = 1, \ldots, N$, in a system. The copy number of the i:th species is denoted by x_i. A transition from one state $\mathbf{x}_r$ to another state $\mathbf{x}$ takes place with probability $w_r(\mathbf{x}_r, t)$ per time unit. The change in $\mathbf{x}_r$ by a transition is written as a chemical reaction

$$
\mathbf{x}_r \xrightarrow{w_r(\mathbf{x}_r,t)} \mathbf{x}, \ \mathbf{n}_r = \mathbf{x}_r - \mathbf{x}. \tag{1}
$$

Only a few states are changed as a result of a transition r. Birth and death processes can be modeled in this way. When a new copy of species i is born then $n_{ri} = -1$ and the other entries of $\mathbf{n}$ are zero. When a copy of i dies then $n_{ri} = 1$. The number of possible transitions is denoted by R.

The probability density function (PDF) $p(\mathbf{x}, t)$ for the system with the transitions (1) to be in the state $\mathbf{x}$ at time t satisfies the ME [6, 11]. Introduce a splitting of $\mathbf{n}_r$ into two parts so that

$$
\mathbf{n}_r = \mathbf{n}_r^+ + \mathbf{n}_r^-, n_{ri}^+ = \max(n_{ri}, 0), n_{ri}^- = \min(n_{ri}, 0),
$$

and let

$$
q_r(\mathbf{x}, t) = w_r(\mathbf{x}, t)p(\mathbf{x}, t). \tag{2}
$$

Then the ME at $(\mathbf{x}, t)$ is

$$
\frac{\partial p}{\partial t} = \sum_{\substack{r=1 \\ \mathbf{x} + \mathbf{n}_r^- \geq 0}}^{R} q_r(\mathbf{x} + \mathbf{n}_r, t) - \sum_{\substack{r=1 \\ \mathbf{x} - \mathbf{n}_r^+ \geq 0}}^{R} q_r(\mathbf{x}, t). \tag{3}
$$

3

The first sum on the right hand side is the inflow of probability from other states that can reach $\mathbf{x}$ by a transition (1) and the last sum is the outflow of probability to other states that can be reached from $\mathbf{x}$. The inequality conditions are such that no inflow by a transition is possible from a state $\mathbf{x}_r$ with one component i such that $x_{ri} < 0$ and no outflow is possible to a state $\mathbf{x}_r'$ with $x_{ri}' = x_i - n_{ri} < 0$ for some i.

The total probability $P_{\text{tot}} = \sum_{\mathbf{x} \in \mathbb{Z}_+^N} p(\mathbf{x}, t)$ is preserved in time with these conditions at the boundaries with $x_i = 0$ for at least one i [5]. If the initial value $p(\mathbf{x}, 0)$ is non-negative, $p(\mathbf{x}, t)$ will remain non-negative for $t > 0$.

3 Discretization

The discretization of the master equation in space and time is described in this section. For notational convenience, we restrict ourselves to two space dimensions and one transition r. How to extend our algorithm to more dimensions is relatively straightforward. The equation for R transitions is obtained by summation over all r as in (3).
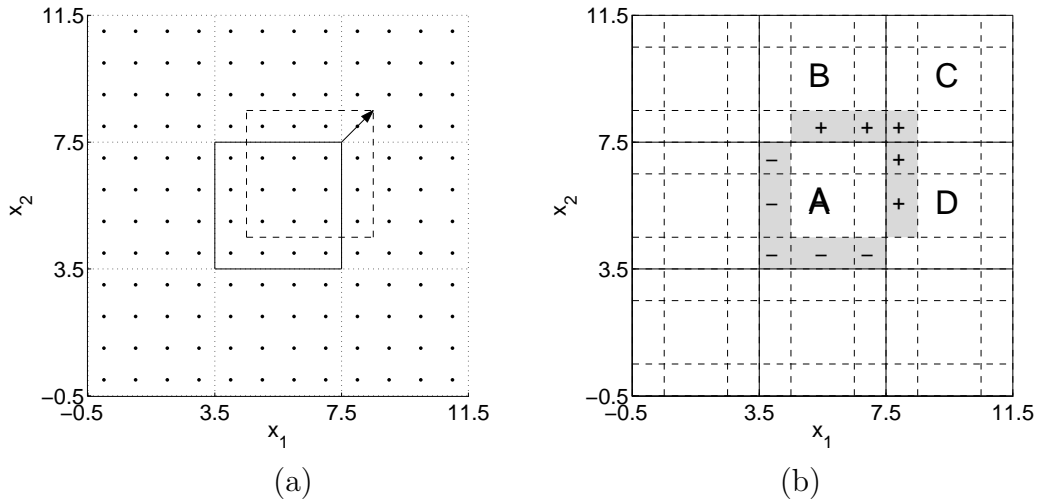
3.1 The master equation for large cells

Fig. 1: The state space is partitioned into cells with $h = 4$. (a) A shifted cell corresponding to $\mathbf{n}_r = (1, 1)^T$ is bounded by dashed lines. (b) The shaded subcells contribute with the indicated sign to the residual in cell A.

The state space $\{\mathbf{x} |\ \mathbf{x} \geq 0\}$ is covered in 2D by square computational cells $\mathcal{C}_{IJ}^h$ with edges of length h. The length is assumed to be a positive integer to fit the block partitioning and the adaptivity but with a slightly different notation the

derivations in this section are valid for any $h \geq 1$. In our examples in Section 6, $h = 1, 2, 4$, or 8. If $h = 1$ then $x_{ij} \in \mathbb{Z}_+$ is the midpoint of the cell

$$\mathcal{C}_{ij} = (i - 0.5, i + 0.5) \times (j - 0.5, j + 0.5), \ i = 0, 1, 2, \ldots, N, \ j = 0, 1, 2, \ldots, N.$$

An outer artificial boundary $x_{\max} > 0$ is introduced in every dimension to limit the computational domain. If $h > 1$ then $\mathcal{C}_{IJ}^h$, $I = 0, 1, 2, \ldots, M_h$, $J = 0, 1, 2, \ldots, M_h$, with $M_h = (N + 1)/h - 1$ consists of the unit cells $\mathcal{C}_{ij}$ such that $hI \leq i < h(I + 1)$, $hJ \leq j < h(J + 1)$. The area of a cell is $|\mathcal{C}_{IJ}^h| = h^2$. This is illustrated in Figure 1, where the parameters are $h = 4, N = 11, M_4 = 2$.

The mean value of $p(\mathbf{x}, t)$ in cell $\mathcal{C}_{IJ}^h$ is $p_{hIJ}(t)$. In $\mathcal{C}_{ij}$ with $h = 1$, $p_{ij} = p(\mathbf{x}_{ij}, t)$ is equal to the mean value. Outflow of probability at the upper boundary due to a reaction is possible only if $\mathbf{x}_{ij} + \mathbf{n}_r^+ < x_{\max}$ and inflow in a reaction must satisfy $\mathbf{x}_{ij} - \mathbf{n}_r^- < x_{\max}$. By summation over the midpoints in the cells with $h = 1$ and including the constraints by the bounded domain, $p_{hIJ}(t)$ satisfies the following ME for one reaction r

$$p_{hIJ} = |\mathcal{C}_{IJ}^h|^{-1} \sum_{x_{ij} \in \mathcal{C}_{IJ}^h} p_{ij},$$

$$\frac{\partial p_{hIJ}(t)}{\partial t} = \frac{1}{|\mathcal{C}_{IJ}^h|} \sum_{x_{ij} \in \mathcal{C}_{IJ}^h} \frac{\partial p_{ij}(t)}{\partial t} =$$

$$\frac{1}{|\mathcal{C}_{IJ}^h|} \left(\sum_{\substack{\{\mathbf{x}_{ij} \in \mathcal{C}_{IJ}^h\} \\ \cap \{\mathbf{x}_{ij} + \mathbf{n}_r^- \geq 0\} \\ \cap \{\mathbf{x}_{ij} + \mathbf{n}_r^+ < x_{\max}\}}} q_r(\mathbf{x}_{ij} + \mathbf{n}_r, t) - \sum_{\substack{\{\mathbf{x}_{ij} \in \mathcal{C}_{IJ}^h\} \\ \cap \{\mathbf{x}_{ij} - \mathbf{n}_r^+ \geq 0\} \\ \cap \{\mathbf{x}_{ij} - \mathbf{n}_r^- < x_{\max}\}}} q_r(\mathbf{x}_{ij}, t) \right). \tag{4}$$

Let $\mathcal{C}_{IJ}^h(\mathbf{n}_r)$ denote the cell $\mathcal{C}_{IJ}^h$ shifted in the $\mathbf{n}_r$-direction, see Figure 1.a, so that $x_{ij} \in \mathcal{C}_{IJ}^h = \mathcal{C}_{IJ}^h(0)$ and $x_{ij} + \mathbf{n}_r \in \mathcal{C}_{IJ}^h(\mathbf{n}_r)$. Then after a change of variables in the first sum in (4) we have

$$\frac{\partial p_{hIJ}(t)}{\partial t} = \frac{1}{|\mathcal{C}_{IJ}^h|} \left(\sum_{\{\mathbf{x}_{ij} \in \mathcal{C}_{IJ}^h(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}} q_r(\mathbf{x}_{ij}, t) - \sum_{\{\mathbf{x}_{ij} \in \mathcal{C}_{IJ}^h(0)\} \cap \mathcal{B}_{ij}} q_r(\mathbf{x}_{ij}, t) \right). \tag{5}$$

The upper and lower boundaries restrict $\mathbf{x}_{ij}$ to the set

$$\mathcal{B}_{ij} = \{\mathbf{x}_{ij} - \mathbf{n}_r^+ \geq 0\} \cap \{\mathbf{x}_{ij} - \mathbf{n}_r^- < x_{\max}\}$$

in (5). Part of the cells $\mathcal{C}_{IJ}^h(\mathbf{n}_r)$ and $\mathcal{C}_{IJ}^h(0)$ may overlap. Introduce the differences between them by

$$\begin{aligned} \Delta \mathcal{C}_{IJ}^{h+}(\mathbf{n}_r) &= \mathcal{C}_{IJ}^h(\mathbf{n}_r) \setminus (\mathcal{C}_{IJ}^h(\mathbf{n}_r) \cap \mathcal{C}_{IJ}^h(0)), \\ \Delta \mathcal{C}_{IJ}^{h-}(\mathbf{n}_r) &= \mathcal{C}_{IJ}^h(0) \setminus (\mathcal{C}_{IJ}^h(\mathbf{n}_r) \cap \mathcal{C}_{IJ}^h(0)). \end{aligned} \tag{6}$$

5

When $h = 1$ or $h \geq 2$ and $n_{ri} \geq h$ for some i, then $\mathcal{C}_{IJ}^h(\mathbf{n}_r) \cap \mathcal{C}_{IJ}^h(0) = \emptyset$ and $\Delta\mathcal{C}_{IJ}^{h+}(\mathbf{n}_r) = \mathcal{C}_{IJ}^h(\mathbf{n}_r)$ and $\Delta\mathcal{C}_{IJ}^{h-}(\mathbf{n}_r) = \mathcal{C}_{IJ}^h(0)$. With this notation, (5) can be written

$$\frac{\partial p_{hIJ}(t)}{\partial t} = \frac{1}{|\mathcal{C}_{IJ}^h|}\left(\sum_{\{\mathbf{x}_{ij} \in \Delta\mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}} q_r(\mathbf{x}_{ij}, t) - \sum_{\{\mathbf{x}_{ij} \in \Delta\mathcal{C}_{IJ}^{h-}(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}} q_r(\mathbf{x}_{ij}, t) \right). \tag{7}$$

The summation of the flux for each reaction r is over the difference between the cell and the shifted cell as illustrated in Figure 1.b for $\mathbf{n}_r = (1,1)^T$. The averages of the PDF are known as averages in the surrounding cells $A, B,$ and C and interpolation is necessary to compute p in the subcells $\Delta\mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)$ (shaded with +-signs in Figure 1.b) and $\Delta\mathcal{C}_{IJ}^{h-}(\mathbf{n}_r)$ (shaded with −-signs in Figure 1.b). The right hand side or residual in (7) is denoted by Φ such that the equation is

$$\frac{\partial p_{hIJ}(t)}{\partial t} = \Phi_{IJ}^h(\mathbf{p}_h). \tag{8}$$

Let $\sum_{I,J} |\mathcal{C}_{IJ}^h| p_{hIJ}$ be the total probability P_{tot} in the domain. Summation over I, J in (7) yields

$$\frac{\partial P_{\text{tot}}}{\partial t} = \frac{\partial}{\partial t} \sum_{I,J=1}^{M_h} h^2 p_{hIJ} =$$

$$\sum_{I,J=1}^{M_h}\left(\sum_{\{\mathbf{x}_{ij} \in \Delta\mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}} q_r(\mathbf{x}_{ij}, t) - \sum_{\{\mathbf{x}_{ij} \in \Delta\mathcal{C}_{IJ}^{h-}(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}} q_r(\mathbf{x}_{ij}, t) \right) = \tag{9}$$

$$\left(\sum_{\{\mathbf{x}_{ij} \in \cup_{I,J=1}^{M_h}\Delta\mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}} q_r(\mathbf{x}_{ij}, t) - \sum_{\{\mathbf{x}_{ij} \in \cup_{I,J=1}^{M_h}\Delta\mathcal{C}_{IJ}^{h-}(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}} q_r(\mathbf{x}_{ij}, t) \right).$$

Every $\mathbf{x}_{ij}$ in $\cup_{I,J=1}^{M_h}\Delta\mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}$ is also in $\cup_{I,J=1}^{M_h}\Delta\mathcal{C}_{IJ}^{h-}(\mathbf{n}_r)\} \cap \mathcal{B}_{ij}$ and conversely. Hence, the two sets are identical and

$$\frac{\partial P_{\text{tot}}}{\partial t} = 0 \tag{10}$$

in (9). The total probability is conserved.

3.2   Fluxes in subcells

According to (7) and Figure 1, the flux q_r shall be computed in the subcells $\Delta\mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)$ and $\Delta\mathcal{C}_{IJ}^{h-}(\mathbf{n}_r)$ for the time evolution of p_{hIJ}. The procedure is derived here for a reaction with $\mathbf{n}_r = (1,0)^T$ to simplify the notation. Let w_{ij} be the

6

propensity of this reaction computed at x_{ij} and t. Let w_{i+h} and w_i be evaluated at the midpoints of each subcell in $\Delta \mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)$ and $\Delta \mathcal{C}_{IJ}^{h-}(\mathbf{n}_r)$ and

$$p_i = h^{-1} \sum_{j=1}^{h} p_{ij}, \; p_{i+h} = h^{-1} \sum_{j=1}^{h} p_{i+h,j}, \; \Delta p_{ij} = p_{ij} - p_i, \; \Delta w_{ij} = w_{ij} - w_i,$$

see Figure 2.

Fig. 2: The subcells i and $i + h$ and the cell I on the x_1-axis.

Then the right hand side of (7) is in the interior of $\{\mathbf{x} \geq 0\}$

$$
\begin{aligned}
&\frac{1}{h^2} \sum_{j=1}^{h} w_{i+h,j} p_{i+h,j} - w_{ij} p_{ij} = \\
&\frac{1}{h^2} \sum_{j=1}^{h} w_{i+h} p_{i+h} - w_i p_i + p_{i+h} \Delta w_{i+h,j} + w_{i+h} \Delta p_{i+h,j} \\
&- p_i \Delta w_{ij} - w_i \Delta p_{ij} + \Delta w_{i+h,j} \Delta p_{i+h,j} - \Delta w_{ij} \Delta p_{ij}.
\end{aligned}
\tag{11}
$$

Two terms in the second sum in (11) vanish by the definition of p_i

$$
\begin{aligned}
\sum_{j=1}^{h} w_i \Delta p_{ij} &= w_i \sum_{j=1}^{h} (p_{ij} - p_i) = 0, \\
\sum_{j=1}^{h} w_{i+h} \Delta p_{i+h,j} &= w_{i+h} \sum_{j=1}^{h} (p_{i+h,j} - p_{i+h}) = 0.
\end{aligned}
$$

Two other terms also vanish if $w(\mathbf{x}, t)$ is constant or varies linearly in x_2

$$
\begin{aligned}
\sum_{j=1}^{h} p_i \Delta w_{ij} &= p_i \sum_{j=1}^{h} (w_{ij} - w_i) = 0, \\
\sum_{j=1}^{h} p_{i+h} \Delta w_{i+h,j} &= p_{i+h} \sum_{j=1}^{h} (w_{i+h,j} - w_{i+h}) = 0.
\end{aligned}
$$

Ignoring the two products of the deviations Δw and Δp in the x_2-direction in (11), the approximation of the ME (7) is when $\mathbf{n}_r = (1, 0)^T$

$$\frac{\partial p_{hIJ}(t)}{\partial t} = \frac{1}{h}(w_{i+h} p_{i+h} - w_i p_i). \tag{12}$$

An analysis of a more general reaction is possible along the same lines as above with a more complicated notation but with the same final approximation: in each subcell evaluate the averages of the densities and the propensities at the midpoint. The error in w vanishes as $h \to 1$.

3.3 Interpolation of p

In order to compute the flux q_r in a subcell, we need w_r and p in the same subcell. An alternative to the approximation in (12) is to interpolate q_r and use (7) but this interpolation will be different for every reaction. By interpolating p once, the same approximation is available for all reactions.

The p-averages in the subcells are recovered from the values in the neighboring cells in one sweep in the x_1-direction and one sweep in the x_2-direction. Then p is known in $\Delta \mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)$ and $\Delta \mathcal{C}_{IJ}^{h-}(\mathbf{n}_r)$ as required by the ME in (7) and (12) for updating p_{IJ}. The interpolation formulas are derived for reactions with $|n_{ri}| \leq 1$ but can be extended in the same manner to reactions with $|n_{ri}| \leq 2$. Firstly, the approximations of p are determined in Figure 3 between the grid lines 0 and 1, 1 and 3, 3 and 4, and so on in the x_1-direction for all cells when $h = 4$. Secondly, using these values the same procedure is repeated in the x_2-direction.
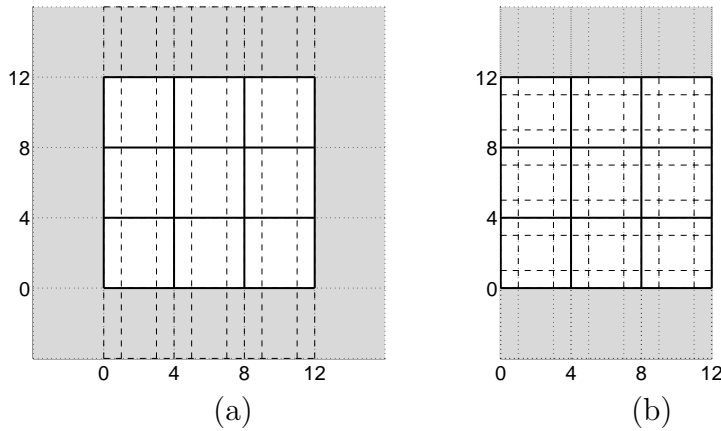

Fig. 3: The interpolation of p to the subcells in two steps from (a) to (b). The values in the shaded cells are also involved in the interpolation formulas.

Let P_i and p_k be the probabilities per unit square in the cells B_i and subcells b_k, respectively, in Figure 4.
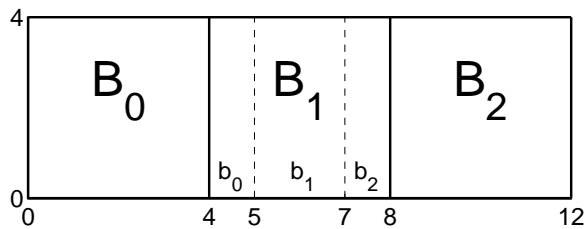

Fig. 4: The cells B_i are involved in the interpolation when a cell is split into the parts b_0, b_1, b_2 between the grid lines 4 and 5, 5 and 7, and 7 and 8.

The one-dimensional interpolation formula is chosen such that it is exact for

8

P_i and p_k being the average

$$\overline{x^j}|_{\mathcal{C}} = \frac{1}{|\mathcal{C}|} \int_{\mathcal{C}} x^j \, dx, \ j = 0, 1, 2,$$

in the corresponding cell or subcell $\mathcal{C}$. Then the formulas are

$$
\begin{aligned}
p_0 &= P_1 + \frac{h-1}{4h}\left(\frac{h-2}{3h}(P_2 - 2P_1 + P_0) - (P_2 - P_0)\right), \\
p_1 &= P_1 + \frac{h-1}{6h^2}(P_2 - 2P_1 + P_0), \\
p_2 &= P_1 + \frac{h-1}{4h}\left(\frac{h-2}{3h}(P_2 - 2P_1 + P_0) + (P_2 - P_0)\right).
\end{aligned}
\tag{13}
$$

The interpolation is conservative in the sense that

$$1 \cdot p_0 + (h-2)p_1 + 1 \cdot p_2 = hP_1.$$

The subcell b_1 in the middle does not exist when $h = 1$ or $h = 2$, but the interpolation formula is still valid for p_0 and p_2. For example, with $h = 1$ we have $p_0 = p_1 = p_2 = P_1$ which are the true values. The approximations of p and w (see Section 3.2) are exact when $h = 1$. Thus, the discretization error in the flux q is zero for unit cells. When h is large, the formula (13) for p_0 can be regarded as an interpolation of p_0 on the edge between B_0 and B_1 and $p_1 = P_1 + \mathcal{O}(h^{-1})$.

Assume that $p = \pi(x_1)$ in the neighborhood of the cell b_0. The interpolation of $\pi(x_1)$ in the x_1-direction in that cell is denoted by $\mathcal{I}_0^1$ such that

$$p_0 = \mathcal{I}_0^1(\pi(x_1)) = \overline{\pi(x_1)}|_{b_0} + \phi_0^1(\pi(x_1)). \tag{14}$$

The interpolation (13) is exact and the error term ϕ_0^1 is zero for $\pi(x_1) = x_1^j$, $j = 0, 1, 2$, and linear combinations of these powers of x_1. When $j = 3, 4$ and x_{10} is the coordinate of the left edge of B_1 and b_0, the error is

$$
\begin{aligned}
\phi_0^1(x_1^3) &= -(h-1)(2h-1)(h+1)/4, \\
\phi_0^1(x_1^4) &= -(h-1)(2h-1)(h+1)(5x_{10} + 2h + 1)/5.
\end{aligned}
\tag{15}
$$

The interpolation error is zero when $h = 1$.

In the x_2-direction, the interpolation $\mathcal{I}_0^2$ is exact for x_2^j, $j = 0, 1, 2$, and for those powers of x_2 its error function ϕ_0^2 is zero. The value p_{00} of p in the lower left corner b_{00} of a cell (e.g. in cell A or C in Figure 1.b) is obtained by applying $\mathcal{I}_0^1$ and subsequently $\mathcal{I}_0^2$. Since $\mathcal{I}_0^1$ is linear, $\mathcal{I}_0^1(\kappa\pi) = \kappa\mathcal{I}_0^1(\pi)$, the two dimensional interpolation of $x_1^j x_2^k$ to obtain its average p_{00} is

$$
\begin{aligned}
p_{00} &= \mathcal{I}_0^2(\mathcal{I}_0^1(x_1^j x_2^k)) = \mathcal{I}_0^2(x_2^k \mathcal{I}_0^1(x_1^j)) = \mathcal{I}_0^2(x_2^k)\mathcal{I}_0^1(x_1^j) \\
&= (\overline{x_1^j}|_{b_{00}} + \phi_0^1(x_1^j))(\overline{x_2^k}|_{b_{00}} + \phi_0^2(x_2^k)) \\
&= \overline{x_1^j x_2^k}|_{b_{00}} + \overline{x_2^k}|_{b_{00}}\phi_0^1(x_1^j) + \overline{x_1^j}|_{b_{00}}\phi_0^2(x_2^k) + \phi_0^1(x_1^j)\phi_0^2(x_2^k).
\end{aligned}
\tag{16}
$$

9

For $0 \leq j$, $k \leq 2$ the formula is exact.

The simplest way of assigning density values to $b_0, b_1,$ and b_2 in Figure 4 would be to let $p_0 = p_1 = p_2 = P_1$ also for $h > 1$ but the accuracy in (13) is much better. The interpolation error for the simple choice is zero only for $\pi(x_1) = \text{const.}$

The residual Φ_{IJ}^h in (8) is linear in $\mathbf{p}_h$ with the approximations in this section. Then the probability densities p_{hIJ} in $\mathbf{p}_h$ satisfy

$$\frac{\partial \mathbf{p}_h}{\partial t} = A(\mathbf{x}, t)\mathbf{p}_h, \tag{17}$$

with a matrix $A \in \mathbb{R}^{M_h \times M_h}$.

3.4 The one-dimensional case

The approximation of the ME in one dimension in cell B_1 in Figure 4 for one reaction with propensity w and $n_r = 1$ is according to (12) and (13)

$$\frac{\partial P_1}{\partial t} = \frac{1}{h}(w_{02}p_{02} - w_{01}p_{01}), \tag{18}$$

where the indices $0j$ denote the left subcells in cells B_j. The propensity is evaluated in the center of the subcells and is exact.

The right hand side in (18) is rewritten as

$$\frac{1}{h}\left(\frac{w_{02} + w_{01}}{2}(p_{02} - p_{01}) + \frac{w_{02} - w_{01}}{2h}h(p_{02} + p_{01})\right).$$

The discretization error τ due to the interpolation of p is then by (14)

$$\frac{1}{h}\left(\frac{w_{02} + w_{01}}{2}(\phi_{02}^1(\pi(x_1)) - \phi_{01}^1(\pi(x_1))) + \frac{w_{02} - w_{01}}{2h}h(\phi_{02}^1(\pi(x_1)) + \phi_{01}^1(\pi(x_1)))\right). \tag{19}$$

For $\pi(x_1) = x_1^3$ we derive from (19) and (15) that

$$\tau_3 = -(h-1)(2h-1)(h+1)\frac{w_{02} - w_{01}}{4h}, \tag{20}$$

and for $\pi(x_1) = x_1^4$ we have

$$\tau_4 = -(h-1)(2h-1)(h+1)\left(\frac{w_{02} + w_{01}}{2} + \frac{10x_{10} + 9h + 2}{5}\frac{w_{02} - w_{01}}{2h}\right). \tag{21}$$

Suppose that p has a Maclaurin expansion in the vicinity of x_{10}

$$p(x_1) = c_0 + c_1 x_1 + c_2 x_1^2 + c_3 x_1^3 + c_4 x_1^4 + \ldots$$

Then the discretization error in (18) is

$$\tau = c_3\tau_3 + c_4\tau_4 + \ldots$$

and it depends on the variation of w in the third order term (20) and on both the mean and the variation of w in the fourth order term (21). Both τ_3 and τ_4 vanish when $h = 1$.

If w is constant then the expression for the time derivative of P_1 is derived from (13)

$$\begin{aligned}\frac{\partial P_1}{\partial t} &= \frac{w}{h}(p_{02} - p_{01}) \\ &= \frac{w}{h}((\gamma - \delta)P_3 + (1 - 3\gamma + \delta)P_2 - (1 - 3\gamma - \delta)P_1 - (\gamma + \delta)P_0),\end{aligned} \tag{22}$$

with $\delta = (h - 1)/4h$ and $\gamma = \delta(h - 2)/3h$. In a Fourier analysis of the space operator in (22), the symbol G of it for the mode $\exp(i\omega h)$ is

$$G(\omega) = 2wh^{-1}s^2(-1 + 4\delta - 4(\delta - \gamma)s^2) + 2ics(1 + 4s^2(\delta - \gamma)), \tag{23}$$

where $s = \sin(\omega h/2)$, $c = \cos(\omega h/2)$. Since $4\delta = (h - 1)/h < 1$ and $\delta - \gamma = \delta(1 - (h - 2)/3h) \geq 0$ for all $h \geq 1$, the real part $\Re G$ of G is non-positive for all values of ωh in $[0, \pi]$. Thus, the differential equation is stable and a sufficiently stable time discretization will produce a bounded numerical solution.

With the simplest choice of p_{02} and p_{01} in (22), $p_{02} = P_2$ and $p_{01} = P_1$ corresponding to $\gamma = \delta = 0$, we have a scheme identical to an upwind discretization of

$$\frac{\partial P}{\partial t} = w\frac{\partial P}{\partial x_1}$$

of first order.

The equation (8) is advanced in time with constant, global time steps Δt bounded by the minimal spatial step $h_{\min}$ such that $\Delta t \leq ch_{\min}$ and the implicit midpoint rule of second order accuracy [8, p. 204]. The discrete version of (17) is at time $t^{n+1} = (n + 1)\Delta t$

$$\mathbf{p}_h^{n+1} = \mathbf{p}_h^n + \frac{\Delta t}{2}A(\mathbf{x}, \frac{1}{2}(t^{n+1} + t^n))(\mathbf{p}_h^{n+1} + \mathbf{p}_h^n). \tag{24}$$

Suppose that $\mathbf{p}_h^0$ is represented by its Fourier expansion at $t = 0$ with the coefficients $\hat{p}_{h\omega}^0$. Then at t_n with the midpoint rule

$$\hat{p}_{h\omega}^n = \left(\frac{1 + 0.5\Delta tG(\omega)}{1 - 0.5\Delta tG(\omega)}\right)^n \hat{p}_{h\omega}^0.$$

11

Since $\Re G \leq 0$ in (23), $|1 + 0.5\Delta t G(\omega)|/|1 - 0.5\Delta t G(\omega)| \leq 1$ and by Parseval's relation we have

$$\|\mathbf{p}_h^n\|_2 = \|\hat{\mathbf{p}}_h^n\|_2 \leq \|\hat{\mathbf{p}}_h^0\|_2 = \|\mathbf{p}_h^0\|_2$$

and the time integration of (22) is stable. A system of linear equations has to be solved in each time step for $\mathbf{p}^{n+1}$ in (24). This system is solved iteratively by the restarted GMRES method [18].

The method with $\gamma = \delta = 0$ is too inaccurate but together with the implicit midpoint rule it preserves the non-negativity of the solution. This cannot be guaranteed with the interpolation of higher accuracy in (22).

4 Block partitioning

The computational domain is partitioned into square blocks. Each block consists of a number of cells $\mathcal{C}_{IJ}^h$ with the same h but h may be different in different blocks. The cell size can jump by a factor two at a block boundary. In this way, a better resolution of the solution can be obtained in parts of the domain where it is necessary for the accuracy and a coarser grid is often sufficient in other parts. Since the smallest h is 1, the grid size will be $h = 2^j$, $j = 0, 1, \ldots$ in the blocks. Savings in computing time and computer memory without loss of accuracy are demonstrated in Section 6 when h is allowed to vary. Implementation on a parallel computer is facilitated by a block partitioning.

Each block is surrounded by ghost subcells, see Figure 5. The fluxes in these ghost cells are needed for the evaluation of the right hand side Φ in the ME (8) in the cells at the boundaries of a block. At the boundary of the domain, the fluxes are zero in the ghost subcells. This corresponds to a subcell $\Delta \mathcal{C}_{IJ}^{h+}(\mathbf{n}_r)$ and a reaction with n_{r1} or n_{r2} negative. If n_{r1} or n_{r2} is positive, then the outflow flux in some subcells is not included in Φ at the boundary cells. The subcells at a block interface are updated by data from the adjacent block.
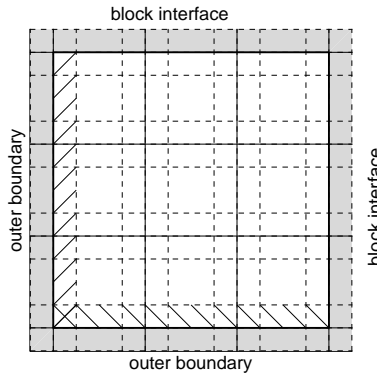

Fig. 5: Fluxes are zero in ghost subcells at the outer boundary (shaded) and sometimes in subcells adjacent to outer boundaries (crossed).

12

The values of p at the inner block interfaces, see Figure 3, are updated from the coinciding cell in the adjacent block. Then p is interpolated as in Section 3.3 to the subcells in the block. Finally, the accumulated flux from all reactions is computed and sent to the receiving cell on the other side of the block interface.

Since the same fluxes are used on both sides of the interface, the total probability P_{tot} is conserved also across the block interfaces. Let $\mathbf{a}$ be a vector of cell areas ordered in the same way as $\mathbf{p}_h$. Then $P_{\text{tot}} = \mathbf{a}^T \mathbf{p}_h$. Since P_{tot} is conserved in time

$$\frac{\partial P_{\text{tot}}}{\partial t} = \mathbf{a}^T \frac{\partial p}{\partial t} = \mathbf{a}^T A \mathbf{p}_h = 0$$

for every $\mathbf{p}_h$, $\mathbf{a}$ is a left eigenvector of A with eigenvalue zero. If A is independent of t, then the corresponding right eigenvector $\mathbf{p}_h^\infty$ is the steady state solution fulfilling $A\mathbf{p}_h^\infty = 0$ (cf. [5]). The discretization (24) has the same steady state solution.
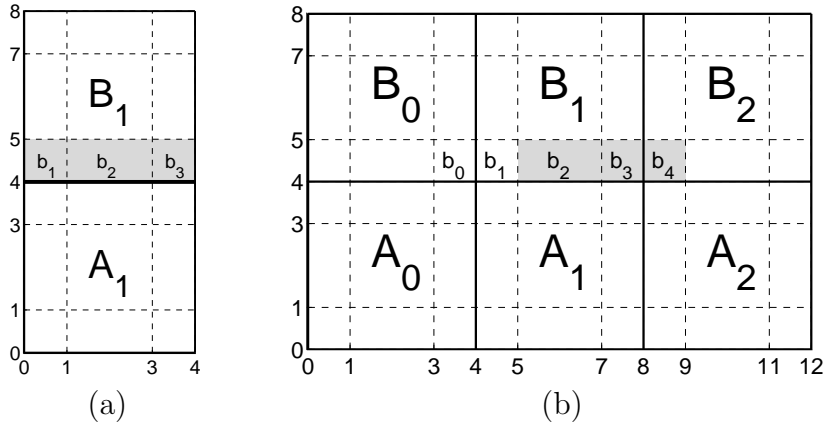

Fig. 6: The fluxes in the shaded areas are used in A_1 for reactions with $\mathbf{n}_r = (0,1)^T$ (a) and $\mathbf{n}_r = (1,1)^T$ (b).

Two examples with equal cell size in two neighboring blocks are found in Figure 6. The cells B_j belong to the upper block and the cells A_j are part of the lower block in the figure. When $\mathbf{n}_r = (0,1)^T$ in Figure 6.a, the q-values in b_1, b_2, and b_3 are determined and added to the residual in cell A_1. The same quantity is subtracted in the residual in cell B_1. In Figure 6.b, $\mathbf{n}_r = (1,1)^T$ and the sum of the fluxes in b_2, b_3, and b_4 is sent to cell A_1. If A_1 is in the upper right corner of the block, then q in b_4 is computed in the block located along the diagonal with only a corner in common with A_1. The communication between diagonal neighbors only involves the flux in a unit cell.
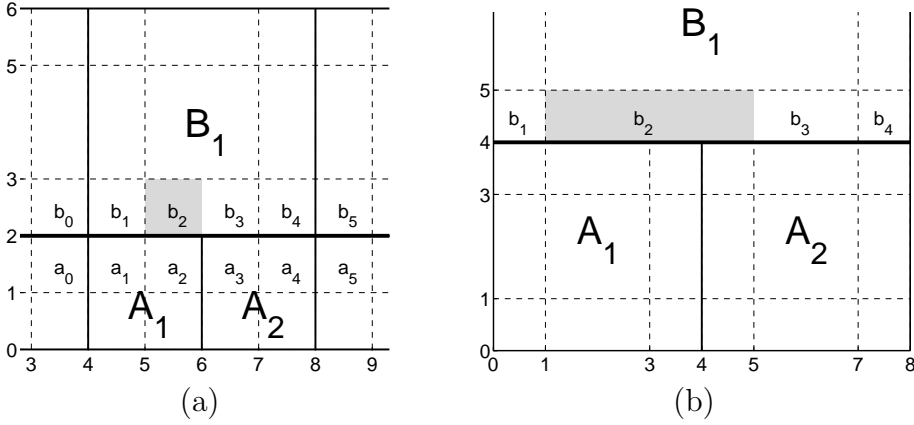
Fig. 7: Cells at block interfaces with jumps in cell size. The shaded areas illustrate cases where a subcell must be divided at the center ($\mathbf{n}_r = (0,1)^T$) (a) or at one unit length from the center ($\mathbf{n}_r = (1,1)^T$) (b).

Additional interpolation in the subcells is necessary when the cell size is doubled across a block interface. Consider the two cases in Figure 7. In the first case, Figure 7.a, $h_l = h/2 = 2$ in the lower block with cells A_1 and A_2 and $h_u = h = 4$ in the upper block with the cell B_1. The flux q_{23} in the subcell $b_{23} = b_2 \cup b_3$ is known using the procedure from Section 3.3 but when $h \geq 4$ and $\mathbf{n}_r$ has a zero component, we need q_j in the cells b_j, $j = 2,3$. This is accomplished by the interpolation

$$q_j = \frac{q_{23}}{2} + (-1)^{j+1} \frac{(h-2)^2}{8(h-1)}(q_4 - q_1), \ j = 2,3. \tag{25}$$

The total flux in b_{23} is conserved since $q_2 + q_3 = q_{23}$. The interpolation is exact for the averages of second degree polynomials. If $\mathbf{n}_r = (0,1)^T$ as in Figure 7.a, then q_1 and q_2 are the inflow fluxes to A_1 and q_3 and q_4 are the corresponding fluxes to A_2.

In the second case, Figure 7.b, b_{23} is split at $x_4 + 1$ where x_4 is the coordinate of the midpoint. Such a division is required for reactions with two nonzero components of $\mathbf{n}_r$ and $h \geq 8$. An interpolation formula is

$$q_2 = \frac{h+2}{2(h-1)}q_{23} + \frac{h-4}{8(h-1)}((h-2)q_1 - (h+2)q_4), \ q_3 = q_{23} - q_2. \tag{26}$$

It is exact for second degree polynomials and preserves the subcell flux in b_{23}. If $\mathbf{n}_r = (1,1)^T$ then q_2 is part of the inflow flux to A_1 and q_3 and q_4 contribute to the change of p in A_2.

5   Adaptation

The cell size in a block is adapted dynamically to control an estimate of the discretization error in the block. If the estimate in one of the cells in a block

14

with $h > 1$ is larger than a predefined error tolerance ε, then h is halved in all cells in the block. If the estimate is smaller than a fraction $\zeta, 0 < \zeta < 1$, of the tolerance in all cells in a block, then h is doubled in all cells. This strategy has been successful in solving partial differential equations [5, 13].

The error estimates are computed after every time step and the cell size in the block is altered if necessary. The new values of p in the smaller cells after refinement are computed by the interpolation formulas (13). When going from a fine to a coarser grid, p in the new grid is given by

$$p_{2hIJ} = (p_{hij} + p_{h,i+1,j} + p_{h,i,j+1} + p_{h,i+1,j+1})/4, \quad \mathcal{C}_{IJ}^{2h} = \bigcup_{k,l=0}^{1} \mathcal{C}_{i+k,j+l}^{h}. \tag{27}$$

In this way, both refinement and coarsening preserve P_{tot}.

The discretization error in the master operator Φ^h on the right hand side of (8) due to a cell size $h > 1$ is estimated by comparing the sum of the space discretizations or residuals in the cells $\mathcal{C}_{i+k,j+l}^{h}$ defined in (27) with the residual in $\mathcal{C}_{IJ}^{2h}$. The quantity ϑ_{IJ} in cell $\mathcal{C}_{IJ}^{2h}$ is defined by

$$\vartheta_{IJ} = \frac{1}{4} \sum_{k,l=0}^{1} \Phi_{i+k,j+l}^{h}(\mathbf{p}_h) - \Phi_{IJ}^{2h}(\mathbf{p}_{2h}), \tag{28}$$

with $\mathbf{p}_h$ and $\mathbf{p}_{2h}$ as in (27).

The discretization error τ^h is the difference between the exact value of the residual $\mathcal{F}$ in a cell $\mathcal{C}_{IJ}^{h}$ computed with an analytical p and the approximation Φ^h given the cell averages $\mathbf{p}_h$ of p (cf. [13])

$$\tau_{IJ}^{h} = \mathcal{F}_{IJ}(p) - \Phi_{IJ}^{h}(\mathbf{p}_h).$$

By the definition of ϑ in (28) and since fluxes in the interior of a cell cancel each other in $\mathcal{F}$

$$\begin{aligned}
\vartheta_{IJ} &= \frac{1}{4} \sum_{k,l=0}^{1} \mathcal{F}_{i+k,j+l}(p) - \mathcal{F}_{IJ}(p) - \frac{1}{4} \sum_{k,l=0}^{1} \tau_{i+k,j+l}^{h} + \tau_{IJ}^{2h} \\
&= \tau_{IJ}^{2h} - \frac{1}{4} \sum_{k,l=0}^{1} \tau_{i+k,j+l}^{h}.
\end{aligned} \tag{29}$$

If the dominant term in τ is $c_\tau \mathcal{Q}_\tau$ where $\mathcal{Q}_\tau$ is a polynomial in h, then by (29) $\vartheta_{IJ} \approx \mathcal{Q}_\tau(2h) - \mathcal{Q}_\tau(h)$ and

$$\tau_{IJ}^{2h} \approx \frac{\mathcal{Q}_\tau(2h)}{\mathcal{Q}_\tau(2h) - \mathcal{Q}_\tau(h)} \vartheta_{IJ}, \qquad \tau_{ij}^{h} \approx \frac{\mathcal{Q}_\tau(h)}{\mathcal{Q}_\tau(2h) - \mathcal{Q}_\tau(h)} \vartheta_{IJ}. \tag{30}$$

In one dimension, $\mathcal{Q}_\tau(h)$ is proportional to $(h-1)(2h-1)(h+1)$ in (20) and (21). If $h = 1$ then $\mathcal{Q}_\tau(h) = 0$ and $\tau_{IJ}^{2h} \approx \vartheta_{IJ}$ and $\tau_{ij}^{h} \approx 0$ in (30). If $h = 2$ then $\tau_{IJ}^{2h} \approx \vartheta_{IJ}$ and $\tau_{ij}^{h} \approx \vartheta_{IJ}/12$.

The true error in Φ is computed in 1D for a reaction with $n_r = 1$ and a constant and a linear propensity w. The dependence of this error on h is compared in Table 1 to the behavior of $\mathcal{Q}$ for increasing h for $p(x) = p_0 \exp(-x^2/50)$ scaled such that P_{tot} is one. The maximum norm of $\boldsymbol{\tau}^h$ in the interval $[-1028, 1028]$ and the quotients $Q_h = \mathcal{Q}(h)/\mathcal{Q}(h/2)$ and $T_h = |\boldsymbol{\tau}^h|_\infty/|\boldsymbol{\tau}^{h/2}|_\infty$ are computed. The agreement between Q_h and T_h is very good.

			$w = 1$		$w = x$	
h	$\mathcal{Q}(h)$	Q_h	$\|\boldsymbol{\tau}^h\|_\infty$	T_h	$\|\boldsymbol{\tau}^h\|_\infty$	T_h
1	0	-	0	-	0	-
2	9	-	2.30e-7	-	8.11e-9	-
4	105	11.7	2.70e-6	10.3	9.40e-8	11.6
8	945	9.0	2.41e-5	8.9	8.08e-7	8.6
16	7905	8.4	1.91e-4	7.9	5.93e-6	7.4

Table 1: Comparison of the true discretization errors with the theoretical estimate in one dimension.

In the experiments in the next section, $\boldsymbol{\vartheta}$ is estimated in $|\cdot|_\infty$. If $|\boldsymbol{\vartheta}|_\infty > \varepsilon$, then the maximum error in the fine grid with cell size h is about $0.1|\boldsymbol{\vartheta}|_\infty$ and the grid is refined. If $|\boldsymbol{\vartheta}|_\infty < \zeta\varepsilon$, then the error is the fine grid is approximately less than $0.1\zeta\varepsilon$ and it can be coarsened.

6 Numerical results

The algorithm for solving the low dimensional ME is applied to a model in sociology by Weidlich [23, ch. 4] for the migration of people. The time dependent solution, the number of cells, and the decay of the residual as the solution approaches the steady state are displayed in four different cases.

Assume that there are two regions $\mathcal{R}_1$ and $\mathcal{R}_2$ and two different populations $\mathcal{P}_1$ and $\mathcal{P}_2$. The populations interact with each other and move between the regions. The system is characterized at each time instant by four states S_i^j, $i, j = 1, 2$, and s_i^j is the number of people from $\mathcal{P}_i$ in region $\mathcal{R}_j$ in state S_i^j. Since the total number of people Ω_i in $\mathcal{P}_i$ is constant, the system has two degrees of freedom. Let

$$x_1 = (s_1^1 - s_1^2)/2, \ -\Omega_1 \le x_1 \le \Omega_1, \ x_2 = (s_2^1 - s_2^2)/2, \ -\Omega_2 \le x_2 \le \Omega_2.$$

Then the system is represented by the difference x_i between the number of people of $\mathcal{P}_i$ in regions $\mathcal{R}_1$ and $\mathcal{R}_2$.

The inclination of an individual in $\mathcal{P}_i$ to move from $\mathcal{R}_1$ to $\mathcal{R}_2$ (and vice versa) depends on the number of $\mathcal{P}_1$ and $\mathcal{P}_2$ in $\mathcal{R}_1$. The changes of the state of the

system can be written as follows for $\Omega = \Omega_1 = \Omega_2$:

$$
\begin{aligned}
S_1^1 & \xrightarrow{w_1} S_1^2, & w_1 &= \gamma(\Omega - x_1)\exp(\Omega^{-1}(\kappa x_1 - \sigma x_2)), & \mathbf{n}_1^T &= (-1, 0), \\
S_1^2 & \xrightarrow{w_2} S_1^1, & w_2 &= \gamma(\Omega + x_1)\exp(-\Omega^{-1}(\kappa x_1 - \sigma x_2)), & \mathbf{n}_2^T &= (1, 0), \\
S_2^1 & \xrightarrow{w_3} S_2^2, & w_3 &= \gamma(\Omega - x_2)\exp(\Omega^{-1}(-\tau\sigma x_1 + \kappa x_2)), & \mathbf{n}_3^T &= (0, -1), \\
S_2^2 & \xrightarrow{w_4} S_2^1, & w_4 &= \gamma(\Omega + x_2)\exp(-\Omega^{-1}(-\tau\sigma x_1 + \kappa x_2)), & \mathbf{n}_4^T &= (0, 1),
\end{aligned}
\tag{31}
$$

according to [23, p. 87]. An interpretation of the model is found in the same reference and in [24]. The parameters κ and σ are real and positive and $\tau = 1$ in a symmetrical case and $\tau = -1$ in an antisymmetrical case. By changing these parameters, four different scenarios are simulated in the numerical examples.

The expected values of the stochastic variables X_1 and X_2 with the PDF given by the master equation (3) approximately satisfy a system of ordinary differential equations (ODEs) in the macroscopic model [6]. This system can be analyzed as a dynamical system [21]. Three of our examples have stable fixed points which are reached when time increases. The solution enters a stable limit cycle in the final example.

The x-domain is $[-\Omega, \Omega] \times [-\Omega, \Omega]$ instead of $[0, x_{\max}] \times [0, x_{\max}]$ as in Section 3 but the boundary conditions are the same. The initial Gaussian density $p(\mathbf{x}, 0)$ is on the diagonal between $(-\Omega, -\Omega)$ and (Ω, Ω) of the (x_1, x_2)-domain in the four main scenarios. The solution is integrated from $t = 0$ on a grid with 5×5 blocks. Isolines of the solution are plotted in the figures. Every second grid line is shown. Three grid levels are used with $h = 1$ at the finest level and $\varepsilon = 1.1 \cdot 10^{-3}$ and $\zeta = 0.1$ in Scenarios 1, 2, and 3. With $h = 1$ everywhere as in a direct solution of the ME, the number of cells would be 14400.
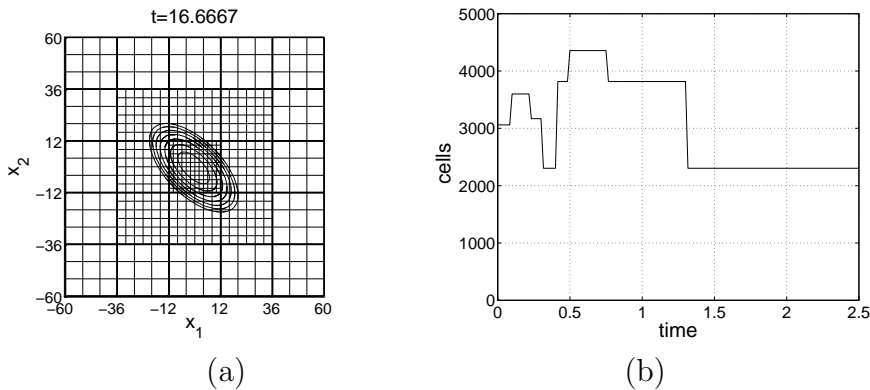
6.1 Scenario 1

(a)   (b)

Fig. 8: The steady state solution of Scenario 1 (a). The time evolution of the number of cells (b).

With the parameters $\kappa = 0.2, \sigma = 0.5$, and $\tau = 1$ the origin is a stable stationary point of the macroscopic problem [23, p. 89]. There is a homogeneous mixture of both populations in both regions in the steady state with a peak in the PDF at $s_1^1 = s_2^1 = s_1^2 = s_2^2$ and $x_1 = x_2 = 0$.

The computing time for $p_a(\mathbf{x}, 17)$ on the adapted grid is 59 percent of the time to determine $p_1(\mathbf{x}, 17)$ when $h = 1$ everywhere. The number of time steps to reach $t = 17$ is almost the same in both cases. For comparison, the solution with $h = 1$ is transfered to the adapted grid $p_1 \rightarrow \tilde{p}_1$ by a restriction as in (27). The relative difference between the solutions defined as

$$\delta p = |\mathbf{p}_a - \tilde{\mathbf{p}}_1|_\infty / |\mathbf{p}_1|_\infty \tag{32}$$

is here $\delta p = 0.0044$. The final grid for p_a consists of 2304 cells. The memory requirements are reduced by more than 80 percent compared to the case with $h = 1$. The steady state solution is displayed in Figure 8.a. The residual Φ measured in $| \cdot |_\infty$ is there about 10^{-10}. The total probability P_{tot} is almost perfectly conserved: the difference between the largest and the smallest P_{tot} is $1.5 \cdot 10^{-12}$ in the time interval.

The variation of the number of cells in the beginning of the simulation is found in Figure 8.b. After about $t = 1.3$ the number is constant.

6.2   Scenario 2

In this scenario, the parameters are $\kappa = 0.5, \sigma = 1, \tau = 1$. The majority of the populations $\mathcal{P}_1$ and $\mathcal{P}_2$ prefers a region $\mathcal{R}_j$ where there are only a few members of the other population, i.e. it is likely that $s_1^1 > s_1^2, s_2^2 > s_2^1$ or $s_2^1 > s_1^1, s_2^1 > s_2^2$. Two different cases are considered: One case with the initial density symmetrically located on the diagonal in Figure 9 and one case with an asymmetrically placed density at $t = 0$ in Figure 10.
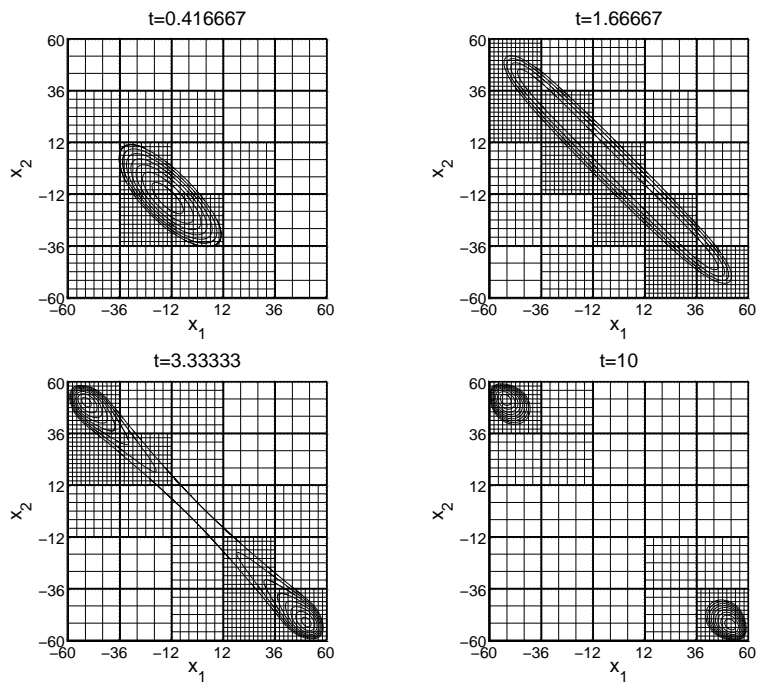
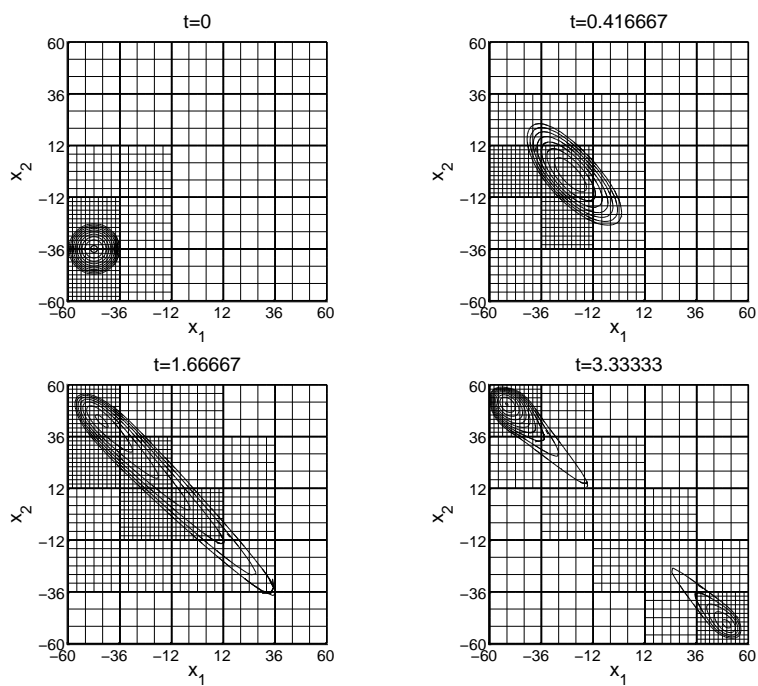Fig. 9: Time evolution of the PDF in scenario 2 with a symmetric p on the diagonal at $t = 0$.

Fig. 10: Time evolution of the PDF in scenario 2 with a p above the diagonal at $t = 0$.

19

The symmetric case in Figure 9 is integrated to $t = 10$. Then δp in (32) is 0.000055, the savings in the number of cells at the steady state is more than 80 percent compared to a uniform grid with $h = 1$, the number of time steps is reduced from 1201 to 1065, and the CPU time decreases by 45 percent. Fewer time steps are needed in the adapted solution since $h_{\min} = 2$ in parts of the time interval. The asymmetry in the initial density is also present after long simulation times, see Figure 10. The maximum difference in P_{tot} is here $4 \cdot 10^{-15}$ in both cases.
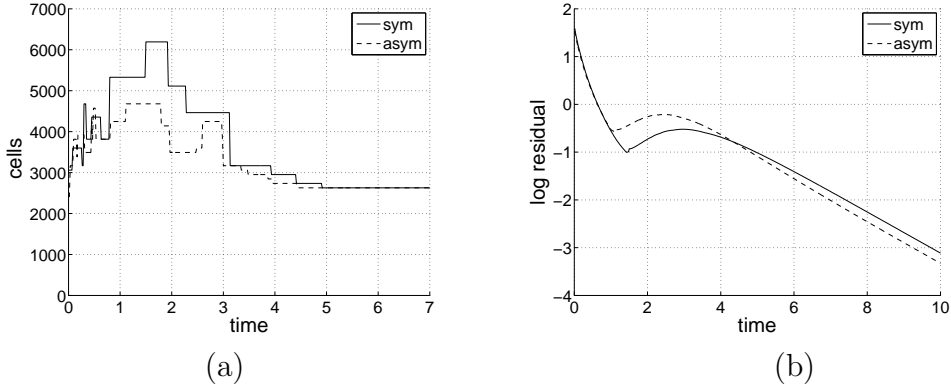

Fig. 11: The time series of the number of cells (a) and the residual Φ (b) for the two initializations of the PDF.

6.3   Scenario 3

There is an asymmetrical interaction between the populations in Scenario 3 with $\kappa = 0.5, \sigma = 1$, and $\tau = -1$. The population $\mathcal{P}_1$ tries to live apart from $\mathcal{P}_2$, but $\mathcal{P}_2$ wants to live together with $\mathcal{P}_1$. The difference from Scenario 1 is that the expected values follow spirals toward the origin which is a stable fixed point in the macroscopic model. The behavior of the algorithm in this example is very similar to the behavior in Scenario 1.
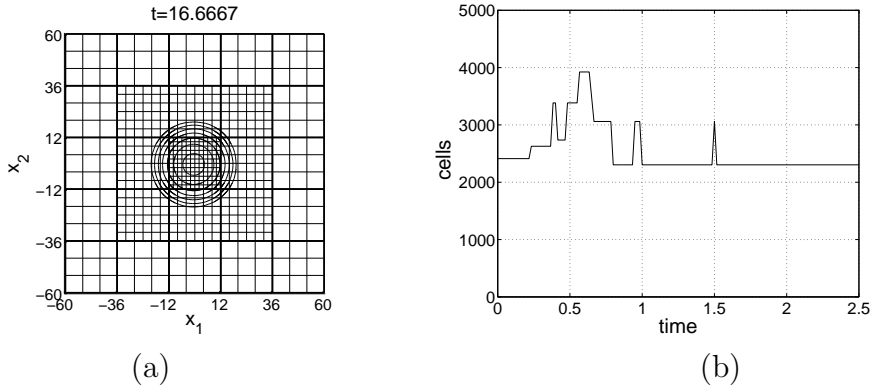

Fig. 12: Scenario 3: The final solution (a). The time series of the number of cells (b).

20

6.4 Scenario 4

The expected values of X_1 and X_2 approach a limit cycle in the final migration example. The solution rotates in the counterclockwise direction about the origin which is an unstable fixed point of the dynamical system. With the parameters $\kappa = 1.2, \sigma = 1$, and $\tau = -1$, the $\mathcal{P}_2$ population follows $\mathcal{P}_1$ which tries to avoid living together with $\mathcal{P}_2$. If both populations are in the upper right quadrant initially in region $\mathcal{R}_1$, then population $\mathcal{P}_1$ moves to $\mathcal{R}_2$ and the upper left quadrant. Population $\mathcal{P}_2$ follows to $\mathcal{R}_2$ and the density moves into the lower left quadrant, but then $\mathcal{P}_1$ returns to $\mathcal{R}_1$ in the lower right quadrant. When $\mathcal{P}_2$ is back in $\mathcal{R}_1$ and $x_1 > 0, x_2 > 0$, one period of the solution is complete.
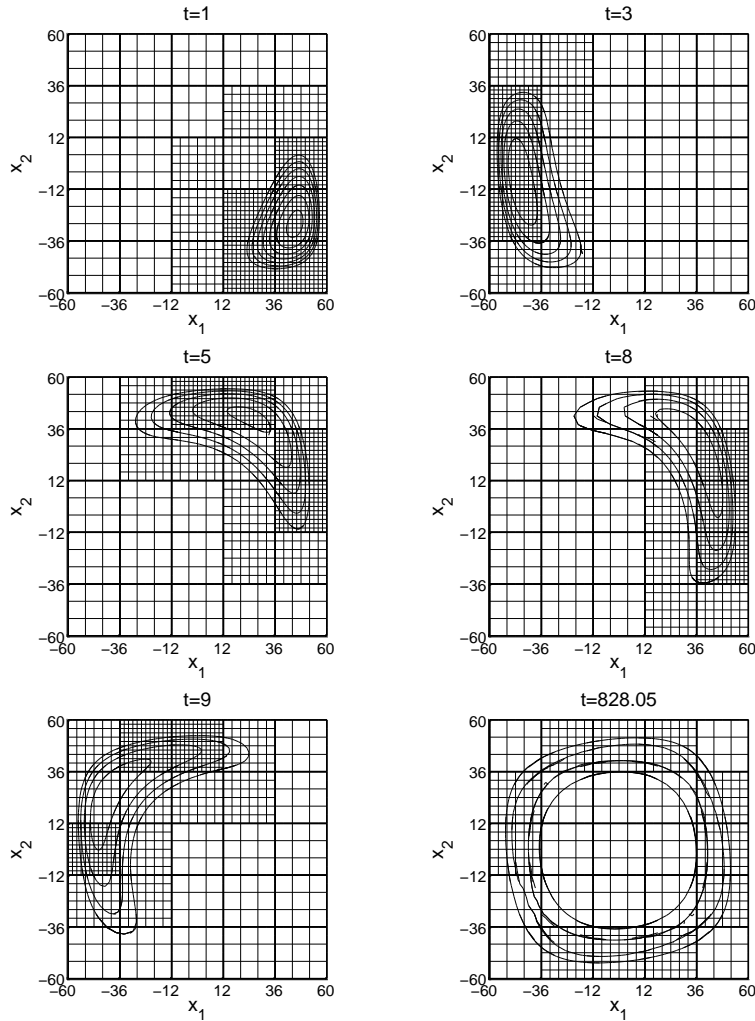

Fig. 13: Solutions of Scenario 4.

The adaptation parameters are $\varepsilon = 2.2 \cdot 10^{-4}$ and $\zeta = 0.05$. Snapshots of the solution are displayed in Figure 13. The initial PDF in the lower left corner

21

rotates and is elongated as time increases. After long time, the residual Φ and the time derivatives vanish (Figure 14.b) and the center of the PDF solution is close to the limit cycle of the approximate dynamical system of the expected values. The total probability is conserved almost perfectly with a variation of about 10^{-12} and the memory requirements are reduced by more than 85 percent at the steady state compared to a direct solution of the ME, see Figure 14.a. The difference δp at $t = 500$ in (32) is 0.021 and the CPU time with adaptation is only about $1/3$ of the time for $h = 1$. Another advantage of the adaptive algorithm is that the time steps are longer when the solution approaches the steady state, see Figure 13, since the cell size of the smallest cells $h_{\min}$ is larger. The number of time steps is reduced from 30000 to 15371.
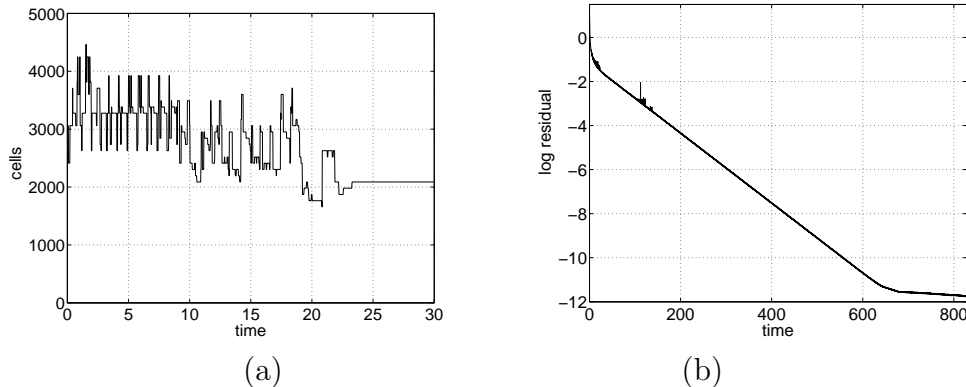

(a)   (b)

Fig. 14: The number of cells (a) and the residual Φ (b) in Scenario 4.

7   Conclusions

A method has been developed for integration of the master equation (ME) suitable for low dimensional problems. The original equation is formulated for cells in space with the length of the edge $h = 1$. The ME is rewritten for cells with $h > 1$ and the average of the PDF is determined in these cells in a way similar to a finite volume discretization of a conservation law. The possibility to use large cells saves memory in particular in problems with large state spaces of the order $10^3 - 10^4$ in each dimension. The computational domain is partitioned into blocks and h is constant in each block. Interpolation is necessary for evaluation of the inflow and outflow of probability in a cell. The accuracy and stability of the interpolation are investigated. The grid is dynamically adapted to the solution and the adaptation is controled by an estimate of the error in the solution due to the large cells. The algorithm is applied to four different examples from the modeling of the migration of people. The relative improvements in computing time, memory requirements, and time steps in three different cases are summarized in Table 2. In Scenario 4, the CPU time is reduced by 65 percent, the storage by

22

85 percent, and the number of time steps by 49 percent.

Scenario	CPU time	Memory	Time steps
1	0.59	0.16	0.99
2	0.55	0.18	0.89
4	0.35	0.15	0.51

Table 2: Quotients for different quantities between the adaptive method with $h \geq 1$ and the direct method with $h = 1$.

References

[1] H.-J. Bungartz, M. Griebel, Sparse grids, *Acta Numerica*, 2004, p. 147–269.

[2] U. Dieckmann, P. Marrow, R. Law, Evolutionary cycling in predator-prey interactions: Population dynamics and the Red Queen, *J. Theor. Biol.*, 176 (1995), p. 91–102.

[3] J. Elf, J. Paulsson, O. G. Berg, M. Ehrenberg, Near-critical phenomena in intracellular metabolite pools, *Biophys. J.*, 84 (2003), p. 154–170.

[4] S. Engblom, A discrete spectral method for the chemical master equation, Technical report 2006-036, Dept of Information Technology, Uppsala University, Uppsala, Sweden, 2006, available at http://www.it.uu.se/research/reports/2006-036/.

[5] L. Ferm, P. Lötstedt, P. Sjöberg, Conservative solution of the Fokker-Planck equation for stochastic chemical reactions, *BIT*, 46 (2006), S61–S83.

[6] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed., Springer, Berlin, 2002.

[7] D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comput. Phys.*, 22 (1976), p. 403–434.

[8] E. Hairer, S. P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations I, Nonstiff Problems*, 2nd ed., Springer-Verlag, Berlin, 1993.

[9] M. Hegland, C. Burden, L. Santoso, S. MacNamara, H. Booth, A solver for the stochastic master equation applied to gene regulatory networks, *J. Comput. Appl. Math.*, 205 (2007), p. 708–724.

[10] A. Hellander, P. Lötstedt, Hybrid method for the chemical master equation, doi:10.1016/j.jcp.2007.07.020, to appear in *J. Comput. Phys.*.

[11] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1992.

[12] P. Lötstedt, L. Ferm, Dimensional reduction of the Fokker-Planck equation for stochastic chemical reactions, *Multiscale Meth. Simul.*, 5 (2006), p. 593–614.

[13] P. Lötstedt, S. Söderberg, A. Ramage, L. Hemmingsson-Fränden, Implicit solution of hyperbolic equations with space-time adaptivity, *BIT*, 42 (2002), p. 134–158.

[14] S. Mac, K. Burrage, R. B. Sidje, Multiscale modeling of chemical kinetics via the master equation, to appear in *Multiscale Meth. Simul.*

[15] H. H. McAdams, A. Arkin, Stochastic mechanisms in gene expression, *Proc. Natl. Acad. Sci. USA*, 94 (1997), p. 814–819.

[16] A. J. McKane, T. J. Newman, Stochastic models in population biology and their deterministic analogs, *Phys. Rev. E*, 70 (2004), 041902.

[17] J. Paulsson, J. Elf, Stochastic modeling of intracellular kinetics, in *System Modeling in Cellular Biology*, eds. Z. Szallasi, J. Stelling, V. Periwal, MIT Press, Cambridge, MA, 2006, p. 149–175.

[18] Y. Saad, M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Stat. Comput.*, 7 (1986), p. 856–869.

[19] P. Sjöberg, P. Lötstedt, J. Elf, Fokker-Planck approximation of the master equation in molecular biology, doi:10.1007/s00791-006-0045-6, to appear in *Comput. Visual. Sci.*.

[20] N. Stollenwerk, V. A. A. Jensen, Meningitis, pathogenicity near criticality: the epidemiology of meningococcal disease as a model for accidental pathogens, *J. Theor. Biol.*, 222 (2003), p. 347–359.

[21] S. H. Strogatz, *Nonlinear Dynamics and Chaos,* Perseus Books, Cambridge, MA, 1994.

[22] M. Thattai, A. van Oudenaarden, Intrinsic noise in gene regulatory networks, *Proc. Nat. Acad. Sci. USA*, 98 (2001), p. 8614–8619.

[23] W. Weidlich, *Sociodynamics. A Systematic Approach to Mathematical Modelling in the Social Sciences*, Taylor and Francis, London, 2002.

[24] W. WEIDLICH, Thirty years of sociodynamics. An integrated strategy of modelling in the social sciences: applications to migration and urban evolution, *Chaos, Solit. Fract.*, 24 (2005), p. 45–56.