PDE and Monte Carlo approaches to solving the master equation applied to gene regulation

Paul Sjöberg

Division of Scientific Computing, Department of Information Technology,
Uppsala University, P.O. Box 337, SE-75105 Uppsala, Sweden.
Paul.Sjoberg@it.uu.se

Abstract. The *Fokker-Planck equation* (FPE) approximation is applied to a subspace of the state space of the *chemical master equation* (CME). The CME-FPE-hybrid method exploits the lower cost of the FPE approximation compared to the full CME. A fourth order finite difference approximation of the FPE part of the hybrid is described and demonstrated on a biologically relevant model in five dimensions.

1 Introduction

In stochastic models of molecular control circuits the the probability for the internal state of the cell model is governed by the *chemical master equation* (CME) [16]. The model is often high-dimensional, since each molecular species adds a dimension to the state space of the CME. The number of dimensions for problems where it is possible to compute a solution to the CME directly is limited. Even if computers get faster and use more and more CPUs in parallel the impact is dwarfed by the exponential growth in problem size with increasing number of molecular species in the model.

It may be fruitful to study model systems in a few dimensions, but the complexity of biochemical pathways and gene regulation will eventually require high-dimensional models. Widely differing scales in copy numbers and frequency of reaction events in biochemical systems will make different approximations suitable for different parts of the model. Such hybrid methods are being developed for different approaches to solve the master equation [1] [2] [7] [8] [10].

One approximation that is close at hand is to approximate the CME with a *partial differential equation* (PDE), the *Fokker-Planck equation* (FPE). The computational performance is problem dependent for both FPE solvers and competing Monte Carlo methods such as [3] and either can be more efficient for a certain problem [13].

The FPE approximation is based on a continuous approximation of the state space, which may be inappropriate in some dimensions, for instance if the dimension can be represented using very few discrete states. Dimensions which are so small that approximation will not make the problem more tractable will here be called *shallow*. What is a shallow dimension depends on the method of approximation. A mean field approximation always reduces problem size, while

the discrete FPE approximation in exceptional cases requires a resolution that exceeds the number of states that are approximated.

Here, a hybrid method is used for approximation of a problem with shallow dimensions. The state space is split into a subspace in the shallow dimensions and a subspace that is well-suited for FPE approximation. This division of the state space enables an operator splitting and derivation of a CME-FPE-hybrid.

Genes are regulated by molecules bound to regulatory regions. Simply put, the regulators bound to DNA determines the transcription rate of the gene. Binding configurations of genes will be referred to as *gene modes*, which in this context is equivalent to a separate molecular species. An obvious application for the hybrid method is adding gene modes to a low-dimensional model of the gene product turnover that is suitable for FPE approximation. The gene modes add dimensions to the state space, but often dimensions with very limited size.

The paper is organized as follows. The splitting of the master operator is described and a fourth order accurate, conservative discretization of the FPE is derived. An alternative is a Monte Carlo method using the *Stochastic Simulation Algorithm* (SSA) by Gillespie [3]. The methods are compared in extensive numerical experiments in the final section.

The supremum norm for a vector v of length N is denoted $\|v\|_\infty$ and the ℓ_1-norm $\|v\|_1 = \sum_{i=1}^{N} |v_i|$. The normal distribution with mean value μ and variance σ^2 is denoted $\mathcal{N}(\mu, \sigma^2)$. The integer numbers are denoted by $\mathbb{Z}$. The copy number of a molecular species is in the set of non-negative integers $\mathbb{Z}_+ = \{0, 1, \ldots\}$.

2 The master and Fokker-Planck equations

The characterizing feature of Markov processes is lack of memory. Modelling biochemical reactions as Markov processes means that only the current state may influence the probability for a reaction event. In practice that is always an approximation which must be justified. For biochemical reactions in living cells the prerequisites for the validity of the assumption is treated e.g. in [16].

The master equation is a differential-difference equation for the time evolution of the probability distribution of the states of a Markov process. Although diffusion in the cell volume is relevant it is common to study a volume that is homogeneous in the the sense that the probability of reactive collision between any pair of reactant molecules is equal and only depends on the reactor volume. When this assumption holds, the reactor is said to be well stirred. In a cell it is reasonable to consider the volume well stirred if there are many inert collisions between reactive events.

The state of a system is defined by the copy number of the different molecular species and a constant volume. Adding a variable volume to the methods presented is straightforward.

Definition 1. *Let the system of volume V contain N chemical species X_i, $i = 1 \ldots N$ and R reaction channels Π_r, $r = 1 \ldots R$. The number of X_i-molecules is denoted by $x_i \in [0, \infty)$ and the state of the system is determined by* $\mathbf{x} =$

$(x_1, x_2, \ldots, x_N)^T$. *Define the propensity* $w_r(\mathbf{x})$ *as the probability for reaction* Π_r *to occur in the next infinitesimal time interval.*

The reaction Π_r is specified by a *reaction propensity* $w_r(\mathbf{x})$ and the distance $\mathbf{n}_r = (n_{r,1}, n_{n,2}, \ldots, n_{r,N})^T$ in the state space from the state after to the state before the reaction. That is a Π_r reaction event will change the state from $\mathbf{x}+\mathbf{n}_r$ to $\mathbf{x}$. The propensity depends on the probability of collision and the probability for reaction given that a collision has occurred. The latter probability depends on the velocity distribution of the reactant molecules (hence the temperature), the activation energy of the reaction, the probability of the molecules to be positioned in a reactive orientation and so on.

A reaction is written as an arrow with the reacting molecules to the left and the product molecules to the right. Above the arrow is the reaction propensity specified. For example the dimerization of A to B would be written

$$2A \xrightarrow{k \cdot a(a-1)} B,$$

where k is a constant and a and b are the copy numbers of A and B respectively.

Consider a system as defined in Definition 1. Summing over all reactions we now can write the CME [16]:

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

There is no probability of having negative copy numbers, that is for any $\mathbf{x}_{ext} \in \mathbb{Z}^N \backslash \mathbb{Z}_+^N$, $p(\mathbf{x}_{ext}, t) = 0$ which implies $w_r(\mathbf{x}_{ext}+\mathbf{n}_r) = 0$. Let $q_r(\mathbf{x}, t) = w_r(\mathbf{x}) p(\mathbf{x}, t)$. By Taylor expansion of the CME and truncation after the second order terms the following FPE is obtained [16] using ∂_{x_i} to denote partial derivation in dimension x_i,

$$\begin{aligned}
\frac{\partial p(\mathbf{x}, t)}{\partial t} &= \sum_{r=1}^{R} \left(\sum_{i=1}^{N} n_{r,i} \partial_{x_i} q_r(\mathbf{x}, t) + \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{n_{r,i} n_{r,j}}{2} \partial_{x_i} \partial_{x_j} q_r(\mathbf{x}, t) \right) \\
&= \sum_{r=1}^{R} \partial_{x_i} (G_{r,i}(\mathbf{x}, t) + 1/2 H_{r,i}(\mathbf{x}, t)) = \sum_{r=1}^{R} \partial_{x_i} F_{r,i}(\mathbf{x}, t) \\
&\equiv \sum_{r=1}^{R} A_{r,FPE}(\mathbf{x})\, p(\mathbf{x}, t),
\end{aligned} \tag{2}$$

where

$$F_{r,i}(\mathbf{x}, t) = G_{r,i}(\mathbf{x}, t) + 0.5 H_{r,i}(\mathbf{x}, t),$$

and

$$G_{r,i}(\mathbf{x}, t) = n_{r,i} \partial_{x_i} q_r,\ H_{r,i}(\mathbf{x}, t) = \sum_j n_{r,i} n_{r,j} \partial_{x_i} \partial_{x_j} q_r(\mathbf{x}, t).$$

The boundary condition is that the probability current $F_{r,i} = 0$ if $x_i = 0$.

The problem size grows exponentially with the number of dimensions, but the FPE approximation can be solved computationally using considerably fewer degrees of freedom than the CME. The SSA for simulation of trajectories through the state space of chemical reaction networks was not originally intended for solving the master equation but is still the standard method for this in computational molecular biology. SSA converges to the exact solution of the master equation and is simple to implement. The convergence properties are poor, but there are several attempts to remedy this but they usually either introduce approximations, complications in the implementation or both [14]. Here, the direct SSA method [3] is used for comparison to the FPE-approximations of the CME.

3 The hybrid of discrete CME and continuous FPE

Biochemical reactions operate in very different scales with respect to copy numbers and reaction times. It is not uncommon that some approximation that is good for some components is poor for the others. The challenge is to develop methods that split the problem into parts where at least one can be approximated in an efficient manner.

Let Y be the subspace constituted by the set of N_Y variables that is suitable for FPE approximation. X is the subspace constituted by the set of the remaining N_X variables. As a consequence the vector $\mathbf{n_r}$ is split into the vectors μ_r and η_r with elements corresponding to the dimensions in X and Y, respectively. The CME (1) for a splitting where $\mathbf{x} \in X$ and $\mathbf{y} \in Y$ is

$$\frac{\partial p}{\partial t}(\mathbf{x}, \mathbf{y}, t) = \sum_{r=1}^{R} q_r(\mathbf{x} + \mu_r, \mathbf{y} + \eta_r, t) - \sum_{r=1}^{R} q_r(\mathbf{x}, \mathbf{y}, t).$$

Add and subtract $q_r(\mathbf{x} + \mu_r, \mathbf{y}, t)$ to the equation and apply the approximation to the variables in Y

$$\begin{aligned}\frac{\partial p}{\partial t}(\mathbf{x}, \mathbf{y}, t) = \sum_{r=1}^{R} & q_r(\mathbf{x} + \mu_r, \mathbf{y} + \eta_r, t) - q_r(\mathbf{x} + \mu_r, \mathbf{y}, t) + \\ & \sum_{r=1}^{R} q_r(\mathbf{x} + \mu_r, \mathbf{y}, t) - q_r(\mathbf{x}, \mathbf{y}, t) \\ \approx & \sum_{r=1}^{R} A_r p(\mathbf{x} + \mu_r, \mathbf{y}, t) + \sum_{r=1}^{R} q_r(\mathbf{x} + \mu_r, \mathbf{y}, t) - q_r(\mathbf{x}, \mathbf{y}, t), \quad (3)\end{aligned}$$

where A_r is an operator that approximates the master equation in the subspace spanned by Y. For the FPE approximation (i.e. $A_r = A_{r,FPE}$, see (2)) the

approximate equation is

$$\frac{\partial p}{\partial t}(\mathbf{x},\mathbf{y},t) = \sum_{r=1}^{R}\left(\sum_{i=1}^{N}\eta_{r,i}\partial_{y_i}q_r + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\eta_{r,i}\eta_{r,j}\partial_{y_i}\partial_{y_j}q_r\right)(\mathbf{x}+\mu_r,\mathbf{y},t) + \sum_{r=1}^{R} q_r(\mathbf{x}+\mu_r,\mathbf{y},t) - q_r(\mathbf{x},\mathbf{y},t). \quad (4)$$

The condition on w near the boundary for the master equation remains for the X-subspace. An estimate of the difference in the solutions of the CME and the FPE is found in [13].

The hybrid methods developed for SSA mainly focus on time scale separation [1] [7] which is close at hand since unimproved SSA suffers from oversampling of the fast time scales. For PDE approximations it is important to reduce the size of the state space so it is natural to use scales in copy numbers to reduce the problem. Many methods are mainly concerned with decreasing the dimension of the problem which benefits both SSA and PDE approaches [2] [8] [10].

4 Discretization of the Fokker-Planck Equation

4.1 Space discretization of fourth order accuracy

Let $\nu = (\nu_1, \ldots, \nu_N)$ be a multi-index of dimension N and $\mathbf{x}_\nu = (x_{1,\nu_1}, \ldots, x_{N,\nu_N})^T$. The shift operator $\mathbb{E}_i^n$ is defined by $\mathbb{E}_i^n f(\mathbf{x}_\nu) = f(x_{1,\nu_1}, \ldots, x_{i,\nu_i+n}, \ldots, x_{N,\nu_N})$. The state space is truncated at high copy numbers in each dimension to form the computational domain Ω for the FPE. Ω is discretized using M_i points in dimension i so that the discretized domain is

$$\Omega_h = \{\mathbf{x}_\nu | \forall i, x_{i,\nu_i} = (\nu_i - 1/2)h_i, \nu_i = 1, \ldots, M_i\}.$$

It will also be useful to define the spaces

$$\Omega_{h,d} = \{\mathbf{x}_\nu | \forall i, x_{i,\nu_i} = (\nu_i - 1/2)h_i, \nu_i = d, \ldots, M_i + 1 - d\}.$$

That is, $\Omega_h = \Omega_{h,1}$ and the set of the boundary nodes is $\partial\Omega_h = \Omega_{h,1} \setminus \Omega_{h,2}$.

The space operator in (2) $\sum_r \sum_i \partial_{x_i} F_{r,i}(\mathbf{x})$ is discretized by M_Ω gridpoints on an N-dimensional grid with the step size h_i in dimension i. Using the shift operator, the following difference operators are defined for the interior $\Omega_{h,2}$

$$\begin{aligned} D_1^i &= D_{1/2}^{i,1} - D_{-1/2}^{i,1} \\ D_2^i &= D_{1/2}^{i,2} - D_{-1/2}^{i,2}, \end{aligned} \quad (5)$$

where

$$\begin{aligned} D_{1/2}^{i,1} &= \tfrac{1}{h_i}\left(\tfrac{-1}{12}\mathbb{E}_i^{-1} + \tfrac{7}{12}\mathbb{E}_i^0 + \tfrac{7}{12}\mathbb{E}_i^1 - \tfrac{1}{12}\mathbb{E}_i^2\right) \\ D_{-1/2}^{i,1} &= \tfrac{1}{h_i}\left(\tfrac{-1}{12}\mathbb{E}_i^{-2} + \tfrac{7}{12}\mathbb{E}_i^{-1} + \tfrac{7}{12}\mathbb{E}_i^0 - \tfrac{1}{12}\mathbb{E}_i^1\right) \\ D_{1/2}^{i,2} &= \tfrac{1}{h_i^2}\left(\tfrac{1}{12}\mathbb{E}_i^{-1} - \tfrac{5}{4}\mathbb{E}_i^0 + \tfrac{5}{4}\mathbb{E}_i^1 - \tfrac{1}{12}\mathbb{E}_i^2\right) \\ D_{-1/2}^{i,2} &= \tfrac{1}{h_i^2}\left(\tfrac{1}{12}\mathbb{E}_i^{-2} - \tfrac{5}{4}\mathbb{E}_i^{-1} + \tfrac{5}{4}\mathbb{E}_i^0 - \tfrac{1}{12}\mathbb{E}_i^1\right). \end{aligned}$$

Now, introduce the gridfunction $w_{r,\nu} = w_r(\mathbf{x}_\nu)$ and $p_\nu(t) = p(\mathbf{x}_\nu, t)$ on Ω_h. The space operator is approximated using

$$\nabla \cdot \mathbf{G}_r(\mathbf{x}_\nu, t) = \sum_{i=1}^{N} \partial_{x_i} G_{r,i}(\mathbf{x}_\nu, t) \approx \sum_{i=1}^{N} n_{r,i} D_1^i w_{r,\nu} p_\nu(t)$$

and

$$\nabla \cdot \mathbf{H}_r(\mathbf{x}_\nu, t) = \sum_{i=1}^{N} \partial_{x_i} H_{r,i}(\mathbf{x}_\nu, t) \approx \sum_{i=1}^{N} n_{r,i}^2 D_2^i w_{r,\nu} p_\nu(t) + \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} n_{r,i} n_{r,j} D_1^i D_1^j w_{r,\nu} p_\nu(t).$$

This is a fourth order accurate scheme. By Taylor expansion the error in the difference approximations of the derivatives of a sufficiently smooth function $u = u(x, y)$ can be determined:

$$\begin{gathered} \|u_x - D_1^1 u(x_i, y_j)\|_\infty \leq \tfrac{1}{30} h_x^4 \|u_{xxxxx}\|_\infty, \\ \|u_{xx} - D_1^2 u(x_i, y_j)\|_\infty \leq \tfrac{1}{90} h_x^4 \|u_{xxxxxx}\|_\infty, \\ \|u_{xy} - D_1^1 D_2^1 u(x_i, y_j)\|_\infty \leq \tfrac{1}{30} h_x^4 \|u_{xxxxxy}\|_\infty + \tfrac{1}{30} h_y^4 \|u_{xyyyyy}\|_\infty. \end{gathered} \tag{6}$$

In conclusion,

$$\frac{\partial p(\mathbf{x}_\nu, t)}{\partial t} = \sum_{r=1}^{R} \Bigg\{ \sum_{i=1}^{N} n_{r,i} D_1^i w_{r,\nu} p_\nu(t) + \sum_{i=1}^{N} n_{r,i}^2 D_2^i w_{r,\nu} p_\nu(t) + \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} n_{r,i} n_{r,j} D_1^i D_1^j w_{r,\nu} p_\nu(t) \Bigg\} + \sum_{i=1}^{N} \sum_{j=1}^{N} O(h_i^4 + h_j^4). \tag{7}$$

The fourth order stencil (7) will require a set of ghost points $\partial\Omega_{h,0} = \Omega_{h,0} \setminus \Omega_h$ outside Ω_h for discretization of the equation in $\Omega_{h,2}$. The solution there is constructed by first order extrapolation. For example if $\mathbf{x}_\nu \in \partial\Omega_{h,0}$ and ν has one index $\nu_j = 0$ and the remaining $\nu_k \in [1, M_k]$:

$$p(\mathbf{x}_\nu, t) = 2\mathbb{E}_j^1 p_\nu(t) - \mathbb{E}_j^2 p_\nu(t).$$

If there are two indices satisfying $\nu_i = \nu_j = 0$ the extrapolation is applied in both directions, that is

$$p(\mathbf{x}_\nu, t) = 4\mathbb{E}_i^1 \mathbb{E}_j^1 p_\nu(t) - 2\mathbb{E}_i^1 \mathbb{E}_j^2 p_\nu(t) - 2\mathbb{E}_i^2 \mathbb{E}_j^1 p_\nu(t) + \mathbb{E}_i^2 \mathbb{E}_j^2 p_\nu(t),$$

and correspondingly for ghost points at other boundaries.

As a comparison, the second order finite volume discretization described in [13] will be used. The errors in this discretization considered as a finite difference approximation are

$$\begin{gathered} \|u_x - \hat{D}_1^1 u(x_i, y_j)\|_\infty \leq \tfrac{1}{6} h_x^2 \|u_{xxx}\|_\infty, \\ \|u_{xx} - \hat{D}_1^2 u(x_i, y_j)\|_\infty \leq \tfrac{1}{12} h_x^2 \|u_{xxxx}\|_\infty, \\ \|u_{xy} - \hat{D}_{1,2}^2 u(x_i, y_j)\|_\infty \leq \tfrac{2}{3} h_x^2 \|u_{xxxy}\|_\infty + \tfrac{2}{3} h_y^2 \|u_{xyyy}\|_\infty, \end{gathered} \tag{8}$$

where $\hat{D}_1^1$, $\hat{D}_1^2$ and $\hat{D}_{1,2}^2$ are the discrete operators approximating the derivatives of u.

4.2 Conservation

Let the grid define a division of computational domain into subvolumes by centering a subvolume around each gridpoint with the width h_i in dimension i. The subvolumes are usually referred to as *cells* in numerical analysis, but for obvious reasons we will make the distinction *computational cells*. The discrete total probability P is defined by

$$P(t) = \sum_{\mathbf{x}_\nu \in \Omega_h} \omega p_\nu(t),$$

where ω is the volume of the computational cells, $\omega = \prod_i h_i$.

For each reaction, the contribution in the right hand side of (7) from each dimension i can be written, see (5),

$$D^i_{\nu_i+1/2} - D^i_{\nu_i-1/2},$$

$$D_{\nu_i+\alpha} = n_{r,i} D^{i,1}_\alpha q_\nu + n^2_{r,i} D^{i,2}_\alpha q_\nu + \sum_{j=1}^{N} n_{r,i} n_{r,j} D^{i,1}_\alpha D^{j,1}_\alpha q_\nu,$$

for $\alpha = 1/2$, $-1/2$. At the lower and upper boundaries of dimension i we have

$$D^i_{1/2} = 0 \quad D^i_{M_i+1/2} = 0.$$

Hence, the contribution from i to P is

$$\sum_{\nu_i}^{M_i} \frac{dp^i_\nu(t)}{dt} = \sum_{\nu_i}^{M_i} D^i_{\nu_i+1/2} - D_{\nu_i-1/2} = 0.$$

Since

$$p_\nu(t) = \sum_{i=1}^{N} p^i_\nu(t),$$

for the total probability of one reaction

$$\begin{aligned}\frac{dP(t)}{dt} = \sum_{\nu_1=1}^{M_1} \cdots \sum_{\nu_N}^{M_N} \omega \frac{dp_\nu(t)}{dt} &= \omega \sum_{i=1}^{N} \sum_{\nu_1=1}^{M_1} \cdots \sum_{\nu_N}^{M_N} \frac{dp^i_\nu(t)}{dt} \\ &= \omega \sum_{i=1}^{N} \sum_{\nu_1=1}^{M_1} \cdots \sum_{\nu_{i-1}=1}^{M_{i-1}} \sum_{\nu_{i+1}=1}^{M_{i+1}} \cdots \sum_{\nu_N=1}^{M_N} \sum_{\nu_i=1}^{M_i} \frac{dp^i_\nu(t)}{dt} = 0\end{aligned}$$

and P is conserved for the sum of the reactions by the discretization.

5 The discretized hybrid method

5.1 Structure of the hybrid operator

The X-subspace also has to be truncated if it is not already finite. The truncation follows the same reasoning as for the truncation of the Y-space. Remaining is an N-dimensional grid with M_X points in the $X-$space and M_Y points in the Y-space. Let $\mathcal{M}_X$ be a linear mapping from $\mathbb{Z}^N$ to $\mathbb{Z}$ defined by

$$\mathcal{M}_X\mathbf{x} = \sum_{j=1}^{N} x_j \cdot \prod_{0<k}^{k<j} M_k,$$

where $\mathbf{x} \in \mathbb{Z}^N$ and M_k are constants. The mapping can be used to assign an index $i = 0, \ldots, M_X - 1$ to the points in the X-subspace if M_k is the number of states in each dimension.

The index is used to order the states and a corresponding mapping $\mathcal{M}_Y$ is used for the Y subspace and the states are ordered into

$$\mathbf{p} = \begin{bmatrix} p_{0,0} \\ \vdots \\ p_{0,M_Y-1} \\ p_{1,0} \\ \vdots \\ p_{M_X-1,M_Y-1} \end{bmatrix}, \quad W_r = \begin{bmatrix} w_{r,0,0} & & & & & \\ & \ddots & & & & \\ & & w_{r,0,M_Y-1} & & & \\ & & & w_{r,1,0} & & \\ & & & & \ddots & \\ & & & & & w_{r,M_X-1,M_Y-1} \end{bmatrix},$$

where $p_{i,j}(t) = p(\mathcal{M}_X^{-1}i, \mathcal{M}_Y^{-1}j, t)$ and $w_{r,i,j} = w(\mathcal{M}_X^{-1}i, \mathcal{M}_Y^{-1}j)$. The discretized FPE-operator for a constant state $\mathbf{x}$ in the X-space is denoted by $A_{r,\mathcal{M}_X^{-1}\mathbf{x}}$. Define the block-diagonal matrix

$$A_r = \begin{bmatrix} A_{r,0} & & & \\ & A_{r,1} & & \\ & & \ddots & \\ & & & A_{r,M_X-1} \end{bmatrix}.$$

The mapping $\mathcal{M}_X$ is also used to find the corresponding state in $\mathbf{p}$ of a state $\mathbf{x}$ for a certain reaction Π_r. The position of the initial state relative to the final state is $\mathcal{M}_X(\mathbf{x} + \mu_r) - \mathcal{M}_X\mathbf{x} = \mathcal{M}_X\mu_r$, if both states are in the truncated domain. The hybrid equation can now be written

$$\frac{d\mathbf{p}}{dt} = \sum_{r=1}^{R} \{(\delta(M_Y \cdot \mathcal{M}_X\mu_r) - I)W_r + \delta(M_Y \cdot \mathcal{M}_X\mu_r)A_r\}\,\mathbf{p}, \tag{9}$$

where I is an identity matrix and $\delta(d)$ is a shifted identity matrix of dimension $M_X M_Y$ with the elements

$$\delta_{i,j}(d) = \begin{cases} 1 \text{ if } i + d = j, \\ 0 \text{ otherwise.} \end{cases}$$

The hybrid discretization is conservative if $A_{r,i}$ is conservative for all i and the reflecting boundary conditions

$$w_r(\mathcal{M}_X^{-1}k, y) = 0 \text{ if } \quad \mathcal{M}_X^{-1}k - \mu_r \notin \mathbb{X} = \{\mathbf{x}|x_i = 1, \ldots, M_i\},$$

are applied.

5.2 Time discretization

The space-discretized FPE-CME-hybrid on a constant grid is a system of time dependent equations

$$\frac{d\mathbf{p}}{dt} = A\mathbf{p},$$

where A is a constant matrix and $\mathbf{p}$ is the vector defined previously. The equation is discretized in time with the implicit second order *backward differentiation formula* (BDF-2) [6] with adaptive time stepping [12]. The adaptive time step was limited by a local error tolerance of one percent measured in the ℓ_1-norm. The implicit BDF-2 is well suited for the stiffness of molecular biological systems. BiCGSTAB [15] is used to solve the system of equations in each time step using incomplete LU-factorization with zero fill-in as preconditioner [5]. As the time step changes, the efficiency of the preconditioner varies and needs to recomputed occasionally, but not too often. The time step will also affect the convergence rate of the iterative equation solver. There is a trade-off between taking longer time steps which makes the number of systems of equations to solve lower and shorter time steps which makes the solution of each system of equations converge faster. In order to avoid time steps that would cause too high iteration counts, the time step is halved every time 90 percent or more of the maximum iteration count is used. The maximum iteration count was here set to 40.

6 Monte Carlo solution of the master equation

The solution of the master equation is approximated by the *stochastic simulation algorithm* (SSA) [3]. The SSA simulates the chemical events of the system. The simulation is a realisation of the stochastic process that is described by the master equation and is a sample of the probability distribution of this process. By generating many such sample trajectories through the state space, the probability distribution can be estimated and the estimate converges to the solution of the master equation.

The computational cells previously defined is used for defining a spatial resolution of SSA solutions. SSA is used to compute the probability for the state of the system to be in each computational cell. The probability for a state on the boundary between computational cells is equally distributed between the bordering cells. L trajectories are computed and partitioned into L_0 groups of $\Delta\lambda$ trajectories. The number of trajectories of group l that is in a state in computational cell ν at the time t is denoted by $\Delta\lambda_\nu^l$. Define

$$P_\nu^l(t) = \Delta\lambda_\nu^l/\Delta\lambda,$$

which is a sequence of independent stochastic variables for $l = 1, 2, \ldots$. According to the central limit theorem [9], the average

$$\bar{P}_\nu^{\Delta\lambda} = \frac{1}{L_0} \sum_{l=1}^{L_0} P_\nu^l \tag{10}$$

is normally distributed $\mathcal{N}(\mu_\nu, \sigma_\nu^2/L_0)$, where σ_ν^2 is the variance in P_ν^l. The expected value μ_ν is the average of the solution of the master equation in computational cell ν, and the estimate (10) of the mean value is the computational solution. The error ϵ in the numerical solution using the ℓ_1-norm is

$$\epsilon = \sum_\nu \omega |\bar{P}_\nu^{\Delta\lambda} - \mu_\nu|.$$

The ℓ_1-norm is used since it preserves the probability mass. Using that $\bar{P}_\nu^{\Delta\lambda} - \mu_\nu$ is normally distributed $\mathcal{N}(0, \sigma_\nu^2/L_0)$, the expected value of this sum is

$$\epsilon \approx \sum_\nu \omega \sqrt{\frac{2}{\pi}} \int_0^\infty x e^{-\frac{x^2}{2\sigma_\nu^2/L_0}} dx = \sqrt{\frac{2}{\pi L_0}} \|\sigma\|_1, \tag{11}$$

where σ is a vector and the mapping $\mathcal{M}$ is used to order the elements in the same way as for p previously. The variance is estimated using the sample variance s_ν^2. With $l_0 \leq L_0$ samples, the sample variance is

$$s_\nu^2 = \frac{1}{l_0 - 1} \sum_{l=1}^{l_0} (P_\nu^l - \bar{P}_\nu^{\Delta\lambda})^2 \approx \sigma_\nu^2.$$

7 Experiments

A model presented in [4] is used to test the numerical methods. It is a model of a gene control circuit where the gene expression of the gene G is controlled by the gene product M through its dimer D. The dimer binds cooperatively to two regulatory sites S_1 and S_2. If no dimer is bound to the regulatory region, the gene is inactivated; if D is bound to S_1 but not to S_2 the gene is active and if D is bound to both regulatory sites, the gene is once again inactivated. The model comprises six molecular species: the gene product monomer (M), the gene product dimer (D), the gene when S_1 and S_2 are unoccupied (DNA), the gene when S_1 is occupied by a D-molecule (DNA_M), the gene when both S_1 and S_2 are occupied by D-molecules (DNA_{2M}) and mRNA (RNA). The number of molecules is denoted by the token for the corresponding species in lower case characters, i.e. m, d, dna, dna_M, dna_{2M} and rna. A dimer bound to S_1 activates transcription of mRNA while a metabolite bound to S_2 shuts down mRNA production. The binding to S_2 is cooperative with binding to S_1 to the degree that no D molecule will bind to S_2 unless S_1 is occupied and a D molecule

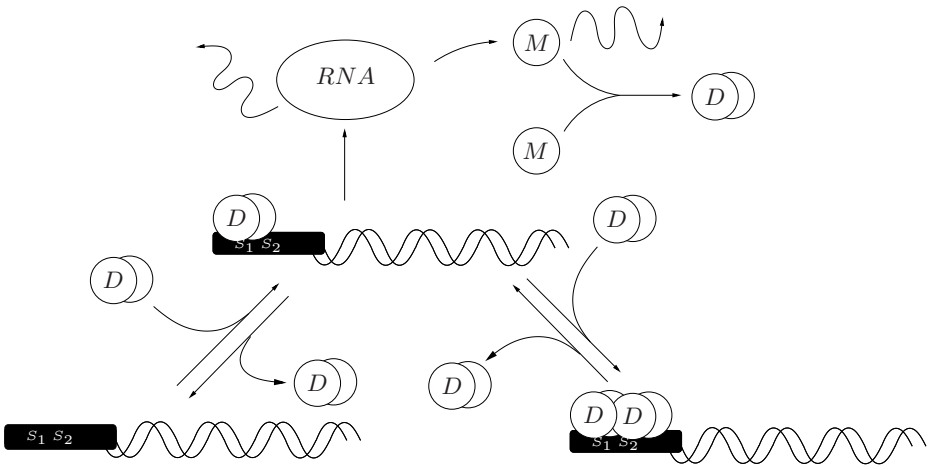

Fig. 1. The reactions of the gene control circuit.

bound to S_1 will not unbind while S_2 is occupied. The metabolite and mRNA are actively degraded.

Exactly two gene copies are assumed in this model. There is no modelling of the replication process and the cell volume is assumed constant. The constant number of gene copies implies a dimension reduction, since the probability for $dna + dna_M + dna_{2M} = 2$ is 1. The state space is really a five-dimensional surface in a six-dimensional state space. The state description is reduced to $\tilde{\mathbf{x}} = (dna, dna_M, rna, m, d)^T$ and dna_{2M} is substituted by $2 - dna - dna_M$ in the propensity functions. Table 1 lists the reactions of the reduced model, that we will call *model A*. Model A is approximated by a FPE in the *rna*-, *m*- and *d*-dimensions. The symbol $\emptyset$ is used when no molecules included in the model are produced by a reaction. A reduced model in two dimensions is used as a test problem. This model will be called *model B* and is defined in Table 2. The FPE of this model problem is

$$\begin{aligned}\frac{\partial p}{\partial t} &= \partial_m\, p + \frac{1}{2}\partial_m^2\, p - \frac{1}{40}\partial_m(mp) + \frac{1}{80}\partial_m^2(mp) + 2\partial_m(m(m-1)p) \\ &- m(m-1)\partial_d\, p + 2\partial_m^2(m(m-1)p) + \frac{1}{2}m(m-1)\partial_d^2\, p - 2\partial_m\partial_d(m(m-1)p) \\ &- \frac{156}{25}d\partial_m\, p + \frac{78}{25}\partial_d(dp) + \frac{156}{25}d\,\partial_m^2\, p + \frac{39}{25}\partial_d^2(dp) - \frac{156}{25}\partial_m\partial_d(dp) \quad (12)\end{aligned}$$

The full model as defined in Table 1 was simulated for $600s$ (model time) using the hybrid method and SSA for $rna = [0, 30]$, $m = [0, 150]$ and $d = [0, 600]$. The initial probability distribution $p(\mathbf{x}, 0)$ is constructed such that $p(\mathbf{x}, 0) =$

Reaction	$\mathbf{n}_r$	Propensity
$RNA \xrightarrow{w_1} RNA + M$	$(0,0,0,-1,0)^T$	$w_1 = 0.05 \cdot rna$
$M \xrightarrow{w_2} \emptyset$	$(0,0,0,1,0)^T$	$w_2 = 0.001 \cdot m$
$DNA_M \xrightarrow{w_3} RNA + DNA_M$	$(0,0,-1,0,0)^T$	$w_3 = 0.1 \cdot dna_M$
$RNA \xrightarrow{w_4} \emptyset$	$(0,0,1,0,0)^T$	$w_4 = 0.005 \cdot rna$
$DNA + M \xrightarrow{w_5} DNA_M$	$(1,-1,0,1,0)^T$	$w_5 = 0.02 \cdot dna \cdot m$
$DNA_M \xrightarrow{w_6} DNA + M$	$(-1,1,0,-1,0)^T$	$w_6 = 0.5 \cdot dna$
$DNA_M + M \xrightarrow{w_7} DNA_{2M}$	$(0,1,0,1,0)^T$	$w_7 = 2 \cdot 10^{-4} \cdot dna_M \cdot m$
$DNA_{2M} \xrightarrow{w_8} DNA_M + M$	$(0,-1,0,-1,0)^T$	$w_8 = 1 \cdot 10^{-11} \cdot (2 - dna - dna_M)$
$M + M \xrightarrow{w_9} D$	$(0,0,0,2,-1)^T$	$w_9 = \frac{5 \cdot 10^7}{AV} \cdot m \cdot (m-1)$
$D \xrightarrow{w_{10}} M + M$	$(0,0,0,-2,1)^T$	$w_{10} = 0.5 \cdot d$

Table 1. The reactions of model A. The model comes from [4]. V is the volume of the biological cell and A is the Avogadro constant.

Reaction	$\mathbf{n}_r$	Propensity
$M \xrightarrow{w_1} \emptyset$	$(1,0)^T$	$w_1 = 1.00 \cdot m$
$\emptyset \xrightarrow{w_2} M$	$(-1,0)^T$	$w_2 = 0.0250$
$M + M \xrightarrow{w_3} D$	$(2,-1)^T$	$w_3 = 1.00 \cdot m \cdot (m-1)$
$D \xrightarrow{w_4} M + M$	$(-2,1)^T$	$w_4 = 3.12 \cdot d$

Table 2. The reactions of model B.

$p(dna,0)p(dna_M,0)p(rna,0)p(m,0)p(d,0)$, where

$$
\begin{aligned}
p(dna,0) &= \begin{cases} 1 \text{ if } dna = 2 \\ 0 \text{ otherwise,} \end{cases} \\
p(dna_M,0) &= \begin{cases} 1 \text{ if } dna = 0 \\ 0 \text{ otherwise,} \end{cases} \\
p(rna,0) &= \mathcal{N}_{h,3}(10,10), \\
p(m,0) &= \mathcal{N}_{h,4}(40,40), \\
p(d,0) &= \mathcal{N}_{h,5}(300,300).
\end{aligned}
$$

$\mathcal{N}_{h,i}(m,\sigma)$ is the probability density function of a normal distribution with mean value m and standard deviation σ in the discretization points of Ω_h in the i:th dimension. $p(\mathbf{x},0)$ is scaled so that the discrete total probability is equal to one. Figure 2 shows different projections of the probability density function of the FPE part of the solution for model A. Figure 3 displays the projection of the numerical solutions on the $rna \times m \times d$-subspace.

7.1 Error estimates

A reference solution $\tilde{p}$ was computed by SSA at time $t = 200s$ using 10^6 trajectories to sample the state space. The states at $t = 200s$ were stored and used to produce probability distributions for the state space and some of its subspaces

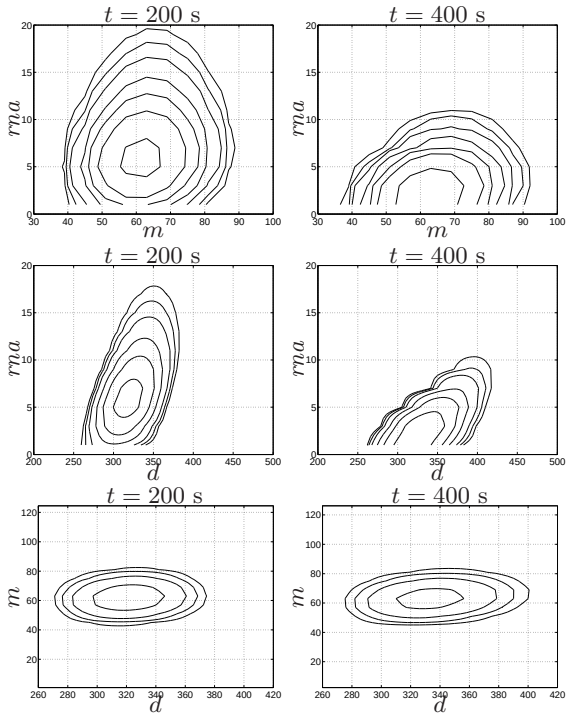

Fig. 2. Snapshots of the numerical solution projected on two dimensional subspaces at $time = 200s$, using isolines $p = \{5.0 \cdot 10^{-5}, 1.0 \cdot 10^{-4}, 2.0 \cdot 10^{-4}, 5.0 \cdot 10^{-4}, 1.0 \cdot 10^{-3}, 2.0 \cdot 10^{-3}, 5.0 \cdot 10^{-3}\}$

in one, two and three dimensions. The corresponding marginal probability functions were computed from the FPE solution at the $t = 200s$.

Figures 4–5 show how the error in an SSA solution decreases as a function of the number of sample trajectories as well as the predicted error in the solution according to (11). The error in the FPE approximation with respect to the reference solution is indicated in the plots by a dashed line. The error in $\tilde{p}$, using (11) is less than 1% in one and two dimensions and about 2.3% in three dimensions. The immediate observation from the convergence studies is that the number of trajectories necessary for reaching the same error as the FPE approximation shows a moderate growth for increasing number of dimensions. This indicates that there is a dimension dependence also for the SSA, which must be attributed to σ_ν. The spatial resolution is also of importance for σ_ν and affects the convergence of SSA. Figure 6 shows how the convergence changes with a refinement in the cell structure. For SSA, a decrease in computational cell size is coupled to an increase in the error.

Another observation is that, for this example, it would be necessary to simulate approximately $5 \cdot 10^4$ trajectories to match the accuracy of the FPE approx-

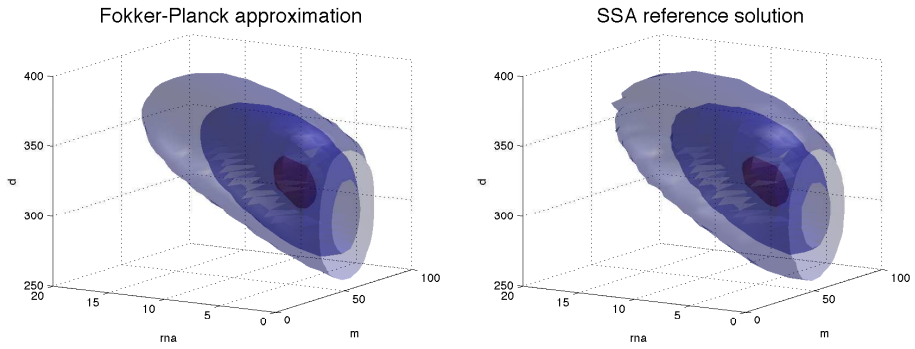

Fig. 3. Snapshot of isosurfaces (10^{-6}, 10^{-5} and 10^{-4}) of the numerical reference solution at time $t = 200s$ projected on the $rna \times m \times d$-subspace.

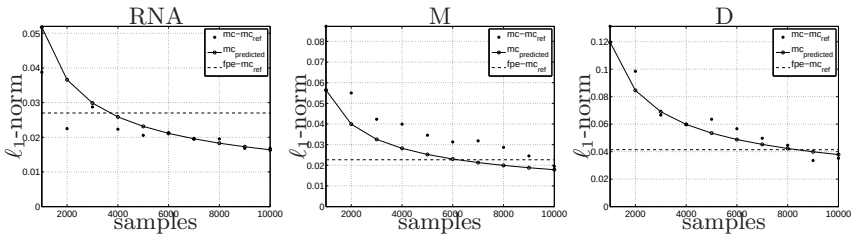

Fig. 4. Convergence of SSA for one-dimensional problems. The errors in the solutions obtained with FPE and SSA simulation are plotted.

imation. The computation of the FPE solution takes about four times longer to compute (not including the generation of the space discretization) and the reference solution $\tilde{p}$ using SSA about twenty times longer. The generation of the matrix in the FPE space discretization is performed by a flexible code but it is not optimized for efficiency. It should in principle be fast compared to the time integration.

The accuracy seems rather poor in the picture painted by the convergence plots, considering that the ℓ_1-norm of the entire solution equals one. Just looking at the solutions, they are quite similar, see Figures 7–8.

The estimates of the truncation error (6) and (8) require some derivatives of the solution which are generally not available. In order to examine the error for model B, the steady state solution was approximated with a normal distribution with the same mean and variance and a correlation coefficient $\rho = 0$. Figure 9 (left) shows that the match is fairly good. The derivatives for the normal distribution are easily calculated analytically and the max-norm in the error estimates can be computed. Figure 9 (right) shows how the max-norm of the analytical derivatives decrease the higher they are. The logarithm of the analytical

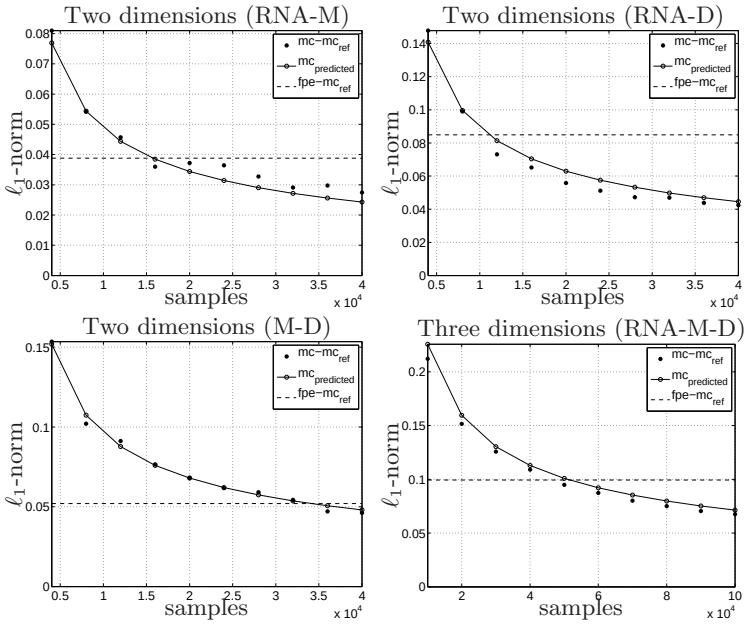

Fig. 5. Convergence of SSA for two- and three-dimensional problems.

derivatives evaluated on an $\mathbb{N}^2$-grid with $h_1 = h_2 = 1$ is plotted as a function of numbers of the applications of ∂_{x_i} and ∂_{x_j} on the normal distribution. For model B, the truncation error for the whole FPE operator using the second order accurate discretization on a $\mathbb{N}^2$-grid is $\|\tau_{2o}\|_\infty = 0.02$. For the fourth order discretization the truncation error is $\|\tau_{4o}\|_\infty = 10^{-4}$ with $h = 2$ we would have $\|\tau_{2o}\|_\infty \approx 0.08$ and $\|\tau_{4,0}\| \approx 0.0016$. For smooth problems as this one, a fourth order method is superior. The $\mathbb{N}^2$-grid is a critical limit for FPE approximation since it is the grid that is used for an unapproximated CME.

For this problem it is necessary to use the fourth order scheme since the error is far too large for the second order scheme to be of practical use, which is illustrated in Figures 10 and 11. They show the mean value and standard deviation for solutions using second order or fourth order discretizations, as a function of time. Figure 10 displays the means and standard deviations for the second order accurate hybrid method on a grid with $3\times3\times15\times20\times150 = 405000$ degrees of freedom, compared to what is obtained by SSA using 10^4 trajectories. Figure 11 shows the same comparison between the fourth order scheme and the same grid, and SSA using 10^4 trajectories. Due to numerical approximation errors the solution is slightly negative at some parts of the state space. In order to calculate the mean, standard deviation and covariances those negative values are set to zero.

The mean values of the *dna* and dna_M decrease to zero and the standard deviation follows. That means that the probability for both gene copies to be in

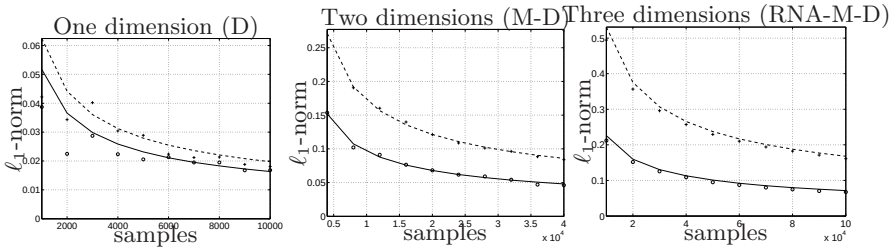

Fig. 6. The theoretical convergence of SSA using the FPE resolution (solid lines) is compared to the theoretical convergence using half the step size in the subspace Y (dashed lines). The measured error in SSA compared to the reference solution for both cases are also plotted ($\circ$ for the coarse and $+$ for the fine resolution).

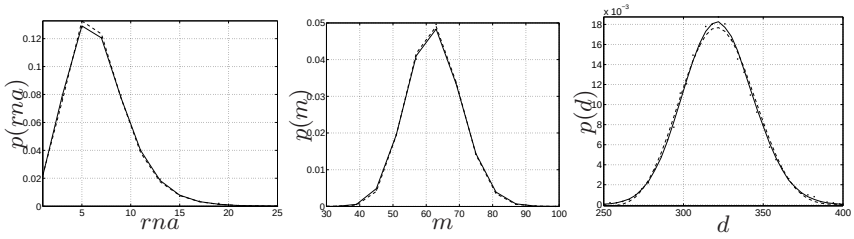

Fig. 7. Comparison of the reference solution (solid lines), the FPE solution (dashed lines) and a SSA solution using 10^4 trajectories ($\bullet$), in one dimension at $t = 200s$.

the binding state with both S_1 and S_2 occupied goes towards one. Hence, the solution shrinks to an almost three-dimensional distribution in the state space for increasing t.

The covariances of the FPE solution is not quite as good a match as the mean values and standard deviations. The results in the left column indicate that the resolution in the *rna*-dimension is too low. Contrary to the m- and d-dimensions, *rna* copy number is low and is distributed close to the boundary where the numerical accuracy is lower than in the interior of the computational domain due to the extrapolation to ghost points near the boundary. The fourth order discretization is derived for an equidistant grid so in order to increase the resolution it necessary to add a substantial number of grid points. As an alternative the reactions where *rna*-concentrations are altered were discretized using the second order discretization, allowing a stretched grid. In Figure 12 the *rna*-dimension is stretched so that the i:th in the *rna*-dimension $h_{rna,i} = (1+\theta)h_{rna,i-1}$, where $h_{rna,0}$ was set to 0.5 and θ is determined so that the axis spans the computational domain in the *rna*-dimension. The accuracy is then much better.

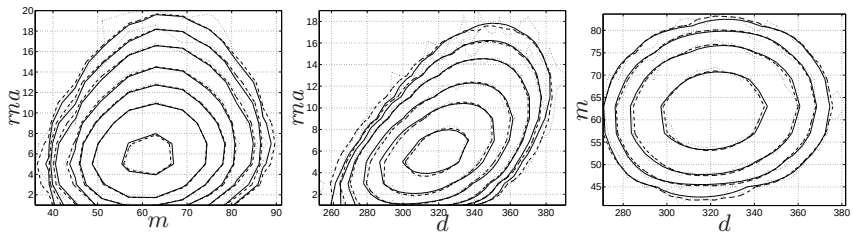

Fig. 8. Comparison of the reference solution (solid lines), the FPE solution (dashed lines) and a SSA solution using 10^4 trajectories (dotted lines), in two dimensions at $t = 200s$ using the isolines $p = \{5.0 \cdot 10^{-5}, 1.0 \cdot 10^{-4}, 2.0 \cdot 10^{-4}, 5.0 \cdot 10^{-4}, 1.0 \cdot 10^{-3}, 2.0 \cdot 10^{-3}, 5.0 \cdot 10^{-3}\}$.

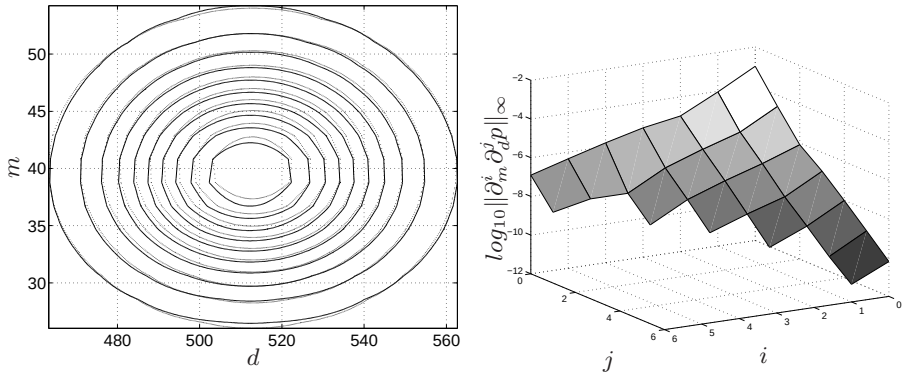

Fig. 9. Left: comparison between function (solid lines) and approximation (dotted lines) and right: supremum norm of the derivatives of the normal distribution approximation.

8 Conclusion

The master equation is partially approximated by the FPE in selected dimensions. The performance of FPE approximation contra SSA is very dependent on the problem [13]. SSA is sensitive to long time correlations in trajectories and the size of the state space, which both affect the efficiency of the sampling. The FPE approach is insensitive to these properties, but is a bit more expensive to set up and also is more sensitive to low resolution of the state space.

This problem is actually rather well suited for SSA. The solution is monomodal and rather compact in the state space. Accordingly the SSA is faster, but not so much faster that FPE can not compete, at least if higher accuracies are required. The foremost drawback of FPE approximations is the higher memory demands. The second order discretization is not sufficient for approximating the gene circuit model without excessive resolution, but the higher order discretiza-

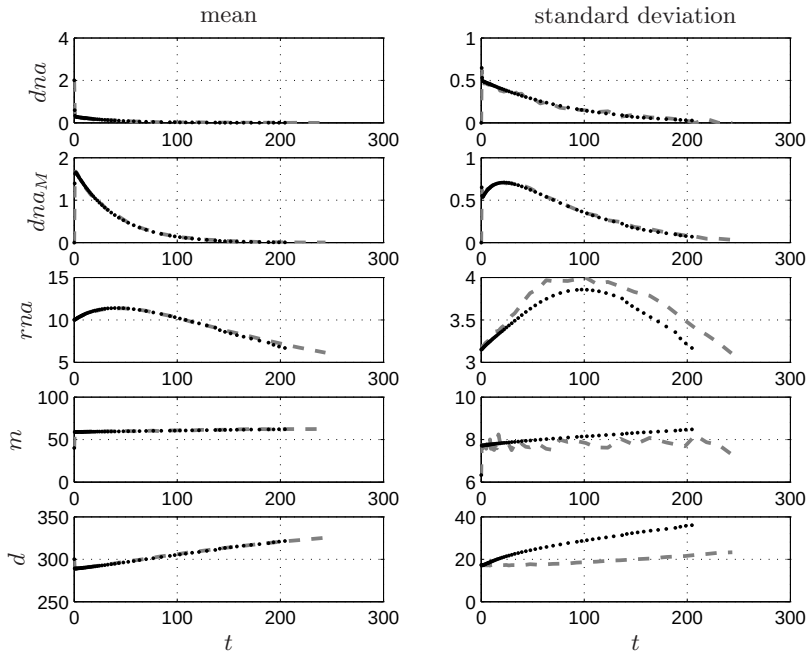

Fig. 10. Central moments for the second order accurate discretization. Mean values (*left*) and standard deviations (*right*) for, from top to bottom, dna, dna_M, rna and m. The hybrid solution (•) and the SSA solution (*dashed gray line*) are plotted for each case.

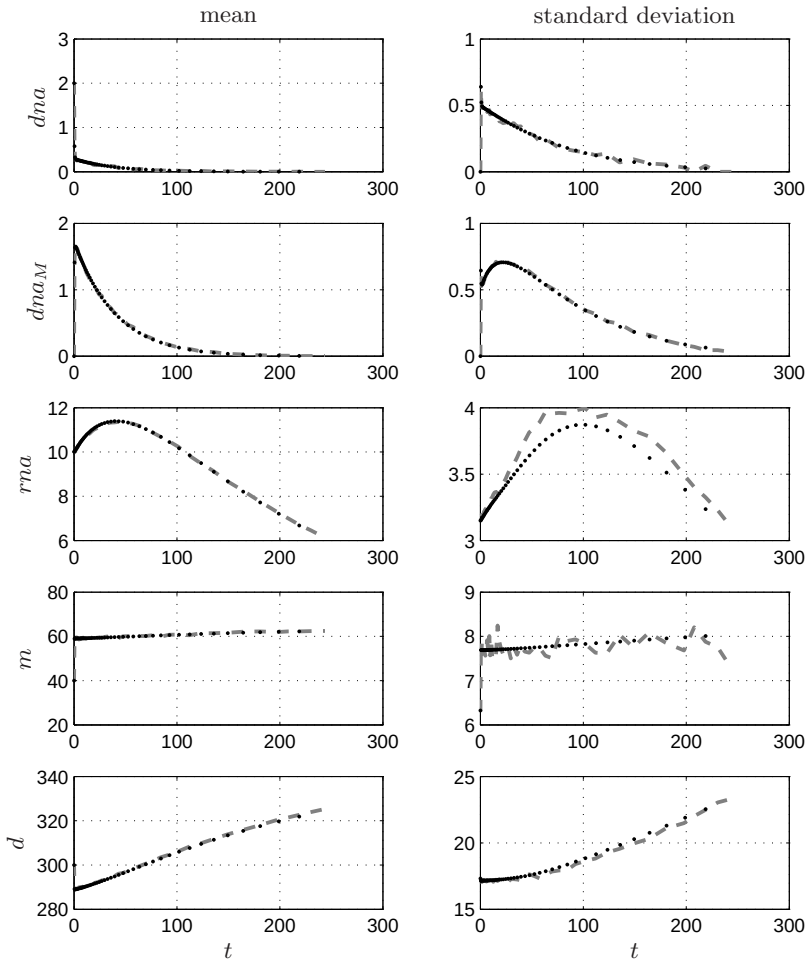

Fig. 11. Central moments for the fourth order accurate discretization. Mean values (*left*) and standard deviations (*right*) for, from top to bottom, *dna*, dna_M, *rna* and m. The hybrid solution (•) and the SSA solution (*dashed gray line*) are plotted for each case.

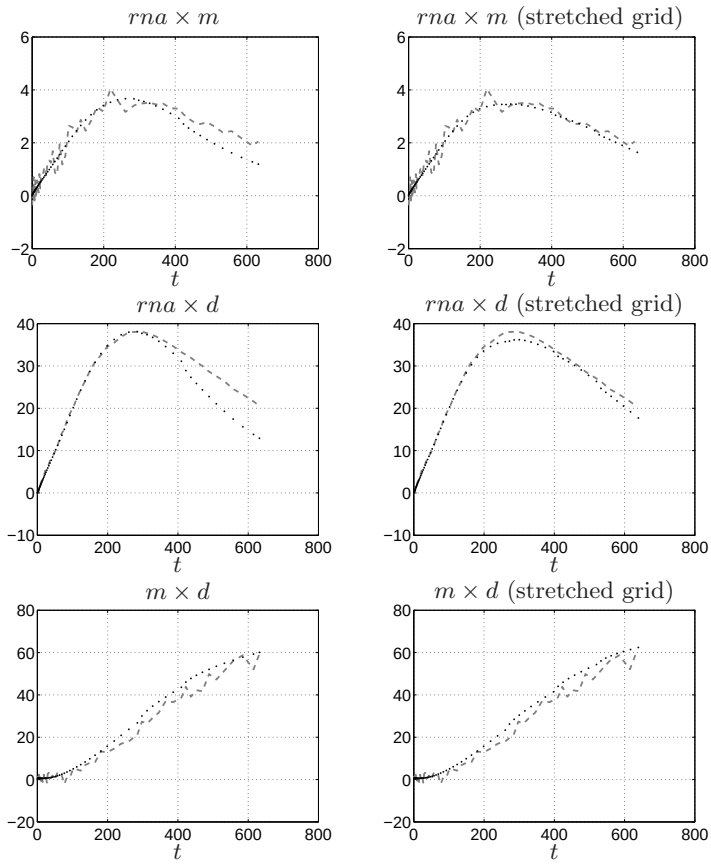

Fig. 12. The covariance for FPE (•) and SSA (*dashed gray line*) using 10^4 trajectories for equidistant grid (left) and a grid stretched in the *rna*-direction (right).

tion makes the FPE approximation a possible solution strategy. An interesting point, though is that the SSA seems to suffer from the increase in dimensionality in a way that is seldom discussed, although it is much weaker than for the FPE.

The example of how stretching the grid compensates for low accuracy is also an example of the flexibility of PDE approximations of the CME. The stretching of the grid was made in an *ad hoc* manner, but yielded an instant and significant improvement. Adaptive grid methods automatically exploit this flexibility. For an example of a related problem in financial mathematics demonstrating grid adaptivity for high-dimensional problems, see [11]. To fully exploit this flexibility adaptive grids may contribute some rigor to the solution process.

Acknowledgements

I want to thank Per Lötstedt for his invaluable advice. This work was funded by the Swedish Research Council, the Swedish National Graduate School in Scientific Computing and the Swedish Foundation for Strategic Research.

References

1. Y. Cao, D. T. Gillespie, and L. R. Petzold. The slow-scale stochastic simulation algorithm. *J. Chem. Phys.*, 122:014116, 2005.
2. L. Ferm and P. Lötstedt. Numerical method for coupling the macro and meso scales in stochastic chemical kinetics. *BIT*, 2007. to appear.
3. D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.
4. J. Goutsias. Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems. *J. Chem. Phys.*, 122:184102, 2005.
5. A. Greenbaum. *Iterative Methods for Solving Linear Systems*. SIAM, Philadelphia, 1997.
6. E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations, Nonstiff Problems*. Springer-Verlag, Berlin, second edition, 1993.
7. E. Haseltine and J. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.*, 117:6959–6969, 2002.
8. A. Hellander and P. Lötstedt. Hybrid method for the chemical master equation. *J. Comput. Phys.* to appear.
9. R. J. Larsen and M. L. Marx. *An Introduction to Mathematical Statistics and Its Applications*. Prentice-Hall, Englewood Cliffs, NJ, second edition, 1986.
10. P. Lötstedt and L. Ferm. Dimensional reduction of the Fokker-Planck equation for stochastic chemical equations. *Multiscale Simul. Model.*, 5:593–614, 2006.
11. P. Lötstedt, J. Persson, L. von Sydow, and J. Tysk. Space–time adaptive finite difference method for European multi-asset options. *Comput. Math. Appl.*, 53:1159–1180, 2007.
12. P. Lötstedt, S. Söderberg, A. Ramage, and L. Hemmingsson-Frändén. Implicit solution of hyperbolic equations with space-time adaptivity. *BIT*, 42:134–158, 2002.
13. P. Sjöberg, P. Lötstedt, and J. Elf. Fokker-Planck approximation of the master equation in molecular biology. *Comput. Visual. Sci.*, 11, 2008. published electronically in February 2007.
14. Z. Szallasi, J. Stelling, and V. Periwal, editors. *Systems Modeling in Cell Biology : From Concepts to Nuts and Bolts*. MIT Press, Cambridge, Mass, 2006.
15. H.A. van der Vorst. BiCGSTAB: A fast and smoothly converging variant of the Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci and Stat. Comp*, 13:631–644, 1992.
16. N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam, second edition, 1992.