

The Linear Noise Approximation for Spatially Dependent Biochemical Networks

Per Lötstedt¹ 

Received: 9 January 2018 / Accepted: 29 March 2018 / Published online: 11 April 2018
© The Author(s) 2018

Abstract An algorithm for computing the linear noise approximation (LNA) of the reaction–diffusion master equation (RDME) is developed and tested. The RDME is often used as a model for biochemical reaction networks. The LNA is derived for a general discretization of the spatial domain of the problem. If M is the number of chemical species in the network and N is the number of nodes in the discretization in space, then the computational work to determine approximations of the mean and the covariances of the probability distributions is proportional to M^2N^2 in a straightforward implementation. In our LNA algorithm, the work is proportional to M^2N . Since N usually is larger than M , this is a significant reduction. The accuracy of the approximation in the algorithm is estimated analytically and evaluated in numerical experiments.

Keywords Linear noise approximation · Spatially dependent · Fast algorithm

Mathematics Subject Classification 60J60 · 65C40 · 92C45

1 Introduction

Many biochemical networks are modeled by ordinary or partial differential equations at a macroscopic level of fidelity. Such continuous models may not be sufficiently accurate when the number of molecules involved in the chemical reactions is small. This is often the case in molecular cell biology (Elowitz et al. 2002; McAdams and

✉ Per Lötstedt
perl@it.uu.se

¹ Division of Scientific Computing, Department of Information Technology, Uppsala University, SE-75105 Uppsala, Sweden

Arkin 1997; Raj and van Oudenaarden 2008; Tsimring 2014). Chemical reactions are then best described as random events, and the discrete number of molecules is important when the copy numbers are low at a mesoscopic level of modeling. The macroscopic equation for the mean values is often satisfactory when the number of molecules is large. Analytical solutions to the governing macroscopic or mesoscopic equations can be obtained only for special systems. Computational methods are needed for quantitative information about the behavior of the systems.

The master equation (ME) or Kolmogorov forward equation is an equation for the time evolution of the probability density function (PDF) for the copy numbers of different species in systems with an intrinsic noise (Gardiner 2004; Gillespie 1992; van Kampen 2004). The systems are modeled as Markov processes with discrete states defined by the copy numbers of the chemical species in continuous time. The particular ME for spatially homogeneous, well-stirred problems in chemistry is the chemical master equation (CME) where reactions between two molecules occur with a propensity that depends on the copy numbers of the species. The ME is generalized in the reaction–diffusion master equation (RDME) to spatially heterogeneous chemical systems by introducing a discretization of the reaction volume into compartments or voxels (Gillespie et al. 2014; Gillespie and Seitaridou 2013). The state is then given by the copy numbers in each one of the voxels.

The computational work and the storage requirements to solve the RDME grows exponentially in the number of species and the number of voxels making the simulation of biochemical systems with the ME prohibitive except for very small systems. Analytical solutions are known only for limited classes of problems such as those with linear propensities. Instead, sample trajectories of well-stirred systems are generated by Gillespie's stochastic simulation algorithm (SSA) (Gillespie 1976, 1977). The original algorithm has been improved in many ways, e.g., for efficiency (Cao et al. 2006; Gibson and Bruck 2000; Gillespie 2001) and for systems with slow and fast reactions (Cao et al. 2005; E et al. 2007). The Gillespie algorithm is generalized to problems with spatial variation due to diffusion in Elf and Ehrenberg (2004), Engblom et al. (2009), Isaacson and Peskin (2006), Lampoudi et al. (2009) and implemented in software (Drawert et al. 2012, 2016; Hattne et al. 2005). The computational effort may be quite large to simulate a system with many chemical species, many molecules, and many voxels, since many realizations of the process are required in a Monte Carlo method like SSA due to slow convergence to the mean and other moments of the distribution. An introduction and an overview of Markov models of chemical reactions are found in Goutsias and Jenkinson (2013). Recent reviews of computational methods at different levels of modeling are Engblom et al. (2017), Gillespie et al. (2013), Mahmutovic et al. (2012) and Sokolowski et al. (2017).

There are ways to approximate the solutions to the CME with deterministic equations. The *linear noise approximation* (LNA) is obtained from the CME by deriving the equations for the moments and then expanding the solution in a large parameter Ω , representing the size of the chemical system (van Kampen 1976, 2004). The means and covariances in LNA are exact for chemical systems with at most first-order reactions where the propensities are constants or linear in the copy numbers. The first and second moments are exact also for other systems with a special structure (Grima 2015). Different modifications have been proposed to improve the accuracy of LNA,

see, e.g., Ferm et al. (2008), Grima (2012). Some of the improvements are compared experimentally in examples in Schnoerr et al. (2015). The LNA and similar approximations are used to quickly study biochemical networks in, e.g., Elf and Ehrenberg (2003), Thomas et al. (2013), Ullah and Wolkenhauer (2009) and more recently as a surrogate model to infer parameters in biochemical models from data in Fearnhead et al. (2014), Fröhlich et al. (2016), Rutter and Opper (2009), Stathopoulos and Girolami (2013). A review of LNA and related methods and their use for inference are found in Schnoerr et al. (2017).

An alternative to the LNA is the EMRE approximation in Grima (2010) extended to spatial problems in Smith et al. (2017). The covariances satisfy the same Lyapunov equation as we derive here. The spatial EMRE algorithm is applied to gene regulation in a cell and to reactions in an aggregation of cells in two space dimensions in Smith et al. (2017). More compartments than one are also found in Challenger et al. (2012). The equations of LNA with spatial variation are derived in Scott et al. (2010) and applied to the modeling of spatial patterns. The equation for the covariances is replaced by an equation for the factorial cumulant. Equations similar to the LNA for spatial problems are used in Butler and Goldenfeld (2009), Anna et al. (2010) to investigate oscillatory systems. Turing patterns are studied in Asllani et al. (2013), Biancalani et al. (2010), McKane et al. (2014), Woolley et al. (2011) with a spatially extended LNA.

Diffusive effects are important for the fidelity of models when the chemical reactions are localized in space in a cell and when the molecular transport is slow compared to the reactions. Some examples where the spatial effects are crucial are found in Fange and Elf (2006), Sturrock et al. (2013), Takahashi et al. (2010). The LNA is a level of modeling suitable for such systems, e.g., to infer parameters for the diffusion and the reactions from measurements, at least in the beginning of the iterative search process for the parameters, thanks to the relative simplicity of LNA.

In this paper, we develop a fast algorithm for computing approximations of the mean and the covariance of the PDF solving the RDME based on the LNA for spatial problems with reactions and diffusion. The equation for the expected values is a system of reaction–diffusion equations, and the equation for the covariances is a time-dependent Lyapunov equation with a source term localized in space. Let M be the number of chemical species and N the number of voxels. The structure of the covariance equations is utilized to compute an approximation of the covariance and to reduce the computational work and the memory requirements from being proportional to M^2N^2 in a straightforward implementation to M^2N in our algorithm. Since N usually is larger than M , this is a substantial reduction. A bound on the deviation of the true covariance from our approximation is proved in a theorem. The accuracy of the covariance approximation is demonstrated in numerical examples in one, two, and three dimensions (1D, 2D, and 3D).

In the next section, the RDME is given and a splitting of the operator is introduced. The equations of the LNA for spatially heterogeneous chemical systems are derived for general shapes of the voxels in Sect. 3. A continuous approximation of the equation for the covariances is analyzed in Sect. 4. The algorithm is presented in Sect. 5 for computation of the mean and the covariance. Numerical results are found in Sect. 6. Finally, some conclusions are drawn.

The notation in the paper is as follows. The i th element of a vector \mathbf{v} is denoted by v_i . The j th column of an array \mathbf{x} with elements x_{ij} is written as \mathbf{x}_j , and \mathbf{x}_i is the i th row. The derivative of $v_i(\mathbf{x})$ with respect to x_j is denoted by $v_{i,j}$. The time derivative $\partial p/\partial t$ of $p(\mathbf{x}, t)$ is written as $\partial_t p$, and \dot{q} is a shorter notation for dq/dt . The Euclidean vector norm is denoted by $\|\mathbf{v}\| = \sqrt{\sum_i v_i^2}$ and the subordinate spectral norm for a matrix \mathbf{A} is $\|\mathbf{A}\|$. The set of integer numbers is written as \mathbb{Z} , and \mathbb{Z}_+ denotes the nonnegative integer numbers. In the same manner, \mathbb{R} denotes the real numbers and \mathbb{R}_+ is the nonnegative real numbers.

2 The Master Equation

Consider a biochemical system with M chemically active species. The system evolves on a line (1D), in an area (2D), or a volume (3D) \mathcal{V} which is partitioned into N voxels (or compartments) \mathcal{V}_j such that they cover \mathcal{V} , $\mathcal{V} = \bigcup_{j=1}^N \mathcal{V}_j$, and are non-overlapping, $\mathcal{V}_j \cap \mathcal{V}_k = \emptyset$, $j \neq k$. The voxels are defined by a computational mesh constructed for numerical discretization of partial differential equations, see Fig. 1. The size of a voxel is $V_i = |\mathcal{V}_i|$, and the diagonal matrix \mathbf{V} has the elements V_i in the diagonal. Each voxel has a node in the center with coordinates $\mathbf{x} \in \mathbb{R}^d$, $d = 1, 2, 3$ and the nodes are connected by edges. The molecular copy number of species i in voxel j is a random integer Y_{ij} . The state of the system is time dependent and is given by $\mathbf{y}(t)$ which is an array of nonnegative integers, $\mathbf{y} \in \mathbb{Z}_+^{M \times N}$. The state changes randomly with reactions between the species in a voxel and with diffusive jumps of the molecules between the voxels.

The CME is a Kolmogorov forward equation for the PDF $p(\mathbf{y}, t)$ for a system to be in the state \mathbf{y} at time t (Gardiner 2004; van Kampen 2004). The state changes at discrete time points after a chemical reaction in a voxel. If $\tilde{\mathbf{y}}$ is the state before the

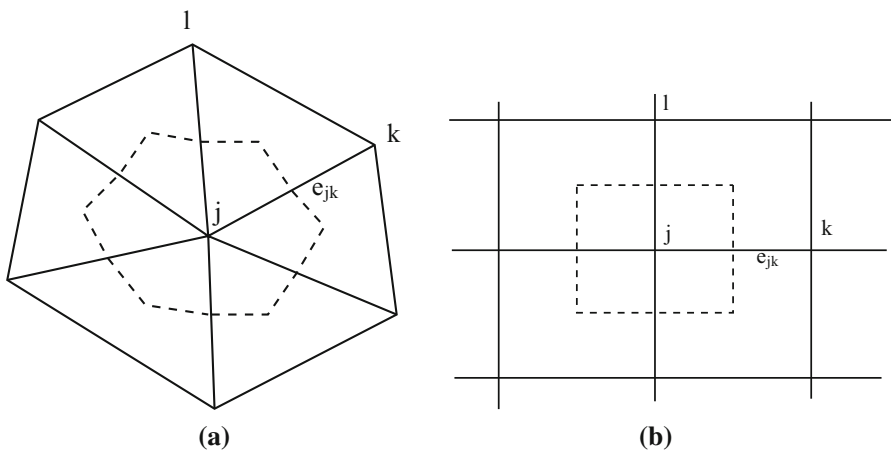


Fig. 1 Meshes with edges (solid), nodes j, k , and l , and voxel \mathcal{V}_j (dashed). The nodes j and k are connected by edge e_{jk} . **a** An unstructured mesh, **b** a structured Cartesian mesh

reaction and \mathbf{y} is the state immediately after the reaction r , then the change in state in \mathcal{V}_j can be written as

$$\tilde{\mathbf{y}}_{\cdot j} \xrightarrow{w_r(\tilde{\mathbf{y}}_{\cdot j})} \mathbf{y}_{\cdot j}, \quad \mathbf{n}_r = \mathbf{y}_{\cdot j} - \tilde{\mathbf{y}}_{\cdot j}. \tag{1}$$

The reaction occurs with the propensity w_r , i.e., with probability $w_r \Delta t$ in a short time interval Δt . The state change vector $\mathbf{n}_r \in \mathbb{Z}^M$ tells how the state is updated after a reaction. Most entries of \mathbf{n}_r are zero and $n_{ri} \neq 0$ only for those species involved in the reaction. In a system with R different reactions, the CME for $p(\mathbf{y}, t)$ is

$$\begin{aligned} \partial_t p(\mathbf{y}, t) &= \sum_{j=1}^N \sum_{r=1}^R w_r(\mathbf{y}_{\cdot j} - \mathbf{n}_r, t) p(\mathbf{y}_{\cdot 1}, \mathbf{y}_{\cdot 2}, \dots, \mathbf{y}_{\cdot j} - \mathbf{n}_r, \dots, \mathbf{y}_{\cdot N}, t) \\ &\quad - \sum_{j=1}^N \sum_{r=1}^R w_r(\mathbf{y}_{\cdot j}, t) p(\mathbf{y}, t) \equiv \mathcal{M} p(\mathbf{y}, t), \end{aligned} \tag{2}$$

defining the master operator \mathcal{M} .

Diffusion of the molecules is modeled as jumps between voxels with a common boundary. Suppose that \mathcal{V}_j and \mathcal{V}_k share a point in 1D, an edge in 2D, or a facet in 3D. Then, a molecule of species i in \mathcal{V}_j jumps to \mathcal{V}_k with propensity $q_{jk} \tilde{y}_{ij}$

$$\tilde{y}_{ij} \xrightarrow{q_{jk} \tilde{y}_{ij}} x_{ik}, \quad \mathbf{n}_{jk} = \mathbf{y}_i - \tilde{\mathbf{y}}_i. \tag{3}$$

The probability for a molecule to jump is given by the jump coefficient q_{jk} . The state change vector has two nonzero components: $n_{jk,j} = -1$, $n_{jk,k} = 1$. The diffusion master equation (DME) in a chemical system without reactions is

$$\begin{aligned} \partial_t p(\mathbf{y}, t) &= \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^N q_{jk} (y_{ij} + 1) p(\mathbf{y}_{1\cdot}, \mathbf{y}_{2\cdot}, \dots, \mathbf{y}_i - \mathbf{n}_{jk}, \dots, \mathbf{y}_{M\cdot}, t) \\ &\quad - \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^N q_{jk} y_{ij} p(\mathbf{y}, t) \equiv \mathcal{D} p(\mathbf{y}, t), \end{aligned} \tag{4}$$

defining the diffusion operator \mathcal{D} . The RDME for the PDF of the state of a system with reactions and diffusion is then

$$\partial_t p(\mathbf{y}, t) = (\mathcal{M} + \mathcal{D}) p(\mathbf{y}, t). \tag{5}$$

The jump coefficients q_{jk} are determined by the geometry of \mathcal{V} and the voxels \mathcal{V}_j and the diffusion coefficient γ . If \mathcal{V} is a rectangle in 2D or a rectangular hexahedron in 3D and the voxels are squares or cubes, then the mesh partitioning \mathcal{V} is Cartesian as in Fig. 1b. When \mathcal{V}_j and \mathcal{V}_k share a boundary and the size of an edge in the mesh is Δx , we have $q_{jk} = \gamma / \Delta x^2$. If $\mathcal{V}_j \cap \mathcal{V}_k = \emptyset$, then $q_{jk} = 0$. For a general shape

of \mathcal{V} , the voxels are defined by an unstructured mesh consisting of triangles (2D) as in Fig. 1a or tetrahedra (3D) in Engblom et al. (2009). If there is an edge e_{ij} between node j in \mathcal{V}_j and node k in \mathcal{V}_k , then $q_{jk} > 0$.

The diffusion equation for $u(\mathbf{x}, t)$ with Neumann boundary conditions in d dimensions is

$$\partial_t u = \gamma \sum_{i=1}^d \frac{\partial^2 u}{\partial x_i^2}, \quad \mathbf{x} \in \mathcal{V}, \quad \mathbf{n} \cdot \nabla u = 0, \quad \mathbf{x} \in \partial \mathcal{V}, \tag{6}$$

where $\partial \mathcal{V}$ is the boundary of \mathcal{V} and \mathbf{n} is the normal of $\partial \mathcal{V}$. The equation is discretized in space by a finite element method in Engblom et al. (2009) to derive q_{kj} . Let u_{ij} be the concentration of y_{ij} in \mathcal{V}_j such that $u_{ij} = y_{ij}/V_j$. Then, \mathbf{u}_i . for one species i satisfies after discretization

$$\dot{\mathbf{u}}_{ij} = \sum_{k \in \mathcal{J}(j)} D_{jk} u_{ik} + D_{jj} u_{ij}, \tag{7}$$

where $\mathcal{J}(j)$ is the set of nodes connected to node j by an edge. It is shown in Engblom et al. (2009) that

$$D_{jk} = \frac{S_{jk}}{V_j}, \quad D_{jj} = - \sum_{k, k \neq j} D_{jk} = \sum_{k \in \mathcal{J}(j)} D_{jk}, \quad S_{jk} = S_{kj}, \quad S_{jj} = - \sum_{k, k \neq j} S_{jk}. \tag{8}$$

With a mesh of good quality, $S_{jk} \geq 0$ for $j \neq k$, see Meinecke et al. (2016). The relation between the jump coefficients in (3) and the diffusion coefficients in (7) is

$$q_{kj} = \frac{V_j}{V_k} D_{jk} = \frac{S_{jk}}{V_k}. \tag{9}$$

To simplify the notation, the assumption here is that the diffusion speed is equal for all species. Otherwise, q_{kj} , D_{jk} , and S_{jk} in (9) would be scaled by γ_i/γ to account for the different diffusion coefficients γ_i of the different species i .

A numerical solution of (5) is seldom computationally feasible due to the high dimension of $\mathbf{y} \in \mathbb{Z}_+^{MN}$. Suppose that the lattice for \mathbf{y} has L points in each coordinate direction, i.e., $y_{ij} \in [0, 1, \dots, L]$. Then, the lattice size for \mathbf{y} is L^{MN} . A simplification is possible by first splitting the operator $\mathcal{M} + \mathcal{D}$ into two parts (Hellander et al. 2014; MacNamara and Strang 2016; Strang 1968). Suppose that the solution is known at t^n . A timestep Δt is chosen, and then, the reaction part (2) is integrated from t^n to $t^{n+1} = t^n + \Delta t$ followed by integration of the diffusion part (4) in

1. $\partial_t p_1(\mathbf{y}, t) = \mathcal{M} p_1(\mathbf{y}, t), \quad t \in [t^n, t^{n+1}], \quad p_1(\mathbf{y}, t^n) = p(\mathbf{y}, t^n),$
2. $\partial_t p(\mathbf{y}, t) = \mathcal{D} p(\mathbf{y}, t), \quad t \in [t^n, t^{n+1}], \quad p(\mathbf{y}, t^n) = p_1(\mathbf{y}, t^{n+1}). \tag{10}$

The splitting error in p at t^{n+1} is of $\mathcal{O}(\Delta t)$ in (10). Second-order accuracy is obtained by evaluating the reaction equation at half the step at $t^{n+1/2} = t^n + 0.5\Delta t$, then solving the diffusion Eq. (4) for a full step, and finally taking half a step with (2) (Strang 1968)

1. $\partial_t p_1(\mathbf{y}, t) = \mathcal{M}p_1(\mathbf{y}, t), \quad t \in [t^n, t^{n+1/2}], \quad p_1(\mathbf{y}, t^n) = p(\mathbf{y}, t^n),$
2. $\partial_t p_2(\mathbf{y}, t) = \mathcal{D}p_2(\mathbf{y}, t), \quad t \in [t^n, t^{n+1}], \quad p_2(\mathbf{y}, t^n) = p_1(\mathbf{y}, t^{n+1/2}),$
3. $\partial_t p(\mathbf{y}, t) = \mathcal{M}p(\mathbf{y}, t), \quad t \in [t^{n+1/2}, t^{n+1}], \quad p(\mathbf{y}, t^{n+1/2}) = p_2(\mathbf{y}, t^{n+1}).$ (11)

The solution has been advanced from t^n to t^{n+1} . The error in $p(\mathbf{y}, t^{n+1})$ is of $\mathcal{O}(\Delta t^2)$.

In (2), the reactions occur independently in every voxel without being influenced by the species in the other voxels. Introduce the ansatz

$$p(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N, t) = \prod_{j=1}^N p(\mathbf{y}_j, t) \tag{12}$$

into (2) to arrive at

$$\begin{aligned} 0 &= \partial_t p(\mathbf{y}, t) - \left(\sum_{j=1}^N \sum_{r=1}^R w_r(\mathbf{y}_j - \mathbf{n}_r) p(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_j - \mathbf{n}_r, \mathbf{y}_N, t) \right. \\ &\quad \left. - \sum_{j=1}^N \sum_{r=1}^R w_r(\mathbf{y}_j) p(\mathbf{y}, t) \right) \\ &= \sum_{j=1}^N \prod_{k=1, k \neq j}^N p(\mathbf{y}_k, t) \left(\partial_t p(\mathbf{y}_j, t) - \left(\sum_{r=1}^R w_r(\mathbf{y}_j - \mathbf{n}_r) p(\mathbf{y}_j - \mathbf{n}_r, t) \right. \right. \\ &\quad \left. \left. - \sum_{r=1}^R w_r(\mathbf{y}_j, t) p(\mathbf{y}_j, t) \right) \right). \end{aligned} \tag{13}$$

Hence, for $t \geq t^n$ and $p(\mathbf{y}_j, t^n)$ given, N separate solutions of the CME can be computed

$$\begin{aligned} \partial_t p(\mathbf{y}_j, t) &= \sum_{r=1}^R w_r(\mathbf{y}_j - \mathbf{m}_r) p(\mathbf{y}_j - \mathbf{m}_r, t) \\ &\quad - \sum_{r=1}^R w_r(\mathbf{y}_j) p(\mathbf{y}_j, t), \quad j = 1, \dots, N, \end{aligned} \tag{14}$$

and then combined in (12) in the first step of (10) (or the first and third steps in (11)). The solution is computed N times on a lattice of size L^M . This is smaller than the lattice for the full problem but may still be prohibitively large for numerical solution.

In the same manner, the species diffuse independently of each other in the second step in (10) and (11). Insert

$$p(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M, t) = \prod_{i=1}^M p(\mathbf{y}_i, t) \tag{15}$$

into (4) and rearrange the terms as in (13) to arrive at M separate equations for the diffusion of the species when $t \geq t^n$ and $p(\mathbf{y}_i, t^n)$ is known

$$\begin{aligned} \partial_t p(\mathbf{y}_i, t) = & \sum_{j=1}^N \sum_{k=1}^N q_{jk}(y_{ij} + 1)p(\mathbf{y}_i - \mathbf{n}_{jk}, t) \\ & - \sum_{j=1}^N \sum_{k=1}^N q_{jk}y_{ij}p(\mathbf{y}_i, t), \quad i = 1, \dots, M. \end{aligned} \tag{16}$$

Since the propensity is linear in \mathbf{y} in (16), there exist analytical solutions to the sub-problems with multinomial and Poisson distributions for p , see Jahnke and Huisinga (2007), but in practice they are not so useful due to the size L^N of the lattice.

3 Linear Noise Approximation

The biochemical systems are assumed to have a scaling with a size parameter Ω as in van Kampen (2004), Kurtz (1970), Kurtz (1971) where $\Omega \gg 1$. In chemical applications, Ω can denote a volume or a typical copy number of the species. Then, the copy numbers are rescaled by Ω , $\mathbf{z} = \Omega^{-1}\mathbf{y}$, and the propensities can be written as

$$w_r(\mathbf{y}) = \Omega v_r(\Omega^{-1}\mathbf{y}) = \Omega v_r(\mathbf{z}), \quad r = 1, \dots, R. \tag{17}$$

Equations for approximation of the PDF of a system are derived below. The computational complexity of their solution is polynomial in M and N instead of exponential as in (5), (14), and (16).

3.1 The Mean Value Equation

Let m_{ij} be the expected value $E[Y_{ij}]$ of the copy number Y_{ij} of species i in voxel j with a PDF satisfying the master Eq. (2). Multiply (2) by y_{ij} and sum over \mathbb{Z}_+^N . Then, m_{ij} satisfies the equation (Ferm et al. 2008; van Kampen 2004)

$$\dot{m}_{ij} = \sum_{r=1}^R n_{ri} E[w_r(\mathbf{Y}_j)]. \tag{18}$$

Suppose that every w_r is linear in \mathbf{y} . Then,

$$E[w_r(\mathbf{Y}_{.j})] = w_r(E[\mathbf{Y}_{.j}]) = w_r(\mathbf{m}_{.j}).$$

and Eq. (18) are exact for the mean values. If w_r is nonlinear in $\mathbf{y}_{.j}$, then an approximation is

$$E[w_r(\mathbf{Y}_{.j})] \approx w_r(\mathbf{m}_{.j}).$$

With this approximation, we obtain the reaction rate equations

$$\dot{m}_{ij} = \sum_{r=1}^R n_{ri} w_r(\mathbf{m}_{.j}). \tag{19}$$

The mean value equations scaled by the size parameter $\mu_{ij} = m_{ij} / \Omega$ are

$$\dot{\mu}_{ij} = \sum_{r=1}^R n_{ri} v_r(\boldsymbol{\mu}_{.j}) \equiv v_i(\boldsymbol{\mu}_{.j}). \tag{20}$$

The mean concentration $u_{ij} = \mu_{ij} / V_j$ satisfies

$$\dot{u}_{ij} = V_j^{-1} \sum_{r=1}^R n_{ri} v_r((\mathbf{uV})_{.j}) = V_j^{-1} v_i(V_j \mathbf{u}_{.j}). \tag{21}$$

3.2 The Linear Noise Approximation

The scaled state variable \mathbf{z}_k in voxel k is split into a deterministic part $\boldsymbol{\mu}_k$ and a random part $\boldsymbol{\eta}$ by van Kampen in van Kampen (1976, 2004) for the chemical reactions. The random term is assumed to be proportional to $\Omega^{-1/2}$. The relation between the copy numbers \mathbf{y}_k , the scaled copy numbers \mathbf{z}_k , the fluctuations $\boldsymbol{\eta}$, and the fluctuations in the concentrations $\boldsymbol{\psi} = V_k^{-1} \boldsymbol{\eta}$ in \mathcal{V}_k is

$$\mathbf{y}_{.k} = \Omega \mathbf{z}_{.k} = \Omega (\boldsymbol{\mu}_{.k} + \Omega^{-1/2} \boldsymbol{\eta}) = \Omega V_k (\mathbf{u}_{.k} + \Omega^{-1/2} \boldsymbol{\psi}). \tag{22}$$

This expansion is inserted into master Eq. (2) with the propensities v_r in (17)

$$\begin{aligned} \partial_t p(\Omega \mathbf{z}_{.k}, t) = & \sum_r \Omega (v_r(\mathbf{z}_{.k} - \Omega^{-1} \mathbf{n}_r) p(\Omega (\mathbf{z}_{.k} - \Omega^{-1} \mathbf{n}_r), t) \\ & - v_r(\mathbf{z}_{.k}) p(\Omega \mathbf{z}_{.k}, t)). \end{aligned} \tag{23}$$

Replace p in (23) by Π in

$$\Pi(\boldsymbol{\eta}, t) = p(\Omega \boldsymbol{\mu}_{.k} + \Omega^{1/2} \boldsymbol{\eta}, t) = p(\mathbf{y}_{.k}, t), \tag{24}$$

and expand the right-hand side of (23) in a Taylor series around $\mu_{\cdot k}$. Terms proportional to $\Omega^{1/2}$ vanish since $\mu_{\cdot k}$ satisfies (20). If Π is the solution to

$$\begin{aligned} \Pi_t &= \sum_{r=1}^R \sum_{i=1}^M \left(\Pi v_{r,i} n_{ri} + \sum_{j=1}^M \Pi_{,i} n_{ri} v_{r,j} \eta_j \right) + 0.5 \sum_{i=1}^M \sum_{j=1}^M W_{ij} \Pi_{,ij}, \\ W_{ij}(\mu_{\cdot k}) &= \sum_r n_{ri} n_{rj} v_r(\mu_{\cdot k}), \end{aligned} \tag{25}$$

then terms of $\mathcal{O}(1)$ cancel out. Terms of $\mathcal{O}(\Omega^{-1/2})$ and smaller are ignored in the expansion. This is the *linear noise approximation* (LNA) for the scaled copy numbers subject to chemical reactions in \mathcal{V}_k .

The solution to (25) is the PDF of a normal distribution

$$\Pi(\eta, t) = \frac{1}{(2\pi)^{M/2} \sqrt{\det \Sigma}} \exp \left(-0.5 \sum_{i=1}^M \sum_{j=1}^M \eta_i (\Sigma^{-1})_{ij} \eta_j \right), \tag{26}$$

see Ferm et al. (2008), van Kampen (2004) and (Risken 1996, p. 156). The dimension of η is M in the CMEs (14) for all N voxels. The matrix Σ for the covariance between the species i and j in \mathcal{V}_k is the solution of

$$\dot{\Sigma}_{ij} = \sum_{l=1}^M v_{i,l} \Sigma_{lj} + \sum_{l=1}^M v_{j,l} \Sigma_{li} + W_{ij}(\mu_{\cdot k}). \tag{27}$$

Since η is normally distributed with $\mathbf{0}$ mean and covariance Σ , $\eta \sim \mathcal{N}(\mathbf{0}, \Sigma)$, it follows from (22) that $\mathbf{Y}_{\cdot k}$, $\mathbf{Z}_{\cdot k}$, and the concentration $\mathbf{U}_{\cdot k} = V_k^{-1} \mathbf{Z}_{\cdot k}$ also have normal distributions

$$\mathbf{Y}_{\cdot k} \sim \mathcal{N}(\Omega \mu_{\cdot k}, \Omega \Sigma), \quad \mathbf{Z}_{\cdot k} \sim \mathcal{N}(\mu_{\cdot k}, \Omega^{-1} \Sigma), \quad \mathbf{U}_{\cdot k} \sim \mathcal{N}(V_k^{-1} \mu_{\cdot k}, \Omega^{-1} V_k^{-2} \Sigma). \tag{28}$$

The covariance of $\mathbf{U}_{\cdot k}$ in (28) is denoted by $\Omega^{-1} \Xi = \Omega^{-1} V_k^{-2} \Sigma$. Then, the differential equation satisfied by Ξ follows from (27)

$$\dot{\Xi}_{ij} = \sum_{l=1}^M v_{i,l} \Xi_{lj} + \sum_{l=1}^M v_{j,l} \Xi_{li} + V_k^{-2} W_{ij}(V_k \mathbf{u}_{\cdot k}). \tag{29}$$

There are M nonlinear ODEs to solve in (20) for $\mu_{\cdot k}$ in every voxel \mathcal{V}_k . The covariance matrix Σ is symmetric, and we have to solve $(M + 1)M/2$ linear ODEs in (27) and (29). The structure of this equation is the same also for other approximations of the CME, e.g., EMRE in Smith et al. (2017). The accuracy of mean value Eq. (20) is improved in Ferm et al. (2008) by adding a term which is linear in the covariance.

3.3 The Diffusion Equation

The notation is simplified if we assume here that there is only one species, $M = 1$, but many voxels, $N > 1$. If $M > 1$, then the diffusion of the other species is treated separately in the same manner, see (16). In diffusion master Eq. (16), the propensity to jump from voxel k to j is linear in $\mathbf{y} \in \mathbb{Z}_+^N$ with $w_r(\mathbf{y}) = q_{kj}y_k$ and $v_r(\mathbf{z}) = q_{kj}z_k$. The linearity implies that there are explicit expressions for the mean value equations, $v_{i,k}$ in (27), and W_{ij} in (25).

The equations for the scaled mean values are obtained from (8), (9), and (20)

$$\dot{\mu}_j = \sum_{k=1}^N S_{jk}\mu_k / V_k, \quad j = 1, \dots, N. \tag{30}$$

The diffusion equation for the mean concentration is derived from (21)

$$\dot{u}_j = V_j^{-1} \sum_{l=1}^N q_{lj}V_l u_l = \sum_{l=1}^N S_{jl}u_l / V_j = \sum_{l=1}^N D_{jl}u_l, \quad j = 1, \dots, N, \tag{31}$$

cf. (7).

The equation for the covariance Σ (27) between voxels i and j depends on $v_{j,k}$ and W_{ij} in (25). The derivative $v_{i,k}$ in (27) and (29) is by (8) and (9)

$$\begin{aligned} v_{i,k} &= \sum_{r=1}^R n_{ri}w_{r,k}(\mathbf{y}) = q_{ki} = S_{ik}/V_k, \quad k \neq i, \\ v_{i,i} &= \sum_{r=1}^R n_{ri}w_{r,i}(\mathbf{y}) = \sum_{j \in \mathcal{J}(i)} -q_{ij} = - \sum_{j \in \mathcal{J}(i)} S_{ji}/V_i = S_{ii}/V_i, \end{aligned} \tag{32}$$

since $n_{ri} = 1$ for the jump from k to i and $n_{ri} = -1$ for all jumps from i to every j connected to node i by an edge e_{ij} . Let \mathcal{E} be the set of all edges in the mesh. The state change vector on edge e_{ij} for a jump from i to j is \mathbf{n}_{ij} with the nonzero components $n_{ij;i} = -1$ and $n_{ij;j} = 1$. The contribution to \mathbf{W} in (27) from e_{ij} jumps in two directions: $i \rightarrow j$ and $j \rightarrow i$. Hence, for all edges

$$\mathbf{W} = \sum_{e_{ij} \in \mathcal{E}} \frac{S_{ij}}{V_j} \mu_j \mathbf{n}_{ij} \mathbf{n}_{ij}^T + \frac{S_{ji}}{V_i} \mu_i \mathbf{n}_{ij} \mathbf{n}_{ij}^T = \sum_{e_{ij} \in \mathcal{E}} S_{ij} \left(\frac{\mu_j}{V_j} + \frac{\mu_i}{V_i} \right) \mathbf{n}_{ij} \mathbf{n}_{ij}^T. \tag{33}$$

The nonzero elements of $\mathbf{N}_{ij} = \mathbf{n}_{ij} \mathbf{n}_{ij}^T$ are $N_{ij;ii} = N_{ij;jj} = 1$ and $N_{ij;ij} = N_{ij;ji} = -1$. Therefore, the elements of the symmetric \mathbf{W} are

$$W_{ij} = -S_{ij} \left(\frac{\mu_j}{V_j} + \frac{\mu_i}{V_i} \right), \quad j \neq i, \quad W_{ii} = \sum_{j \in \mathcal{J}(i)} S_{ij} \left(\frac{\mu_j}{V_j} + \frac{\mu_i}{V_i} \right). \tag{34}$$

The random component of the concentrations $\boldsymbol{\psi} = \mathbf{V}^{-1}\boldsymbol{\eta}$ is normally distributed $\mathcal{N}(\mathbf{0}, \mathbf{V}^{-1}\boldsymbol{\Sigma}\mathbf{V}^{-1})$. The equation for $\boldsymbol{\Xi} = \mathbf{V}^{-1}\boldsymbol{\Sigma}\mathbf{V}^{-1}$ is derived from (27)

$$\dot{\boldsymbol{\Xi}}_{ij} = \sum_{l=1}^M V_i^{-1}v_{i,l}V_l\boldsymbol{\Xi}_{lj} + \sum_{l=1}^M V_j^{-1}v_{j,l}V_l\boldsymbol{\Xi}_{li} + V_i^{-1}W_{ij}(\mathbf{V}\mathbf{u})V_j^{-1}. \tag{35}$$

The coefficients in (35) multiplying $\boldsymbol{\Xi}$ are

$$V_i^{-1}v_{i,l}V_l = S_{il}/V_i = D_{il}, \quad l \neq i, \quad V_i^{-1}v_{i,i}V_i = S_{ii}/V_i = D_{ii}. \tag{36}$$

In (35), \mathbf{W} is scaled by \mathbf{V}

$$\begin{aligned} V_i^{-1}W_{ij}V_j^{-1} &= -V_i^{-1}S_{ij}V_j^{-1}(u_i + u_j) = -D_{ij}V_j^{-1}(u_i + u_j), \quad j \neq i, \\ V_i^{-1}W_{ii}V_i^{-1} &= 2 \sum_{j \in \mathcal{J}(i)} V_i^{-1}S_{ij}V_i^{-1}u_i = -2V_i^{-1}S_{ii}V_i^{-1}u_i \\ &= -2D_{ii}V_i^{-1}u_i. \end{aligned} \tag{37}$$

Then, the scaled \mathbf{W} -term in (37) can be written in a symmetric form

$$\begin{aligned} V_i^{-1}W_{ij}V_j^{-1} &= -D_{ij}V_j^{-1}(u_i + u_j) = -S_{ij}V_i^{-1}V_j^{-1}(u_i + u_j) \\ &= -\left(u_i D_{ij}V_j^{-1} + u_j D_{ji}V_i^{-1}\right). \end{aligned} \tag{38}$$

The covariance equation corresponding to (29) for diffusion is by (35), (36), and (38)

$$\begin{aligned} \dot{\boldsymbol{\Xi}}_{ij} &= \sum_{l=1}^N D_{il}\boldsymbol{\Xi}_{lj} + \sum_{l=1}^N D_{jl}\boldsymbol{\Xi}_{li} - f_{ij}, \\ f_{ij} &= u_i D_{ij}V_j^{-1} + u_j D_{ji}V_i^{-1}, \quad i, j = 1, \dots, N. \end{aligned} \tag{39}$$

The copy numbers \mathbf{Y}_k . and \mathbf{Z}_k . are normally distributed as in (28). The covariance of the concentrations in space of a species k , \mathbf{U}_k ., is $\Omega^{-1}\boldsymbol{\Xi}$ and $\mathbf{U}_k. \sim \mathcal{N}(\mathbf{u}_k., \Omega^{-1}\boldsymbol{\Xi})$.

In the stationary equation, $\dot{\boldsymbol{\Xi}}_{ij} = 0$ in (39) and it is a Lyapunov equation. A stationary solution of (31) is $u_i = \text{const.}$ and of (39) is

$$\boldsymbol{\Xi}_{ij} = u_i V_i^{-1} \delta_{ij}, \tag{40}$$

where δ_{ij} is the Kronecker delta. If the initial data $\boldsymbol{\Xi}_{ij}(0)$ are symmetric, then the solution to (39) is symmetric for all $t > 0$. At the stationary solution of (40)

$$\boldsymbol{\Sigma} = \mathbf{V}\boldsymbol{\Xi}\mathbf{V} = \text{diag}(\mathbf{V}\mathbf{u}), \quad \boldsymbol{\mu} = \mathbf{V}\mathbf{u},$$

where $\text{diag}(\mathbf{x})$ is a diagonal matrix with x_i in the diagonal. Thus, the stationary distribution of the copy numbers in the voxels \mathbf{Y}_k . for species k follows from (28)

$$\mathbf{Y}_k. \sim \mathcal{N}(\Omega\mathbf{V}\mathbf{u}, \Omega \text{diag}(\mathbf{V}\mathbf{u})). \tag{41}$$

The stationary copy numbers in different voxels are uncorrelated, have a multivariate normal distribution and are therefore independent, and are approximately Poisson distributed since $Y_{ki} \sim \mathcal{N}(\Omega V_i u_i, \Omega V_i u_i)$ with equal mean and variance. If $\Xi_{ij}(0) = 0$, then the time-dependent solutions to (39) will be proportional to u_i/V_i and the mean and the covariance of Y_{ij} are both proportional to $\Omega V_i u_i$. The distributions of the solution to the DME in (4) are discussed in Engblom et al. (2017) based on the theory for linear propensities in Jahnke and Huisinga (2007). Their distributions are either multinomial, Poisson, or a combination. The stationary distribution is multinomial according to Anderson et al. (2010) and approximately Poissonian (Engblom et al. 2017).

The components of the solution \mathbf{u} of (31) are the node values of the finite element approximation of $u(\mathbf{x}, t)$ solving diffusion Eq. (6) for one species. Let $\Xi(\mathbf{x}_1, \mathbf{x}_2, t)$ be the covariance between the solutions at the coordinates $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d$. Then, Ξ_{ij} in (39) can be interpreted as the value of $\Xi(\mathbf{x}_1, \mathbf{x}_2, t)$ at the nodes at \mathbf{x}_{1i} and \mathbf{x}_{2j} . The coefficient D_{ij} in f_{ij} in (39) is negative when $\mathbf{x}_{1i} = \mathbf{x}_{2j}$, and positive when $\mathbf{x}_{1i} \neq \mathbf{x}_{2j}$. When D_{ij} is nonzero, the difference $\xi_{ij} = \frac{1}{\sqrt{2}}(\mathbf{x}_{1i} - \mathbf{x}_{2j})$ is small. On a regular mesh with a typical length of an edge equal to Δx , the positive weight D_{ij} depends approximately only on $r_{ij} = \|\xi_{ij}\|$ and V_i varies smoothly with a typical size V . On such a mesh, $D_{ij} \propto \Delta x^{-2}$ and $u_i D_{ij} V_j^{-1}$ is approximated here by a continuous function $u(\mathbf{x}_1, t)\varphi(r_{ij})$ with

$$\varphi(r) = -\frac{\gamma}{\Delta x^2 V} \exp\left(\frac{-r^2}{\sigma^2}\right) \cos(\pi \omega r), \tag{42}$$

and the solution $u(\mathbf{x}, t)$ to (6). The scalings σ and ω are chosen such that $\sigma \propto \Delta x$ and $\omega = \Delta x^{-1}$. When $r = 0$ then $D_{ii} V_i \approx \varphi(0) = -\gamma/\Delta x^2 V$ and when $r = \Delta x$ we have

$$D_{ij} V_j^{-1} \approx \varphi(\Delta x) = \frac{\gamma}{\Delta x^2 V} \exp\left(-\left(\frac{\Delta x}{\sigma}\right)^2\right) < |\varphi(0)|.$$

Then, the continuous equation corresponding to discrete Eq. (39) is

$$\partial_t \Xi = \gamma \sum_{i=1}^d \frac{\partial^2 \Xi}{\partial x_{1i}^2} + \frac{\partial^2 \Xi}{\partial x_{2i}^2} - (u(\mathbf{x}_1, t) + u(\mathbf{x}_2, t))\varphi(r). \tag{43}$$

If exact initial conditions of the distribution of molecules are known, then $\Xi(\mathbf{x}_1, \mathbf{x}_2, 0) = 0$.

Let $u_\infty(\mathbf{x})$ be the stationary solution to the diffusion equation. Then, one can show that an approximate stationary solution to (43) is

$$\Xi_\infty(\mathbf{x}_1, \mathbf{x}_2) = \frac{u_\infty(\mathbf{x}_1)}{V} \exp\left(-\frac{\|\mathbf{x}_1 - \mathbf{x}_2\|^2}{2\sigma^2}\right), \tag{44}$$

As $\Delta x \rightarrow 0$, this solution approaches $u_\infty(\mathbf{x}_1)\delta(\|\mathbf{x}_1 - \mathbf{x}_2\|)$ where δ is the Dirac measure. The solution in (40) to discrete Eq. (39) is similar to (44).

4 Analysis of the Covariance Equation

A property of the continuous approximation $\Xi(\mathbf{x}_1, \mathbf{x}_2, t)$ of the covariance in (43) is derived in this section. We show that Ξ decays exponentially when $\|\mathbf{x}_1 - \mathbf{x}_2\|$ grows, indicating that the discrete variance Ξ_{ij} in (39) is small when $\|\mathbf{x}_{1i} - \mathbf{x}_{2j}\| > \Delta x$.

Consider (43) in free space $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d$ and for $t \geq 0$ with initial data $\Xi(\mathbf{x}_1, \mathbf{x}_2, 0)$. The concentration $u(\mathbf{x}_1, t)$ is nonnegative and is assumed to be bounded by C_u for all \mathbf{x}_1 and $t \geq 0$. Introduce a change of variables

$$\begin{pmatrix} \xi_{1j} \\ \xi_{2j} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{1j} \\ x_{2j} \end{pmatrix}, \quad j = 1, \dots, d. \tag{45}$$

The diffusion equation in (43) is in the new variables

$$\partial_t \Xi = \gamma \sum_{j=1}^d \frac{\partial^2 \Xi}{\partial \xi_{1j}^2} + \frac{\partial^2 \Xi}{\partial \xi_{2j}^2} - f(\xi_1, \xi_2, t)\varphi(\|\xi_1\|). \tag{46}$$

Here $f(\xi_1, \xi_2, t) = u(\mathbf{x}_1, t) + u(\mathbf{x}_2, t)$ is nonnegative and bounded by $2C_u$. The factor φ in the source term vanishes quickly when $\|\xi_1\|$ increases.

With the fundamental solution of the diffusion equation in $2d$ dimensions (Evans 2010; Stakgold 2000), the solution of (46) can be written as a sum of two integrals depending on the initial data and the source term

$$\Xi(\xi_1, \xi_2, t) = I_{\text{ini}} + I_{\text{src}}, \tag{47}$$

where

$$I_{\text{ini}} = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{1}{(4\pi\gamma t)^d} \exp\left(-\left(\|\xi_1 - \zeta_1\|^2 + \|\xi_2 - \zeta_2\|^2\right)/4\gamma(t)\right) \Xi(\zeta_1, \zeta_2, 0) d\zeta_1 d\zeta_2, \tag{48}$$

and

$$I_{\text{src}} = - \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\gamma}{(4\pi\gamma(t-s))^d} \exp\left(-\left(\|\xi_1 - \zeta_1\|^2 + \|\xi_2 - \zeta_2\|^2\right)/(4\gamma(t-s))\right) \cdot \frac{f(\zeta_1, \zeta_2, t)}{\Delta x^2 V} \cos(\omega\|\zeta_1\|) \prod_{j=1}^d \exp(-\zeta_{1j}^2/\sigma^2) d\zeta_1 d\zeta_2 ds. \tag{49}$$

The integral with the source term is bounded by

$$\begin{aligned}
 |I_{\text{src}}| &\leq \frac{2C_u\gamma}{\Delta x^2 V} \int_0^t \frac{1}{(4\pi\gamma(t-s))^d} \int_{\mathbb{R}^d} \exp\left(-\|\xi_2 - \zeta_2\|^2/(4\gamma(t-s))\right) d\zeta_2 \\
 &\quad \cdot \int_{\mathbb{R}^d} |\cos(\omega\|\xi_1\|)| \prod_{j=1}^d \exp\left(-(\xi_{1j} - \zeta_{1j})^2/(4\gamma(t-s))\right) \\
 &\quad \times \exp\left(-\zeta_1^2/\sigma^2\right) d\zeta_1 ds. \tag{50}
 \end{aligned}$$

The spatial integral I_{d2} over $\zeta_2 \in \mathbb{R}^d$ in (50) is

$$\begin{aligned}
 I_{d2}(t-s) &= \prod_{j=1}^d \int_{\mathbb{R}} \exp(-(\xi_{2j} - \zeta)^2/(4\gamma(t-s))) d\zeta = \prod_{j=1}^d 2\sqrt{\pi\gamma(t-s)} \\
 &= (4\pi\gamma(t-s))^{d/2}. \tag{51}
 \end{aligned}$$

The integral I_{d1} of the product over $\zeta_1 \in \mathbb{R}^d$ in (50) is

$$\begin{aligned}
 &I_{d1}(t-s, \sigma) \\
 &= \int_{\mathbb{R}^d} \prod_{j=1}^d \exp\left(-(\xi_{1j} - \zeta_{1j})^2/(4\gamma(t-s))\right) \exp\left(-\zeta_{1j}^2/\sigma^2\right) d\zeta_1 \\
 &= \prod_{j=1}^d \int_{\mathbb{R}} \exp\left(-(\xi_j - \zeta)^2/(4\gamma(t-s))\right) \exp\left(-\zeta^2/\sigma^2\right) d\zeta = \prod_{j=1}^d I_{1j}(\sigma). \tag{52}
 \end{aligned}$$

With $\tau = 4\gamma(t-s)$ and $\alpha = \tau^{-1} + \sigma^{-2}$, we have

$$\begin{aligned}
 I_{1j}(\sigma) &= \int_{\mathbb{R}} \exp\left(-(\xi_{1j} - \zeta)^2/\tau - \zeta^2/\sigma^2\right) d\zeta \\
 &= \exp\left(-\xi_{1j}^2/\tau + \xi_{1j}^2/\alpha\tau^2\right) \int_{\mathbb{R}} \exp\left(-\alpha(\zeta - \xi_{1j}/\alpha\tau)^2\right) d\zeta \\
 &= \frac{\exp\left(-\xi_{1j}^2/(\tau + \sigma^2)\right)}{\sqrt{\alpha}} \int_{\mathbb{R}} \exp(-z^2) dz \\
 &= \frac{\sigma\sqrt{\pi\tau}}{\sqrt{\tau + \sigma^2}} \exp\left(-\xi_{1j}^2/(\tau + \sigma^2)\right). \tag{53}
 \end{aligned}$$

Using (53), (52), and (51), a bound on I_{src} in (50) is

$$\begin{aligned}
 |I_{src}| &\leq \frac{2C_u}{\Delta x^2 V} \int_0^t \frac{\gamma}{(4\pi\gamma(t-s))^d} I_{d1}(t-s, \sigma) I_{d2}(t-s) ds \\
 &= \frac{2C_u}{\Delta x^2 V} \int_0^{4\gamma t} \frac{(\pi\tau)^{d/2}}{4(\pi\tau)^d} \prod_{j=1}^d I_{1j}(\sigma) d\tau \\
 &= \frac{C_u\sigma^d}{2\Delta x^2 V} \int_0^{4\gamma t} \frac{\exp(-\|\xi_1\|^2/(\tau + \sigma^2))}{(\tau + \sigma^2)^{d/2}} d\tau \\
 &\leq \frac{C_u\sigma^d}{2\Delta x^2 V} \exp(-\|\xi_1\|^2/(4\gamma t + \sigma^2)) \int_0^{4\gamma t} \frac{d\tau}{(\tau + \sigma^2)^{d/2}} \\
 &= \begin{cases} \frac{C_u\sigma^2}{2\Delta x^2 V} \exp(-\|\xi_1\|^2/(4\gamma t + \sigma^2)) \log\left(1 + \frac{4\gamma t}{\sigma^2}\right), & d = 2, \\ \frac{C_u\sigma^2}{\Delta x^2 V(d-2)} \exp(-\|\xi_1\|^2/(4\gamma t + \sigma^2)) \left(1 - 1/\left(1 + \frac{4\gamma t}{\sigma^2}\right)^{d/2-1}\right), & d = 1 \\ & d \geq 3. \end{cases} \tag{54}
 \end{aligned}$$

Assume that the initial data are localized close to $\xi_1 = 0$ such that $|\Xi(\mathbf{x}_1, \mathbf{x}_2, 0)| \leq \Xi_0 \exp(-\|\xi_1\|^2/\chi^2)$ for some $\chi > 0$. A bound on the integral in (47) due to the initial data is then

$$\begin{aligned}
 |I_{ini}| &\leq \frac{\Xi_0}{(4\pi\gamma t)^d} \int_{\mathbb{R}^d} \exp\left(-\left(\|\xi_1 - \xi_1\|^2/4\gamma t + \|\xi_1\|^2/\chi^2\right)\right) d\xi_1 \\
 &\quad \cdot \int_{\mathbb{R}^d} \exp\left(-\|\xi_2 - \xi_2\|^2/4\gamma t\right) d\xi_2 = \frac{\Xi_0}{(4\pi\gamma t)^d} I_{d1}(t, \chi) I_{d2}(t) \\
 &= \frac{\Xi_0}{(4\pi\gamma t)^d} \prod_{j=1}^d I_{1j}(\chi) \cdot (4\pi\gamma t)^{d/2} \\
 &= \Xi_0 \exp\left(-\|\xi_1\|^2/(4\gamma t + \chi^2)\right) / \left(1 + \frac{4\gamma t}{\chi^2}\right)^{d/2}. \tag{55}
 \end{aligned}$$

Hence, a bound on the covariance solution in (47) is obtained by (54) and (55). The assumptions and conclusions are summarized in a theorem:

Theorem 1 Assume that $|u(\mathbf{x}, t)| \leq C_u$ and that the initial data satisfy

$$|\Xi(\mathbf{x}_1, \mathbf{x}_2, 0)| \leq \Xi_0 \exp(-\|\xi_1\|^2/\chi^2).$$

The relations between the \mathbf{x} and ξ coordinates are $\xi_1 = \frac{1}{\sqrt{2}}(\mathbf{x}_1 - \mathbf{x}_2)$, $\xi_2 = \frac{1}{\sqrt{2}}(\mathbf{x}_1 + \mathbf{x}_2)$. Then, the solution of (46) with φ defined by (42) for $t > 0$ is bounded by

$$\begin{aligned}
 |\Xi(\mathbf{x}_1, \mathbf{x}_2, t)| &\leq \frac{C_u\sigma^2}{\Delta x^2 V} f_d \left(1 + \frac{4\gamma t}{\sigma^2}\right) \exp\left(-(\|\xi_1\|/\sigma)^2/(1 + 4\gamma t/\sigma^2)\right) \\
 &\quad + \frac{\Xi_0}{(1 + 4\gamma t/\chi^2)^{d/2}} \exp\left(-(\|\xi_1\|/\chi)^2/(1 + 4\gamma t/\chi^2)\right), \tag{56}
 \end{aligned}$$

where

$$f_d(\zeta) = \frac{1}{2} \log(\zeta), \quad d = 2, \quad f_d(\zeta) = \frac{1}{d-2} \left(1 - \frac{1}{\zeta^{d/2-1}} \right), \quad d = 1, \quad d \geq 3. \tag{57}$$

□

The function f_d depends on the dimension d and is 0 at $t = 0$. The first term in (56) is proportional to $C_u/V = \sup u/V$ since $\sigma \propto \Delta x$. The solution in (56) decays exponentially in $\|\xi_1\| = \frac{1}{\sqrt{2}}\|\mathbf{x}_1 - \mathbf{x}_2\|$ for a fixed t in all dimensions and is small when $\|\xi_1\| > \sigma \propto \Delta x$. For a given ξ_1 , the first term in (56) increases slowly with t in 1 and 2 dimensions and is bounded by $C_u\sigma^2/\Delta x^2 V(d-2)$ when $d \geq 3$. The second term in (56) decreases when $t \geq 0.5\|\xi_1\|^2 - 0.25\chi^2$ for $d = 1$ and for all $t \geq 0$ when $d \geq 2$.

Our bounded domain \mathcal{V} for \mathbf{x}_1 and \mathbf{x}_2 has a boundary that is not taken into account in (56). The bound on Ξ is a good estimate when the main part of the solution is away from the boundary, e.g., when t is not too large and $u(\mathbf{x}, t)$ is nonzero only in the middle of \mathcal{V} .

Since (43) is a continuous approximation of (39) we expect the discrete variance Ξ_{ij} to behave in a similar way and be negligible when the nodes i and j are not neighbors and not directly connected by an edge in the mesh. This property will be exploited in the algorithm in the next section.

5 Algorithm

The algorithm to compute the solution to the LNA for both reactions and diffusion is based on the operator splitting in Sect. 2, the derivations in Sects. 3.2 and 3.3, and Theorem 1 in Sect. 4.

The mean value equation in (21) is added to the diffusion equation in (31) to obtain the reaction–diffusion equation for the concentration u_{ik} of species i in voxel k with MN components

$$\dot{u}_{ik} = V_k^{-1} v_i(V_k \mathbf{u}_k) + \sum_{\beta=1}^N D_{k\beta} u_{i\beta}, \quad i = 1, \dots, M, \quad j = 1, \dots, N. \tag{58}$$

The covariance between the concentrations of the species i and j in voxels k and l is written as $\Xi_{ij;kl}$ and has $M^2 N^2$ components. The equation satisfied by the covariance is obtained from (29) and (39)

$$\begin{aligned} \dot{\Xi}_{ij;kl} &= \sum_{\alpha=1}^M v_{i,\alpha} \Xi_{\alpha j;kl} + \sum_{\alpha=1}^M v_{j,\alpha} \Xi_{\alpha i;kl} + \delta_{kl} V_k^{-2} W_{ij}(V_k \mathbf{u}_k) \\ &+ \sum_{\beta=1}^N D_{k\beta} \Xi_{ij;\beta l} + \sum_{\beta=1}^N D_{l\beta} \Xi_{ij;\beta k} - \delta_{ij} \left(u_{ik} D_{kl} V_l^{-1} + u_{jl} D_{lk} V_k^{-1} \right), \\ &i, j = 1, \dots, M, \quad k, l = 1, \dots, N. \end{aligned} \tag{59}$$

The reaction source term vanishes if the concentrations are from different voxels, since there is no reaction between molecules in separate voxels. The diffusion source term is zero if the species are different since a diffusion event occurs when the same species changes location between adjacent voxels by a jump.

The equations for the mean and the covariance (58) and (59) are solved by splitting the operator on the right-hand side and advancing the solution one timestep from t^n to $t^{n+1} = t^n + \Delta t$ as in (10) with $u_{ik}(t^n)$ and $\Xi_{ij;kl}(t^n)$ as initial data, cf. (10). The algorithm is

Algorithm 1

$$1. \quad \dot{\tilde{u}}_{ik} = V_k^{-1} v_i (V_k \tilde{\mathbf{u}}_k), \quad t \in [t^n, t^{n+1}], \quad \tilde{u}_{ik}(t^n) = u_{ik}(t^n) \quad (60)$$

$$2. \quad \dot{\tilde{\Xi}}_{ij;kk} = \sum_{\alpha=1}^M v_{i,\alpha} \tilde{\Xi}_{\alpha j;kk} + \sum_{\alpha=1}^M v_{j,\alpha} \tilde{\Xi}_{\alpha i;kk} + V_k^{-2} W_{ij} (V_k \tilde{\mathbf{u}}_k), \\ t \in [t^n, t^{n+1}], \quad \tilde{\Xi}_{ij;kk}(t^n) = \Xi_{ij;kk}(t^n) \quad (61)$$

$$3. \quad \dot{\tilde{\Xi}}_{ij;kl} = \sum_{\alpha=1}^M v_{i,\alpha} \tilde{\Xi}_{\alpha j;kl} + \sum_{\alpha=1}^M v_{j,\alpha} \tilde{\Xi}_{\alpha i;kl}, \quad k \neq l, \\ t \in [t^n, t^{n+1}], \quad \tilde{\Xi}_{ij;kl}(t^n) = \Xi_{ij;kl}(t^n) \quad (62)$$

$$4. \quad \dot{u}_{ik} = \sum_{\beta=1}^N D_{k\beta} u_{i\beta}, \quad t \in [t^n, t^{n+1}], \quad u_{ik}(t^n) = \tilde{u}_{ik}(t^{n+1}) \quad (63)$$

$$5. \quad \dot{\tilde{\Xi}}_{ii;kl} = \sum_{\beta=1}^N D_{k\beta} \Xi_{ii;\beta l} + \sum_{\beta=1}^N D_{l\beta} \Xi_{ii;\beta k} - (u_{ik} D_{kl} V_l^{-1} + u_{il} D_{lk} V_k^{-1}), \\ t \in [t^n, t^{n+1}], \quad \Xi_{ii;kl}(t^n) = \tilde{\Xi}_{ii;kl}(t^{n+1}) \quad (64)$$

$$6. \quad \dot{\tilde{\Xi}}_{ij;kl} = \sum_{\beta=1}^N D_{k\beta} \Xi_{ij;\beta l} + \sum_{\beta=1}^N D_{l\beta} \Xi_{ij;\beta k}, \quad i \neq j, \\ t \in [t^n, t^{n+1}], \quad \Xi_{ij;kl}(t^n) = \tilde{\Xi}_{ij;kl}(t^{n+1}) \quad (65)$$

The discretization error in $u_{ik}(t^{n+1})$ and $\Xi_{ij;kl}(t^{n+1})$ will be of $\mathcal{O}(\Delta t)$. The ODEs in steps 1, 2 and 3 update u and Ξ in a voxel (steps 1, 2) and Ξ between two adjacent voxels (step 3) due to the reactions as in step 1 of (10) and (14). In the ODEs in steps 4, 5 and 6, u and Ξ change due to diffusion between voxels without any influence of the other species as in step 2 of (10) and (16). A more accurate splitting than in Algorithm 1 with an error of $\mathcal{O}(\Delta t^2)$ is possible in the same manner as in (11).

It follows from Theorem 1 that if $\Xi_{ij;kl}(t^n)$ decays rapidly when the nodes \mathbf{x}_k and \mathbf{x}_l are separated then this property is preserved in $\tilde{\Xi}_{ij;kl}(t^{n+1})$ where $C_u > 0$ in (56) in step 2 and $C_u = 0$ without the source term in step 3. Using the same arguments in steps 5 and 6, we find that if $\Xi_{ij;kl}(t^n)$ decays rapidly when $\|\mathbf{x}_k - \mathbf{x}_l\|$ increases, then after one timestep $\Xi_{ij;kl}(t^{n+1})$ also decays rapidly in $\|\mathbf{x}_k - \mathbf{x}_l\|$.

Supported by the analysis in Sect. 4, we assume that $\Xi_{ij;kl}$ is negligible when node l and node k are not neighbors, $l \notin \mathcal{J}(k)$, and we let $\hat{\Xi}_{ij;kl} = 0$ in a sparse approximation of $\Xi_{ij;kl}$. Then, only elements of $\Xi_{ij;kl}$ when $k = l$ and $l \in \mathcal{J}(k)$ need to be stored and updated in $\hat{\Xi}_{ij;kl}$ by Algorithm 1. The sparsity (or nonzero) pattern of $\hat{\Xi}_{ij;kl}$ for each pair i, j is the same as that of \mathbf{S} and \mathbf{D} in (8) since S_{kl} and D_{kl} are nonzero only on the diagonal and if nodes the k and l are neighbors connected by an edge in the mesh and $l \in \mathcal{J}(k)$. Moreover, $\Xi_{ij;kl}$ is symmetric in both i and j and k and l . With M different species and N voxels, $\Xi_{ij;kl}$ in general has $\frac{1}{2}MN(MN + 1)$ different elements but $\hat{\Xi}_{ij;kl}$ has only $C_d M^2 N$ nonzero elements that are necessary to store taking the symmetry into account. The coefficient C_d depends on the dimension and the structure of the mesh. In a Cartesian mesh, $C_d = 2(1D), 3(2D)$, or $4(3D)$ and in an unstructured mesh $C_d = 2$ in 1D but C_d depends on the particular mesh in 2D and 3D. The mean value vectors \mathbf{u} and $\tilde{\mathbf{u}}$ have the dimension MN .

In order to estimate the computational work in the steps of the algorithm, we assume that v_i depends on a limited number of u_{jk} independent of M . Then, there are also a limited number of derivatives $v_{i,j}$ different from zero and independent of M . Thus, the work to compute the right-hand side (RHS) in step 1 in (60) is independent of M and N and it is computed once for every species i and voxel k , i.e., MN times. Since there are a limited number of nonzeros in $v_{i,\alpha}$, the sums and \mathbf{W} in step 2 in (61) are computed independently of M and N . Hence, the work is proportional to $M^2 N$ for the covariances between the species in every voxel. In step 3 in (62), M^2 covariances are computed for every combination of voxels k and l where $\tilde{\Xi}$ is nonzero. This is the case when k and l are neighbors and each k has a limited number of neighbors. This number is independent of N . Therefore, the work to compute the full RHS in step 3 is of $\mathcal{O}(M^2 N)$. The number of $D_{k\beta} \neq 0$ in the RHS in step 4 in (63) is independent of N according to the previous paragraph. The work to determine all derivatives of u_{ik} is then proportional to MN . For $\Xi_{ii;kl}$ to be nonzero in step 5 in (64), voxels k and l are neighbors, $l \in \mathcal{J}(k)$. Furthermore, the products in the sums are nonzero only if $\beta \in \mathcal{J}(k) \cap \mathcal{J}(l)$. The work to calculate the sums is independent of N , and the RHS is computed $\mathcal{O}(MN)$ times. In the same manner, the RHS in step 6 in (65) is computed $\mathcal{O}(M^2 N)$ times. The conclusion is that the work to determine the RHS in the ODEs for u_{ik} and $\hat{\Xi}_{ij;kl}$ in the algorithm has linear complexity in N and is proportional to $M^2 N$.

Since there are additional administrative costs in Algorithm 1, the straightforward algorithm ignoring the sparsity of Ξ will be faster when $N < N_*$ for some small N_* which is problem dependent. However, for $N > N_*$ Algorithm 1 will be the winner and its advantage is greater, the greater the N is.

If the diffusion coefficient is different for different species i , then $D_{k\beta}$ and $D_{l\beta}$ in steps 4, 5 and 6 would depend on i but the algorithm and its properties remain the same.

In summary, the algorithm in words is for one timestep Δt from t^n to $t^{n+1} = t^n + \Delta t$:

1. Solve the ODE in (60) numerically for the mean values with initial data $\mathbf{u}(t^n)$ to obtain $\tilde{\mathbf{u}}(t)$
2. Solve the ODE in (61) numerically for the covariances between the species in the same voxel k with $\tilde{\mathbf{u}}$ from step 1 and initial data $\Xi(t^n)$ to determine $\tilde{\Xi}_{\dots;kk}(t)$

3. Solve the ODE in (62) numerically for the covariances between the species in different voxels k and l satisfying $l \in \mathcal{J}(k)$ with $\tilde{\mathbf{u}}$ from step 1 and initial data $\tilde{\Xi}(t^n)$ to determine $\tilde{\Xi}_{\cdot,;kl}(t)$
4. Solve the ODE in (63) numerically for the mean values with initial data $\tilde{\mathbf{u}}(t^{n+1})$ from (60) to obtain $\mathbf{u}(t)$ and $\mathbf{u}(t^{n+1})$
5. Solve the ODE in (64) numerically for the covariances between voxels k and l satisfying $l \in \mathcal{J}(k)$ for the same species i with \mathbf{u} from step 4 and initial data $\tilde{\Xi}(t^{n+1})$ from steps 2 and 3 to determine $\Xi_{ii;\cdot}(t^{n+1})$
6. Solve the ODE in (65) numerically for the covariances between voxels k and l satisfying $l \in \mathcal{J}(k)$ for different species i and j with \mathbf{u} from step 4 and initial data $\tilde{\Xi}(t^{n+1})$ from steps 2 and 3 to determine $\Xi_{ij;\cdot}(t^{n+1})$

In the first three steps, the mean values and the covariances change due to the reactions and in the last three steps due to the diffusion.

Theorem 1 and numerical experiments in Sect. 6.1 indicate that the accuracy in $\hat{\Xi}$ increases when the dimension grows. By storing and updating only the sparse approximation in steps 3, 5 and 6 in Algorithm 1, considerable savings are possible in computing time and computer memory when N is large, e.g., in 3D.

5.1 Example

Consider the reversible reaction for association and dissociation of the species A , B , and C



with copy numbers $\boldsymbol{\mu}_{\cdot,j}^T = (a_j, b_j, c_j)$ in voxel j and propensities and state change vectors

$$v_1 = k_a' a_j b_j, \quad \mathbf{n}_1^T = (-1, -1, 1), \quad v_2 = k_d' c_j, \quad \mathbf{n}_2^T = (1, 1, -1). \tag{67}$$

The macroscopic reaction coefficients are $k_a = V_k k_a'$ and $k_d = k_d'$. Then, Eq. (21) for the concentrations in step 1 of the above algorithm in \mathcal{V}_k is

$$\dot{\mathbf{u}}_k = \begin{pmatrix} -k_a u_{1k} u_{2k} + k_d u_{3k} \\ -k_a u_{1k} u_{2k} + k_d u_{3k} \\ k_a u_{1k} u_{2k} - k_d u_{3k} \end{pmatrix}, \quad k = 1, \dots, N. \tag{68}$$

Order the means and the covariances such that

$$\mathbf{u} = \begin{pmatrix} \mathbf{u}_1. \\ \mathbf{u}_2. \\ \mathbf{u}_3. \end{pmatrix}, \quad \Xi = \begin{pmatrix} \Xi_{11;\cdot} & \Xi_{12;\cdot} & \Xi_{13;\cdot} \\ \Xi_{21;\cdot} & \Xi_{22;\cdot} & \Xi_{23;\cdot} \\ \Xi_{31;\cdot} & \Xi_{32;\cdot} & \Xi_{33;\cdot} \end{pmatrix}. \tag{69}$$

The Jacobian \mathbf{J} of the propensities with $J_{ij} = v_{i,j}$ in (31) and the source term in step 2 in \mathcal{V}_k are

$$\mathbf{J} = \begin{pmatrix} -k_a u_{2k} & -k_a u_{1k} & k_d \\ -k_a u_{2k} & -k_a u_{1k} & k_d \\ k_a u_{2k} & k_a u_{1k} & -k_d \end{pmatrix},$$

$$g_k(\mathbf{u}) = V_k^{-1}(k_a u_{1k} u_{2k} + k_d u_{3k}), \quad V_k^{-2} \mathbf{W} = g_k(\mathbf{u}) \mathbf{n}_1 \mathbf{n}_1^T, \tag{70}$$

since $\mathbf{n}_1 \mathbf{n}_1^T = \mathbf{n}_2 \mathbf{n}_2^T$. Introduce \mathbf{K} and \mathbf{G} using the identity matrix \mathbf{I}_N of size N , \mathbf{J} , and \mathbf{W} in (70)

$$\mathbf{K} = \begin{pmatrix} -k_a \text{diag}(\mathbf{u}_2) & -k_a \text{diag}(\mathbf{u}_1) & k_d \mathbf{I}_N \\ -k_a \text{diag}(\mathbf{u}_2) & -k_a \text{diag}(\mathbf{u}_1) & k_d \mathbf{I}_N \\ k_a \text{diag}(\mathbf{u}_2) & k_a \text{diag}(\mathbf{u}_1) & -k_d \mathbf{I}_N \end{pmatrix}$$

$$= \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \otimes (k_a \text{diag}(\mathbf{u}_2) k_a \text{diag}(\mathbf{u}_1) - k_d \mathbf{I}_N),$$

$$\mathbf{G}(\mathbf{u}) = \mathbf{n}_1 \mathbf{n}_1^T \otimes \text{diag}(\mathbf{g}(\mathbf{u})), \tag{71}$$

where \otimes denotes the Kronecker product. Then, the equation in steps 2 and 3 in matrix form is

$$\dot{\tilde{\mathbf{E}}} = \mathbf{K} \tilde{\mathbf{E}} + (\mathbf{K} \tilde{\mathbf{E}})^T + \mathbf{G}(\tilde{\mathbf{u}}). \tag{72}$$

Define the matrices \mathbf{D}_3 and \mathbf{H} by

$$\mathbf{D}_3 = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D} \end{pmatrix} = \mathbf{I}_3 \otimes \mathbf{D},$$

$$\mathbf{H}(\mathbf{u}) = \begin{pmatrix} \mathbf{H}_{1;\dots} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{2;\dots} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_{3;\dots} \end{pmatrix}, \quad H_{i;kl} = D_{kl} V_l^{-1}(u_{ik} + u_{il}). \tag{73}$$

The submatrix \mathbf{D} is the approximation of the Laplacian in (31) and (39). If the diffusion varies between the species, then \mathbf{D} in (73) would be replaced by $\gamma_i/\gamma \mathbf{D}$, $i = 1, 2, 3$, on the diagonal. The sparsity or nonzero pattern in $\mathbf{H}_{i;\dots}$ is the same as in \mathbf{D} . The diffusion equation for the mean values in step 4 in Algorithm 1 is as in (31)

$$\dot{\mathbf{u}} = \mathbf{D}_3 \mathbf{u}. \tag{74}$$

The matrix form of steps 5 and 6 in Algorithm 1 is (cf. (38) and (39))

$$\dot{\tilde{\mathbf{E}}} = \mathbf{D}_3 \tilde{\mathbf{E}} + (\mathbf{D}_3 \tilde{\mathbf{E}})^T - \mathbf{H}(\mathbf{u}). \tag{75}$$

In 1D, \mathbf{D} is a tridiagonal matrix and if $V = \Delta x$ is constant then

$$D_{\alpha k} = D_{k\alpha} = \gamma/\Delta x^2, \quad \alpha = k + 1, \quad k = 1, \dots, N - 1,$$

$$D_{kk} = -2\gamma/\Delta x^2, \quad k = 2, \dots, N - 1, \quad D_{11} = D_{NN} = -\gamma/\Delta x^2.$$

6 Numerical Results

The algorithm is tested for computing the mean and the approximation of the covariance in the LNA of systems with diffusion in 1D, 2D, and 3D and a system in 2D with the reversible reaction (66).

6.1 Diffusion

A Cartesian grid in d dimensions is generated with a constant step size Δx and a diffusion coefficient $\gamma = 0.01$. The number of dimensions is $d = 1, 2, 3$, and the domain is the unit cube $[0, 1]^d$. The number of grid points is $n = 1/\Delta x + 1$ in each dimension yielding $N = n^d$ components in \mathbf{u} . A straightforward implementation of Algorithm 1 in steps 5 and 6 will generate N^2 elements in Ξ . By updating only those elements of Ξ that correspond to nonzeros in \mathbf{D} and \mathbf{S} , the number of nonzero elements in the approximation $\hat{\Xi}$ will be of $\mathcal{O}(N)$.

The initial data $\mathbf{u}(0)$ are sampled from a uniform distribution $u_k(0) \sim \mathcal{U}[0, 1]$ and $\Xi = \mathbf{0}$. The ODEs in (74) and (75) are solved numerically for $t \geq 0$ by the forward Euler method for simplicity. Then, the RHS in each step of Algorithm 1 is evaluated once requiring a computational work proportional to $M^2 N$ in every timestep from t^n to t^{n+1} . Better numerical accuracy is achieved by splitting the computations according to Strang (1968) as in (11) and by using a higher-order method. Better numerical stability is obtained by an implicit method.

6.1.1 1D

The covariance $\Xi(x_1, x_2, t)$ is computed in 1D on a grid with $\Delta x = 0.025$ and $N = n = 41$ using the full Ξ without zeros, the sparse $\hat{\Xi}$ with the same nonzero pattern as \mathbf{S} , i.e., the diagonal, the subdiagonal, and the superdiagonal are nonzero in a tridiagonal matrix as proposed in Sect. 6.1, and the sparse $\check{\Xi}$ where another two diagonals below and above the diagonal are nonzero in a pentadiagonal matrix. One row of the three matrices is shown in Fig. 2. The approximations $\hat{\Xi}(0.5, x_2, t)$ and $\check{\Xi}(0.5, x_2, t)$ agree fairly well with $\Xi(0.5, x_2, t)$ in particular for larger t in Fig. 2.

The PDF of the multivariate normal distribution $\mathcal{N}(\mathbf{u}, \Xi)$ is

$$p(\boldsymbol{\eta}, t) = \frac{1}{(2\pi)^{N/2} \sqrt{\det \Xi}} \exp\left(-\frac{1}{2}(\boldsymbol{\eta} - \mathbf{u})^T \Xi^{-1}(\boldsymbol{\eta} - \mathbf{u})\right). \quad (76)$$

The covariance matrix is factorized by $\Xi = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^T$ where \mathbf{Q} is orthogonal and $\boldsymbol{\Lambda}$ has the positive eigenvalues of Ξ on the diagonal. The expression in (76) in the exponential defines surfaces of ellipsoids in \mathbb{R}^N with equal probability, and the eigenvalues of Ξ are the lengths of the principal axes of the ellipsoids. Another way of comparing Ξ and its approximations is then to compare the eigenvalues to see the difference between the lengths of these axes.

The eigenvalues of Ξ , $\hat{\Xi}$, and $\check{\Xi}$ are displayed in Fig. 3. They also agree well except for the one or two smallest ones in the figure. Using five diagonals and four neighbors

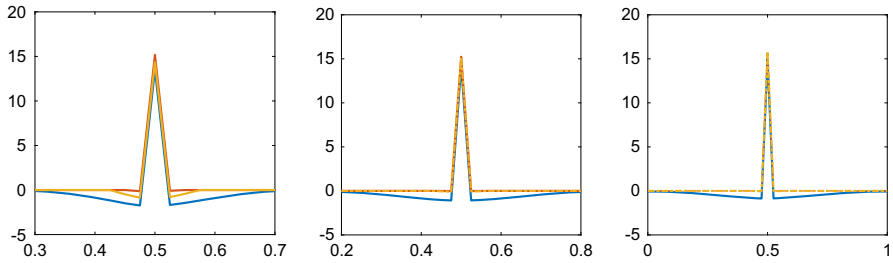


Fig. 2 Comparison of the covariance $\Xi(0.5, x_2, t)$ for diffusion in 1D with $x_2 \in [0, 1]$ on the abscissa at different time points and different approximations: without sparse approximation (blue), with two extra diagonals in $\hat{\Xi}(0.5, x_2)$ (red), and with four extra diagonals in $\check{\Xi}(0.5, x_2)$ (yellow). The time is $t = 8$ (left), $t = 20$ (middle), and $t = 40$ (right)

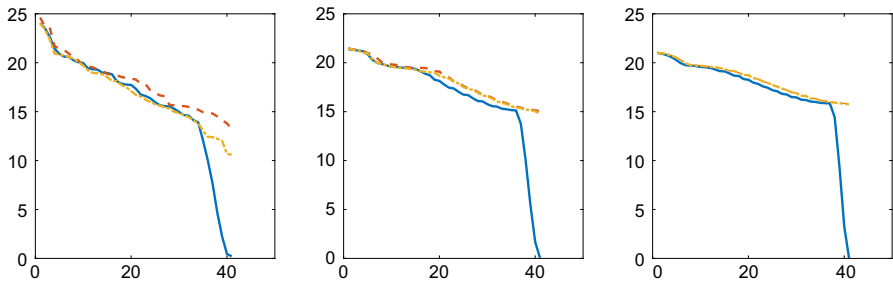


Fig. 3 Comparison of the N eigenvalues of the covariance $\Xi(0.5, x_2, t)$ for diffusion in 1D at different time points and different approximations: without sparse approximation (solid blue), with two extra diagonals in $\hat{\Xi}$ (dashed red), and with four extra diagonals in $\check{\Xi}$ (dash-dotted yellow). The time is $t = 8$ (left), $t = 20$ (middle), and $t = 40$ (right). The curves for $\hat{\Xi}$ and $\check{\Xi}$ at $t = 40$ are indistinguishable with the precision in the figure

in $\check{\Xi}$ improves the approximation somewhat compared to $\hat{\Xi}$. Including more than the nearest neighbors in 2D and 3D with an unstructured mesh is possible but makes the algorithm more complicated.

6.1.2 2D and 3D

In 2D, $\Xi(\mathbf{x}_1, \mathbf{x}_2, t)$ is computed with the full Ξ matrix and with the approximation $\hat{\Xi}$ that has the same sparsity pattern as \mathbf{S} on a grid with $\Delta x = 0.05$ and $N = n^2 = 441$. One row of Ξ corresponds to one coordinate \mathbf{x}_{1k} and its covariance with the 2D \mathbf{x}_2 . The variance is high at $\Xi(\mathbf{x}_1, \mathbf{x}_1, t)$, and the covariance $\Xi(\mathbf{x}_1, \mathbf{x}_2, t)$ is very low when $\mathbf{x}_1 \neq \mathbf{x}_2$. This is depicted in the left panel of Fig. 4 where \mathbf{x}_1 and x_{22} are fixed and x_{21} varies in $\mathbf{x}_2^T = (x_{21}, x_{22})$. The differences in covariance between Ξ and $\hat{\Xi}$ are very small and not visible in the figure. Since Ξ is symmetric, the result is similar in other directions in \mathbf{x}_2 and for other \mathbf{x}_1 . The steady-state solution (40) with $u = 0.5$ and $V = 1/400$ is here 200.

One section of the 3D covariances Ξ and $\hat{\Xi}$ is shown in the right panel of Fig. 4. As in 2D, \mathbf{x}_1, x_{22} , and x_{23} in $\mathbf{x}_2^T = (x_{21}, x_{22}, x_{23})$ are fixed and $\Xi(\mathbf{x}_1, \mathbf{x}_2, t)$ is plotted as a function of x_{21} . The step size in the grid is $\Delta x = 0.0833$ and $N = n^3 = 2197$.

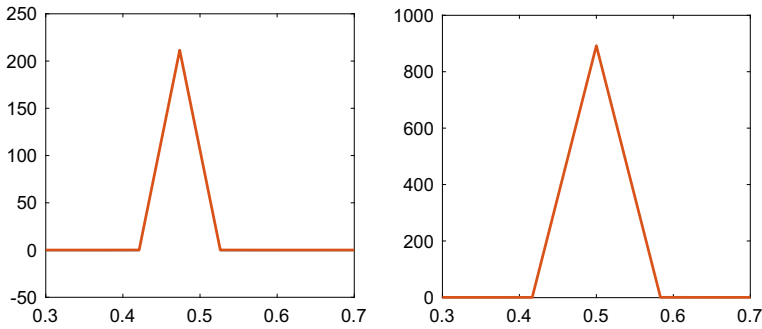


Fig. 4 The covariances $\Xi(\mathbf{x}_1, \mathbf{x}_2, t)$ for diffusion computed with the full matrix Ξ and its sparse approximation $\hat{\Xi}$ in 2D (left) and 3D (right). The coordinate \mathbf{x}_1 chosen in the middle of the domain and with one degree of freedom in \mathbf{x}_2 on the abscissa at $t = 2.5$ with $\Delta x_2 = 1/20$ (left) and $t = 0.2$ with $\Delta x_2 = 1/12$ (right). The difference between Ξ and $\hat{\Xi}$ is not discernible in the figures

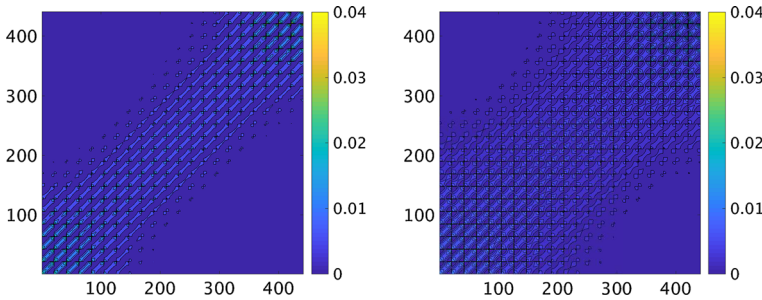


Fig. 5 Scaled difference $\Delta \Xi_{ij}$ of the covariance matrices Ξ and $\hat{\Xi}$ in (77) when $i, j = 1, \dots, N$ for diffusion in 2D at $t = 2.5$ (left) and $t = 5$ (right)

After a short time, the covariances of Ξ and $\hat{\Xi}$ agree very well as in 2D. The stationary solution in (40) with the data here is 864.

The scaled difference $\Delta \Xi$ between the covariances Ξ and $\hat{\Xi}$ is defined by

$$\Xi = \hat{\Xi}(\mathbf{I}_N + \Delta \Xi), \quad \Delta \Xi = \hat{\Xi}^{-1}(\Xi - \hat{\Xi}). \tag{77}$$

The dominant elements in $\hat{\Xi}$ are the variances on the diagonal. With small elements in $\Delta \Xi$ compared to 1, the difference between the covariances in Ξ and $\hat{\Xi}$ is small relative to the variances in $\hat{\Xi}$. In Fig. 5, $\Delta \Xi$ for the 2D example is shown at two time points. The values of $\Delta \Xi$ are low in blue color in most parts of the matrix. The peaks in the left panel are at 0.035 in isolated points. In the right panel, $\max \Delta \Xi_{ij} < 0.02$.

The eigenvalues of Ξ and $\hat{\Xi}$ in 2D and 3D are compared in Fig. 6. The sparse approximation captures all the eigenvalues except for one or two of the smallest ones.

The covariance of the fluctuations in concentration between different parts of the domain is well approximated by the sparse $\hat{\Xi}$, especially in 2D and 3D in Figs. 4 and 6. This is expected from Theorem 1 in Sect. 4 where the decay of $\Xi(\mathbf{x}_1, \mathbf{x}_2, t)$ is slower in 1D than in 2D and 3D when $\xi_1 = \frac{1}{\sqrt{2}}(\mathbf{x}_1 - \mathbf{x}_2)$ is growing.

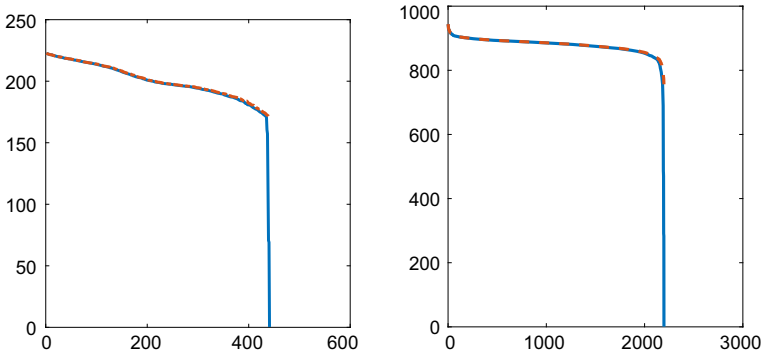


Fig. 6 Comparison of the N eigenvalues of the covariance matrices Ξ and $\hat{\Xi}$ for diffusion in 2D at $t = 2.5$ (left) and 3D at $t = 0.2$ (right) and different approximations: with sparse approximation (dashed red) and without (solid blue)

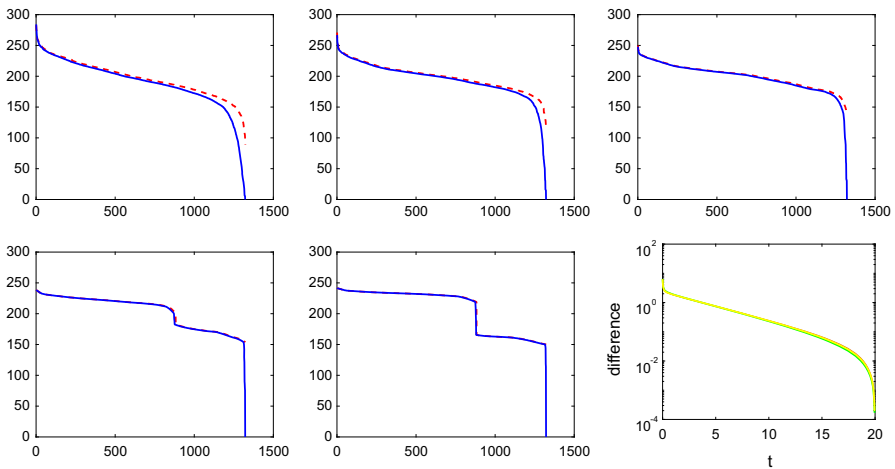


Fig. 7 Comparison of the MN eigenvalues of the covariance matrices Ξ and $\hat{\Xi}$ for the reaction (66) and diffusion in 2D at $t = 0.25$ (upper left), $t = 0.5$ (upper middle), $t = 1$ (upper right), $t = 3$ (lower left), $t = 6$ (lower middle) and different approximations: with sparse approximation (solid blue) and without (dashed red). The convergence of the solution to the steady state for the species A , B , and C as a function of t (lower right)

6.2 Reactions and Diffusion in 2D

The time evolution of the chemical reaction (66) on the Cartesian mesh in 2D in Sect. 6.1.2 is computed with the LNA as in the example in Sect. 5.1. The parameters are $k_a = k_d = 0.1$, and the diffusion is low with $\gamma = 0.01$. The dimension of \mathbf{u} is $MN = Mn^2 = 1323$, and the initial values in $\mathbf{u}(0)$ are uniformly distributed between 0 and 1 and $\Xi(0) = \mathbf{0}$.

The eigenvalues of the covariances with the full matrix Ξ and with the sparse matrix $\hat{\Xi}$ are compared at different t in Fig. 7. An approximate stationary solution \mathbf{u}_∞ is determined at $t = 20$. The convergence of the three subvectors \mathbf{u}_1 , \mathbf{u}_2 , and \mathbf{u}_3 .

in (69) corresponding to the concentrations of A , B , and C is displayed in the lower right panel in the figure. The difference between \mathbf{u} and \mathbf{u}_∞ is measured in $\|\cdot\|$ for the species. In the resolution of the figure, it is not possible to distinguish between the differences in convergence between the species. The balance equation $k_d \bar{a} \bar{b} = k_d \bar{c}$ is satisfied with a relative error less than 0.008 by the mean values \bar{a} , \bar{b} , and \bar{c} of the components in $\mathbf{u}_i(20)$, $i = 1, 2, 3$. At ∞ , $\mathbf{u}_{\infty 1}$, $\mathbf{u}_{\infty 2}$, and $\mathbf{u}_{\infty 3}$ are constant in space.

The convergence plot in Fig. 7 shows that the variation in the solution is larger for small t and decreases with t . The covariances Ξ and $\hat{\Xi}$ agree well for large eigenvalues for small t and they agree well for all eigenvalues when t grows. The off-diagonal submatrices $\Xi_{ij;\dots}$, $i \neq j$, in (69) are comparable in size to the diagonal submatrices $\Xi_{ii;\dots}$ when t is small but as t grows $\Xi_{ii;\dots}$, $i = 1, 2, 3$, will dominate and be closer and closer to diagonal matrices. There is a jump in the spectrum for larger t , e.g., at $t = 6$. This is explained by the difference in the size of the stationary values \bar{a} , \bar{b} , and \bar{c} where $\bar{a} \approx \bar{b}$, and $\bar{c}/\bar{a} \approx 0.6$. The approximation in the covariance equation behaves in the same way with reactions as in Sect. 6.1.2 without the reactions.

7 Conclusions

The master equation is a model for biochemical reactions and diffusion but the numerical solution of it is impossible except for simple, well-stirred systems with special properties. An alternative for large systems with spatial variation is to use the linear noise approximation (LNA). We have derived the equations for the LNA for diffusion and chemical reactions on general meshes. The reactions involve M species, and the mesh consists of N voxels. The covariance of the concentrations is approximated by a sparse representation in an algorithm such that the computational complexity is reduced from $\mathcal{O}(M^2 N^2)$ in a straightforward implementation to $\mathcal{O}(M^2 N)$ here. Also the memory to store the solution is reduced in the same way. The approximation is supported by analytical expressions showing that the higher the dimension is, the better the approximation is. Consequently, the quality of the approximation and the savings in work and storage are more prominent in 3D when N is large. The accuracy of the approximation is evaluated by comparing the elements and the eigenvalues of the full covariance matrix and its sparse approximation in numerical examples with only diffusion in 1D, 2D, and 3D and an example in 2D with a reversible reaction and slow diffusion.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Anderson DF, Craciun G, Kurtz TG (2010) Product-form stationary distributions for deficiency zero chemical reaction networks. *Bull Math Biol* 72:1947–1970

- Asllani M, Biancalani T, Fanelli D, McKane AJ (2013) The linear noise approximation for reaction–diffusion systems on networks. *Eur Phys J B* 86:476
- Biancalani T, Fanelli D, Patti FD (2010) Stochastic Turing patterns in the Brusselator model. *Phys Rev E* 86:046215
- Butler T, Goldenfeld N (2009) Robust ecological pattern formation induced by demographic noise. *Phys Rev E* 80:030902(R)
- Cao Y, Gillespie D, Petzold L (2005) Multiscale stochastic simulation algorithm with stochastic partial equilibrium assumption for chemically reacting systems. *J Comput Phys* 206:395–411
- Cao Y, Gillespie D, Petzold L (2006) Efficient step size selection for the tau-leaping simulation method. *J Chem Phys* 124:044109
- Challenger JD, McKane AJ, Pahle J (2012) Multi-compartment linear noise approximation. *J Stat Mech* 2012(11):P11010
- de Anna P, Di Patti F, Fanelli D, McKane AJ, Dauxios T (2010) Spatial model of autocatalytic reactions. *Phys Rev E* 81:056110
- Drawert B, Engblom S, Hellander A (2012) URDME: a modular framework for stochastic simulation of reaction-transport processes in complex geometries. *BMC Syst Biol* 6:76
- Drawert B, Hellander A, Bales B, Banerjee D, Bellesia G, Daigle B, Douglas G, Gu M, Gupta A, Hellander S, Horuk C, Narayanasamy G, Nath D, Wu S, Lötstedt P, Krintz C, Petzold L (2016) Stochastic simulation service: bridging the gap between the computational expert and the biologist. *PLoS Comput Biol* 12:e1005220
- E W, Liu D, Vanden-Eijnden E (2007) Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates. *J Comput Phys* 221:158–180
- Elf J, Ehrenberg M (2003) Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res* 13:2475–2484
- Elf J, Ehrenberg M (2004) Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Syst Biol* 1:230–236
- Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* 297:1183–1186
- Engblom S, Ferm L, Hellander A, Lötstedt P (2009) Simulation of stochastic reaction–diffusion processes on unstructured meshes. *SIAM J Sci Comput* 31:1774–1797
- Engblom S, Hellander A, Lötstedt P (2017) Multiscale simulation of stochastic reaction–diffusion networks. In: Holcman D (ed) *Stochastic dynamical systems, multiscale modeling, asymptotics and numerical methods for computational cellular biology*. Springer, Cham, pp 55–79
- Engblom S, Lötstedt P, Meinecke L (2017) Mesoscopic modeling of random walk and reactions in crowded media. In: Technical report, Department of Information Technology, Uppsala University, Uppsala. [arXiv:1707.05998](https://arxiv.org/abs/1707.05998)
- Evans LC (2010) *Partial differential equations*, Graduate studies in mathematics, 2nd edn. American Mathematical Society, Providence, RI
- Fange D, Elf J (2006) Noise-induced min phenotypes in *E. coli*. *PLoS Comput Biol* 2:637–648
- Fearnhead P, Giagos V, Sherlock C (2014) Inference for reaction networks using the linear noise approximation. *Biometrics* 70:457–466
- Ferm L, Lötstedt P, Hellander A (2008) A hierarchy of approximations of the master equation scaled by a size parameter. *J Sci Comput* 34:127–151
- Fröhlich F, Thomas P, Kazerooni A, Thies FJ, Grima R, Hasenauer J (2016) Inference for stochastic chemical kinetics using moment equations and systems size expansion. *PLoS Comput Biol* 12:e1005030
- Gardiner CW (2004) *Handbook of stochastic methods*, Springer series in synergetics, 3rd edn. Springer, Berlin
- Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem* 104(9):1876–1889
- Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys* 22(4):403–434
- Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81:2340–2361
- Gillespie DT (1992) A rigorous derivation of the chemical master equation. *Phys A* 188:404–425
- Gillespie DT (2001) Approximate accelerated stochastic simulation of chemically reacting systems. *J Chem Phys* 115(4):1716–1733
- Gillespie DT, Seitaridou E (2013) *Simple Brownian diffusion*. Oxford University Press, Oxford

- Gillespie DT, Hellander A, Petzold LR (2013) Perspective: stochastic algorithms for chemical kinetics. *J Chem Phys* 138:170901
- Gillespie DT, Petzold LR, Seitaridou E (2014) Validity conditions for stochastic chemical kinetics in diffusion-limited systems. *J Chem Phys* 140:054111
- Goutsias J, Jenkinson G (2013) Markovian dynamics on complex reaction networks. *Phys Rep* 529:199–264
- Grima R (2010) An effective rate equation approach to reaction kinetics in small volumes: theory and application to biochemical reactions in nonequilibrium steady-state conditions. *J Chem Phys* 133:035101
- Grima R (2012) A study of the moment-closure approximations for stochastic chemical kinetics. *J Chem Phys* 136:154105
- Grima R (2015) Linear-noise approximation and the chemical master equation agree up to second-order moments for a class of chemical systems. *Phys Rev E* 92:042124
- Hattne J, Fange D, Elf J (2005) Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics* 21:2923–2924
- Hellander A, Lawson MJ, Drawert B, Petzold L (2014) Local error estimates for adaptive simulation of the reaction–diffusion master equation via operator splitting. *J Comput Phys* 266:89–100
- Isaacson SA, Peskin CS (2006) Incorporating diffusion in complex geometries into stochastic chemical kinetics simulations. *SIAM J Sci Comput* 28:47–74
- Jahnke T, Huisinga W (2007) Solving the chemical master equation for monomolecular reaction systems analytically. *J Math Biol* 54:1–26
- Kurtz TG (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes. *J Appl Prob* 7:49–58
- Kurtz TG (1971) Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J Appl Prob* 8:344–356
- Lampoudi S, Gillespie D, Petzold L (2009) The multinomial simulation algorithm for discrete stochastic simulation of reaction–diffusion systems. *J Chem Phys* 130:094104
- MacNamara S, Strang G (2016) Operator splitting. In: Glowinski R, Osher S, Yin W (eds) *Splitting methods in communication, imaging, science, and engineering*. Springer, Cham, pp 95–114
- Mahmutovic A, Fange D, Berg OG, Elf J (2012) Lost in presumption: stochastic reactions in spatial models. *Nat Meth* 9(12):1–4
- McAdams HH, Arkin A (1997) Stochastic mechanisms in gene expression. *Proc Natl Acad Sci USA* 94:814–819
- McKane AJ, Biancalani T, Rogers T (2014) Stochastic pattern formation and spontaneous polarisation: the linear noise approximation and beyond. *Bull Math Biol* 76:895–921
- Meinecke L, Engblom S, Hellander A, Lötstedt P (2016) Analysis and design of jump coefficients in discrete stochastic diffusion models. *SIAM J Sci Comput* 38:A55–A83
- Raj A, van Oudenaarden A (2008) Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135(2):216–226
- Risken H (1996) *The Fokker–Planck equation*. Springer, Berlin
- Rutter A, Opper M (2009) Efficient statistical inference for stochastic reaction processes. *Phys Rev Lett* 103:230601
- Schnoerr D, Sanguinetti G, Grima R (2015) Comparison of different moment-closure approximations for stochastic chemical kinetics. *J Chem Phys* 143:185101
- Schnoerr D, Sanguinetti G, Grima R (2017) Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. *J Phys A Math Theor* 50:093001
- Scott M, Poulin FJ, Tang H (2010) Approximating intrinsic noise in continuous multispecies models. *Proc R Soc A* 467:718–737
- Smith S, Cianci C, Grima R (2017) Analytical approximations for spatial stochastic gene expression in single cells and tissues. *J R Soc Interface* 13:20151051
- Sokolowski TR, ten Wolde PR (2017) Spatial-stochastic simulation of reaction–diffusion systems. Technical report [arXiv:1705.08669v1](https://arxiv.org/abs/1705.08669v1)
- Stakgold I (2000) *Boundary value problems of mathematical physics, vol II*. SIAM, Philadelphia
- Stathopoulos V, Girolami MA (2013) Markov chain Monte Carlo inference for Markov jump processes via the linear noise approximation. *Philos Trans R Soc A* 371:20110541
- Strang G (1968) On the construction and comparison of difference schemes. *SIAM J Numer Anal* 5:506–517
- Sturrock M, Hellander A, Matzavinos A, Chaplain MA (2013) Spatial stochastic modelling of the Hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation. *J R Soc Interface* 10:20120988

- Takahashi K, Tănase-Nicola S, ten Wolde PR (2010) Spatio-temporal correlations can drastically change the response of a MAPK pathway. *Proc Natl Acad Sci USA* 107(6):2473–2478
- Thomas P, Straube AV, Timmer J, Fleck C, Grima R (2013) Signatures of nonlinearity in single cell noise-induced oscillations. *J Theor Biol* 335:222–234
- Tsimring LS (2014) Noise in biology. *Rep Prog Phys* 77:026601
- Ullah M, Wolkenhauer O (2009) Investigating the two-moment characterisation of subcellular biochemical networks. *J Theor Biol* 260:340–352
- van Kampen NG (1976) The expansion of the master equation. *Adv Chem Phys* 34:245–309
- van Kampen NG (2004) *Stochastic processes in physics and chemistry*, 5th edn. Elsevier, Amsterdam
- Woolley TE, Baker RE, Gaffney EA, Maini PK (2011) Stochastic reaction and diffusion on growing domains: understanding the breakdown of robust pattern formation. *Phys Rev E* 84:046216