PATHWISE ERROR BOUNDS IN MULTISCALE VARIABLE SPLITTING METHODS FOR SPATIAL STOCHASTIC KINETICS

AUGUSTIN CHEVALLIER AND STEFAN ENGBLOM

Abstract. Stochastic computational models in the form of pure jump processes occur frequently in the description of chemical reactive processes, of ion channel dynamics, and of the spread of infections in populations. For spatially extended models, the computational complexity can be rather high such that approximate multiscale models are attractive alternatives. Within this framework some variables are described stochastically, while others are approximated with a macroscopic point value.

We devise theoretical tools for analyzing the pathwise multiscale convergence of this type of variable splitting methods, aiming specifically at spatially extended models. Notably, the conditions we develop guarantee well-posedness of the approximations without requiring explicit assumptions of a priori bounded solutions. We are also able to quantify the effect of the different sources of errors, namely the multiscale error and the splitting error, respectively, by developing suitable error bounds. Computational experiments on selected problems serve to illustrate our findings.

1. Introduction

Mesoscopic spatially extended stochastic models are in frequent use in many fields, with notable examples found in cell biology, neuroscience, and epidemiology. The traditional macroscopic description is a partial differential equation (PDE) governing the flow of concentration field variables in a generalized reaction-transport process. Whenever a certain concentration is small enough, discrete stochastic effects become more pronounced, thus invalidating the assumptions behind the macroscopic model. An alternative is then to turn to a mesoscopic stochastic model, a continuous-time Markov chain over a discrete state-space. This model often remains accurate at an acceptable computational complexity.

In the traditional non-spatial, or well-stirred setting, early work by Kurtz connected these two descriptions via limit theorems, showing essentially that continuous approximations emerge in the limit of large molecular numbers, sometimes referred to as the “thermodynamic limit”. Strong approximation theorems in the same setting were later also developed (for more of this, see the monograph [14] and the references therein).

Multiscale-, or hybrid descriptions, in which the two scales are blended has since attracted many researchers. The focus of the research tend to fall into one of
two categories; either “theoretical” and concerning error bounds and rate of convergence, or more “practical” by developing actual implementations and general software.

In the first category, tentative analysis of specific examples are found in [4], while [17, 19] are of more general character and based on averaging techniques, and conditional expectations, respectively. A related analysis in the sense of mean-square convergence for operator splitting techniques is found in [11]. In [18] the issue of a proper scaling is stressed and similar remarks are made in [15], where notably, a practical multiscale simulation algorithm is also devised.

Towards the more algorithmic side, an early suggestion for a hybrid method in [16] came to be followed up by several others [1, 23, 24]. Related multiscale algorithms based on quasi equilibrium assumptions are found in [7, 9], and the method in [13] relied on the macroscale description as a preconditioner to bring out parallelism.

With few exceptions [3, 25], the main body of work has been done in the well-stirred (or 0-dimensional) setting. Since the work [12] and the software described in [8], however, it is fairly well understood how spatial models are to be developed. Here the computational complexity is much higher such that multiscale methods appear as a very attractive alternative. This is the starting point for the present contribution.

The goal with the analysis of the paper is twofold. We will firstly deal with the multiscale analysis required for the splitting of the state variable into a stochastic and a deterministic part, respectively. Secondly, we will also deal with the numerical analysis relied upon when designing a basic but representative time-discretization of this approximating process.

The paper is organized as follows: below we first summarize the main results of the paper. In §2 we work through the description of mesoscopic reactive processes as continuous-time Markov chains with a focus on the spatial case. A substantial effort is made to avoid any possibly circular assumptions on the solution regularity, but rather to prove all results within a single coherent framework. The analysis of the multiscale approximation is found in §3, where error bounds for both the multiscale and the splitting errors are developed. Our approach is pathwise in the sense that the errors are measured in $L^2$ over a single probability space. Selected numerical examples are presented in §4, and a concluding discussion is offered in §5.

1.1. Summary of main results. A brief orientation of the technical results of the paper is as follows:

1. Theorem 2.4 proves a strong regularity result for the type of spatial reactive processes considered in the paper.
2. Theorem 2.5 proves the corresponding result in the setting of a multiscale framework. In particular, this reveals partial assumptions for when a multiscale description is meaningful.
3. Theorems 2.6 and 2.7 similarly develop regularity results for the multiscale and the split-step approximations, respectively.
4. Theorems 3.4 and 3.5 provide for a multiscale convergence theory when parts of the dynamics is approximated via deterministic terms.
5. Theorems 3.6 and 3.7 similarly provides for a convergence theory of split-step methods in a general multiscale setting.
In this list, items 1–3 prove well-posedness and stability for the various involved processes. Following the celebrated Lax principle, items 4–5 next prove convergence and error estimates by an investigation of the consistency in the different approximations.

2. Mesoscopic spatial stochastic kinetics

We devote this section to some technical developments; §§2.1–2.2 summarize reaction-transport type modeling over irregular lattices, and regularity results under suitable model assumptions are developed in §§2.3. The variable splitting setup to be studied is similarly detailed in §§2.4–2.5, where the corresponding regularity results are evaluated anew.

Throughout the paper we shall remain in the framework of continuous-time Markov processes on a discrete state-space, albeit with some special structure imposed from the spatial context. Assuming a process $X(t) \in \mathbb{Z}^D_+$ counting at time $t$ the number of entities in each of $D$ compartments, a set of $R$ state transitions $X \mapsto X - N_r$ is generally prescribed by

$$
\mathbb{P} \{ X(t + dt) = x - N_r \mid X(t) = x \} = w_r(x) \, dt + o(dt),
$$

for $r = 1 \ldots R$. To enforce a conservative chain which remains in $\mathbb{Z}^D_+$, we assume $w_r(x) = 0$ whenever $x - N_r \not\in \mathbb{Z}^D_+$.

2.1. Continuous-time Markov chains on irregular lattices. In the traditional well-stirred setting we have $D$ species interacting according to $R$ chemical reactions in some fixed volume $V_{tot}$. Given an initial state $X(0)$, the dynamics is then fully described by the stoichiometric matrix $N \in \mathbb{Z}^{D \times R}$, and $w(x) \equiv [w_1(x), \ldots, w_R(x)]^T$, the set of propensities. Assuming a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting $R$-dimensional Poisson processes, the state is evolved according to [14, Chap. 6.2]

$$
X_i(t) = X_i(0) - \sum_{r=1}^{R} N_{ri} \Pi_r \left( \int_0^t w_r(X(s)) \, ds \right),
$$

for species $i = 1 \ldots D$ and with standard unit-rate independent Poisson processes $\Pi_r$, $r = 1 \ldots R$.

If the assumption of a spatially uniform distribution no longer holds a notation for spatial dependency needs to enter. The given continuous volume $V_{tot}$ is discretized into $J$ smaller voxels $(V_j)_{j=1}^J$ and the state $X \in \mathbb{Z}_{+}^{D \times J}$, where $X_{ij}$ is the number of molecules of the $i$th species in the $j$th voxel. The assumption of global homogeneity is replaced with a local assumption about uniformity in each voxel such that the dynamics (2.2) may be used anew on a per-voxel basis. Adding suitable terms covering any specified transport process we get

$$
X_{ij}(t) = X_{ij}(0) - \sum_{r=1}^{R} N_{ri} \Pi_{rj} \left( \int_0^t w_{rj}(X_{-j}(s)) \, ds \right) \\
- \sum_{k=1}^{J} \Pi'_{ijk} \left( \int_0^t q_{ijk} X_{ij}(s) \, ds \right) \\
+ \sum_{k=1}^{J} \Pi'_{ikj} \left( \int_0^t q_{ikj} X_{ik}(s) \, ds \right),
$$

for $i = 1 \ldots D$ and with suitable notation for spatial dependency.
where \( q_{ijk} \) is the rate per unit of time for species \( i \) to move from the \( j \)th voxel to the \( k \)th.

An important consequence of the integral representation \((2.3)\) is Dynkin’s formula \([6, \text{Chap. 9.2.2}]\). For \( f : \mathbb{Z}^{\mathbb{R}^d} \to \mathbb{R} \) a suitable function,

\[
E \left[ f(X(t)) - f(X(0)) \right] = 
\]

\((2.4)\)

\[
\int_0^t \sum_{j=1}^J \sum_{r=1}^R w_{rj}(X_{rj}(s)) \left[ f(X(s) - N_r \cdot \mathbb{I}_r^T) - f(X(s)) \right] \, ds 
\]

\[
+ \int_0^t \sum_{j,k=1}^J \sum_{i=1}^D q_{ijk} X_{ijk}(s) \left[ f(X(s) - \mathbb{1}_i \cdot \mathbb{I}_j^T + \mathbb{1}_i \cdot \mathbb{I}_k^T) - f(X(s)) \right] \, ds 
\]

\[
+ \int_0^t \sum_{j,k=1}^J \sum_{i=1}^D q_{ikj} X_{ikj}(s) \left[ f(X(s) + \mathbb{1}_i \cdot \mathbb{I}_j^T - \mathbb{1}_i \cdot \mathbb{I}_k^T) - f(X(s)) \right] \, ds,
\]

expressed in terms of the stopped process \( X(t) = X(t \wedge \tau_P) \) for a stopping time \( \tau_P := \inf_{t \geq 0} \{ \|X(t)\| > P \} \) in some suitable norm, and \( P > 0 \) an arbitrary real number. In \((2.4)\), \( \mathbb{1}_j \) is an all-zero column vector of suitable height and with a single 1 at position \( j \).

### 2.2. Mesh regularity

The subdivision of the total volume \( V_{tot} \) into smaller voxels is in principle arbitrary. However, any meaningful analysis will clearly depend to some extent on the regularity of this discretization.

**Definition 2.1 (Mesh regularity parameters).** We consider a geometry in \( d \) dimensions and total volume \( V_{tot} \), discretized by any member in the set of meshes \( \mathcal{M} \). For any such mesh \( M \in \mathcal{M} \) consisting of voxel volumes \( (V_j)_{j=1}^J \) we assume that it holds that

\[
\begin{align*}
(2.5) & \quad m_V \bar{V}_M \leq V_j \leq M_V \bar{V}_M, \\
(2.6) & \quad m_h V_j^{1/d} \leq \text{diam}(V_j) \leq M_h V_j^{1/d}, \\
(2.7) & \quad |\{k; q_{ijk} \neq 0\}| \leq M_D,
\end{align*}
\]

for constants \( 0 < m_V \leq M_V, 0 < m_h \leq M_h, M_D \), and average voxel volume \( \bar{V}_M = J^{-1} \sum_{j=1}^J V_j \). Hence under this parametrization we may write

\[ \mathcal{M} = \mathcal{M}(m_V, M_V, m_h, M_h, M_D). \]

Informally, \((2.5)\) measures how far the meshes in \( \mathcal{M} \) are from being uniform, \((2.6)\) ensures that no single voxel collapses into a voxel in less than \( d \) dimensions, and \((2.7)\) that the connectivity of the mesh is bounded. In the present paper \((2.6)\) is not used explicitly; this assumption assures a connection to the macroscopic viewpoint in that a concentration variable may be meaningfully defined everywhere.

### 2.3. Solution regularity

We next ensure the well-posedness of \((2.3)\) by deriving some pathwise bounds on this process. To get some feeling for what is going on we first look briefly at the corresponding PDE-setting.

Assume for simplicity that the transport rates \( q_{ijk} \) have been chosen as a consistent discretization of the operator \( \sigma_i \Delta \) under homogeneous Neumann conditions at the mesh \( M \). Denoting a deterministic time-dependent concentration variable
by \( v_i = v_i(t, x) \) for \( x \in \mathbb{R}^d \) and \( i = 1 \ldots D \), a macroscopic reaction-diffusion PDE corresponding to (2.3) reads

\[
\frac{\partial v_i}{\partial t} = \sigma_i \Delta v_i - \sum_{r=1}^{R} \mathbb{N}_{r_i} u_r(v_i), \quad \text{in } V_{\text{tot}}, \quad \frac{\partial v_i}{\partial n} = 0, \quad \text{on } \partial V_{\text{tot}},
\]

for certain nonlinear rates \( u_r, r = 1 \ldots R \) to be prescribed below. Equipped with suitable initial data, (2.8) can be expected to be a well-posed initial-boundary value problem in \( L^\infty([0, T]) \times L^p(V_{\text{tot}}) \) for any \( p \geq 1 \).

For the stochastic case (2.3), and in the non-spatial setting, an analysis in the form of assumptions and various \emph{a priori} bounds has been developed previously [10]. We borrow many ideas from this work in what follows.

The propensities in (2.3) generally obey the \emph{density dependent} scaling such that \( w_{rj}(x) = V_j u_r(V_j^{-1} x) \) for some dimensionless function \( u_r \) [14, Chap. 11]. We further expect from a physically realistic model that the number of molecules in an isolated volume \( V_j \) can somehow be bounded \emph{a priori}. To this end we postulate the existence of a weighted norm

\[
\| x \|_w := w^T x, \quad x \in \mathbb{R}^D,
\]

normalized such that \( \min_i w_i = 1 \). Following [10] we formulate

**Assumption 2.2 (Reaction regularity).** For a mesh \( M \in \mathcal{M} \) consisting of voxel volumes \( (V_j)_{j=1}^J \) we assume the density dependent scaling,

\[
w_{rj}(x) = V_j u_r(V_j^{-1} x),
\]

where \( u \) is \emph{independent of the mesh} and further satisfies,

\[
-w^T N u(x) \leq A + \alpha \| x \|_w,
\]

\[
( -w^T N )^2 u(x) / 2 \leq B + \beta_1 \| x \|_w + \beta_2 \| x \|_w^2,
\]

\[
u_r(x) - u_r(y) \leq L_r(P) \| x - y \|, \quad \text{for } r = 1 \ldots R, \quad \text{and } \| x \|_w \vee \| y \|_w \leq P.
\]

With the exception of \( \alpha \), all parameters \( \{ A, B, \beta_1, \beta_2, L \} \) are assumed to be non-negative.

When considering spatially varying solutions, the natural analogue to (2.9) is

\[
\| X \|_{w, 1} = \sum_{j=1}^J \| X_{\cdot j} \|_w = w^T 1,
\]

for \( 1 \) an all-unit column vector of suitable height. Our starting point is Dynkin’s formula (2.4). We find

\[
\mathbb{E} \left[ \| X(t) \|_{w, 1}^p \right] = \mathbb{E} \left[ \| X(0) \|_{w, 1}^p \right] + \mathbb{E} \left[ \int_0^t F(X(s)) \, ds \right]
\]

where

\[
F(X) = \sum_{j=1}^J \sum_{r=1}^R w_{rj} (X_{\cdot j}(s)) \left( \| X(s) \|_{w, 1} - w^T N_r \right)^p - \| X(s) \|_{w, 1}^p.
\]

We quote the following convenient inequality.
Lemma 2.1 (Lemma 4.6 in [10]). Let \( f(x) \equiv (x + y)^p - xp \) with \( x \in \mathbb{R}_+ \) and \( y \in \mathbb{R} \). Then for integer \( p \geq 1 \) we have the bounds
\[
\begin{align*}
(2.16) & \quad f(x) \leq p(x^{p-1} + 2p-4(p-1)y^2) \left[ x^{p-2} + |y|^{p-2} \right], \\
(2.17) & \quad |f(x)| \leq p|y|^{p-2} \left[ x^{p-1} + |y|^{p-1} \right].
\end{align*}
\]

Using Lemma 2.1 (2.16), Assumption 2.2 (2.10)–(2.12), and Definition 2.1 (2.5) we obtain, where for brevity \( x \equiv \|X\|_{\mathcal{W},1} \),
\[
(2.18) \quad F(X) \leq p(AV_{tot} + \alpha x)x^{p-1} + C_p(BV_{tot} + \beta_1 x + \beta_2 x^2)(x^{p-2} + C_{p-2}),
\]
where \( C_p := 2p^{-3}p(p-1), \beta_2 := \beta_2 m_{V}^{-1} \tilde{V}_M^{-1} \), and \( C_N := \|w^T N\|_{\infty} \). Combining (2.15) and (2.18) and using Young’s inequality several times we may obtain a bound of the form
\[
(2.19) \quad \mathbb{E} \left[ \|X(\hat{t})\|_{\mathcal{W},1}^p \right] \leq \mathbb{E} \left[ \|X(0)\|_{\mathcal{W},1}^p \right] + \mathbb{E} \left[ \int_0^{\hat{t}} C(1 + \|X(\hat{\tau})\|_{\mathcal{W},1}^p) \, ds \right],
\]
for some \( C > 0 \). Using Gronwall’s inequality and letting \( P \to \infty \) we arrive at

**Theorem 2.2.** Let \( X(t) \) obey (2.3) under Assumption 2.2. Then for any integer \( p \geq 1 \),
\[
(2.20) \quad \mathbb{E} \left[ \|X(t)\|_{\mathcal{W},1}^p \right] \leq \left( \mathbb{E} \left[ \|X(0)\|_{\mathcal{W},1}^p \right] + 1 \right) \exp(\lambda t) - 1,
\]
where the constant \( C > 0 \) depends on \( p \) and on the constants in the assumptions.

**Proof.** It remains to prove that \( \hat{t} \to t \) almost surely as \( P \to \infty \). Suppose to the contrary that \( \hat{t} = \tau_P \land t \) does not converge a.s. to \( t \) as \( P \to \infty \). Define \( A \equiv \{ \omega : \forall P : \tau_P(\omega) < t \} \). By the assumption \( P(A) > 0 \) and, for any \( \omega \in A \), and for all \( P > 0 \),
\[
\sup_{0 \leq s \leq t} \|X_s(\omega)\| > P, \text{ or simply, } \sup_{0 \leq s \leq t} \|X_s(\omega)\| = \infty.
\]
In other words, \( X(\hat{t}, \omega) \to \infty \) for every \( \omega \in A \), and \( \|X(\hat{t}, \omega)\| \) forms an increasing sequence with respect to \( P \). Using the Lebesgue monotone convergence theorem together with \( P(A) > 0 \), we get that \( \mathbb{E}[\|X(\hat{t})\|] \geq \mathbb{E}[\|X(\hat{t})\|1_{\omega \in A}] \to \infty \). However, \( \mathbb{E}[\|X(\hat{t})\|] \) is bounded from above independently of \( P \) and thus we have a contradiction. \( \square \)

Notably, when small voxels \( V_j \) are present and quadratic reactions which are not \( w \)-neutral are allowed (i.e. \( \beta_2 \neq 0 \)), then an investigation of \( C \) in (2.19) reveals that the second order moment and higher may grow fast as \( \exp(\beta_2 V_j^{-1} t) \).

To achieve pathwise convergence results we will need a stronger regularity guarantee which requires control of the martingale part via Burkholder’s inequality. To this end we define the quadratic variation of a real-valued process \( \{Y_t\}_{t \geq 0} \) by
\[
(2.21) \quad [Y]_t = \lim_{\|P\| \to 0} \sum_{k=0}^{n-1} (Y_{t_{k+1}} - Y_{t_k})^2,
\]
where the partition \( P = \{0 = t_0 < t_1 < \cdots < t_n = t\} \) for which \( \|P\| := \max_k |t_{k+1} - t_k| \) and where the limit is in probability.
Lemma 2.3. Let $X(t)$ satisfy (2.3) under Assumption 2.2. Then the quadratic variation of $\|X(t)\|_{w,1}^p$ is bounded by

\begin{equation}
E \left( \|X\|_{w,1}^{1/2} \right)^2 \leq E \left[ \int_0^t C(1 + \|X(s)\|_{w,1}^p + \beta_2 V(\|X(s)\|_{w,1}^{p+1})) \, ds \right],
\end{equation}

where $C > 0$ again depends on $p$ and on the constants in Assumption 2.2, but not on the mesh resolution, and where $\beta_2 V := \beta_2 m^{-1} V^{-1}$. 

Proof. Let $t_0 = 0$ and $t_i$ for $i = 1, 2, \ldots$ be the successive jump times of $X$. Then

\[ \|X(t_i)\|_{w,1}^p \leq \left( \sum_{0 < s_i \leq t_i} \left( \|X(s_i)\|_{w,1}^p - \|X(s_{i-1})\|_{w,1}^p \right)^2 \right)^{1/2}. \]

Under the stopping time $X$ is non-explosive with probability 1 and the number of jumps is finite in $[0, t]$. Thus we can use the inequality $\| \cdot \|_2 \leq \| \cdot \|_1$ to get

\[ \|X(t_i)\|_{w,1}^p \leq \sum_{0 < s_i \leq t_i} \left( \|X(s_i)\|_{w,1}^p - \|X(s_{i-1})\|_{w,1}^p \right). \]

The right-hand side can be written as a Lebesgue-Stieltjes integral,

\[ \int_0^t \sum_{j=1}^J \sum_{r=1}^{R_j} \left( \|X(s)\|_{w,1}^p - \|X(s)\|_{w,1}^p \right) \, dY_{rj}(s), \]

with $Y_{rj}$ the counting process $Y_{rj}(t) = \Pi_{rj} \left( \int_0^t w_{rj}(X_j(s)) \, ds \right)$. Taking the expectation yields

\[ E \left( \|X\|_{w,1}^{1/2} \right)^2 \leq E \left[ \int_0^t \sum_{j=1}^J \sum_{r=1}^{R_j} \left( \|X(s)\|_{w,1}^p - \|X(s)\|_{w,1}^p \right) \, ds \right]. \]

Using Lemma 2.1 (2.17) and Assumption 2.2 (2.10) and (2.12),

\[ \leq E \left[ \int_0^t \sum_{j=1}^J \sum_{r=1}^{R_j} p \|w_{rj} \|_{w} \|X(s)\|_{w,1} \left( \|X(s)\|_{w,1}^p + \|w_{rj} \|_{w} \right)^{2p-2} \right] \, ds \]

\[ \leq E \left[ \int_0^t C_p (BV_{tot} + \beta_1 \|X(s)\|_{w,1} + \beta_2 V(\|X(s)\|_{w,1}^{p+1} + C_{loc}^{p-1}) \right) \, ds \].

Relying on the moment bound in Theorem 2.2 we let $P \to \infty$ to arrive at the stated bound. \hfill \Box

We consider the following strong sense of pathwise locally bounded processes:

\begin{equation}
S_{p, loc}^p (Z_{+}^{D \times J}) = \left\{ X(t, \omega) : X(t) \in Z_{+}^{D \times J} \text{ is } F_t \text{-adapted such that } \sup_{t \in [0, T]} \|X(t)\|_{w,1}^p < \infty \text{ for } \forall T < \infty \right\}.
\end{equation}

Theorem 2.4 (Regularity). Let $X(t)$ be a solution to (2.3) under Assumption 2.2 with $\beta_2 = 0$. Then if $E[\|X(0)\|_{w,1}^p] < \infty$, \{ $X(t)$ \}_{t \geq 0} \in S_{p, loc}^p (Z_{+}^{D \times J})$. If $\beta_2 > 0$ then the conclusion remains under the additional requirement that $E[\|X(0)\|_{w,1}^{p+1}] < \infty$. 

Proof: This result follows as a combination of Theorem 2.2 and Lemma 2.3. We find that
\[
\|X(\hat{t})\|_{w,1}^p = \|X(0)\|_{w,1}^p + \int_0^{\hat{t}} F(X(s)) \, ds + M_{\hat{t}},
\]
with \( F \) defined in (2.15). The quadratic variation of the local martingale \( M_{\hat{t}} \) can be estimated via Lemma 2.3,
\[
(2.24) \quad \mathbb{E} \left( [M]^{1/2}_{\hat{t}} \right) \leq \mathbb{E} \left[ \int_0^{\hat{t}} C \left( 1 + \|X(s)\|_{w,1}^p + \beta_2 Y_s \|X(s)\|_{w,1}^{p+1} \right) ds \right].
\]
Assume first that \( \beta_2 = 0 \). Using the previously developed bound in (2.18) and (2.19) for the drift part we get
\[
\|X(\hat{t})\|_{w,1}^p \leq \|X(0)\|_{w,1}^p + \int_0^{\hat{t}} C \left( 1 + \|X(s)\|_{w,1}^p \right) ds + |M_{\hat{t}}|.
\]
Combining with (2.24) we find after using Burkholder’s inequality [20, Chap. IV.4],
\[
\mathbb{E} \left[ \sup_{s \in [0,\hat{t}]} \|X(s)\|_{w,1}^p \right] \leq \mathbb{E}[\|X(0)\|_{w,1}^p] + \int_0^{\hat{t}} C \left( 1 + \mathbb{E} \left[ \sup_{s' \in [0,s]} \|X(s')\|_{w,1}^p \right] \right) ds.
\]
For clarity, writing \( \|X\|_{w,1}^p (t) := \sup_{s \in [0,t]} \|X(s)\|_{w,1}^p \) we find that
\[
\mathbb{E}[\|X\|_{w,1}^p (\hat{t})] \leq \mathbb{E}[\|X(0)\|_{w,1}^p] + \int_0^{\hat{t}} C \left( 1 + \mathbb{E}[\|X\|_{w,1}^p (\bar{s})] \right) ds.
\]
Gronwall’s inequality now implies that \( \mathbb{E}[\|X\|_{w,1}^p (\hat{t})] \) is bounded in terms of the initial data and time \( t \). By Fatou’s lemma the claim follows by letting \( P \to \infty \).

We next consider \( \beta_2 > 0 \). Using Theorem 2.2 we still have the bound (2.24) which yields
\[
\mathbb{E} \left( [M]^{1/2}_{\hat{t}} \right) \leq \int_0^{\hat{t}} C \left( 1 + \mathbb{E}[\|X(s)\|_{w,1}^{p+1}] \right) ds \leq (e^{C\hat{t}} - 1)(\mathbb{E}[\|X(0)\|_{w,1}^{p+1}] + 1),
\]
where we similarly obtain a bound in terms of \( \mathbb{E}[\|X(0)\|_{w,1}^{p+1}] \). \( \square \)

2.4. Scaling. We shall now regard the transport rates, the reaction rates, and the magnitude of the state variables as problem parameters which may induce a scale separation. Although a completely general multiscale analysis is possible within the current framework, to fix our ideas and in the interest of a transparent presentation, we consider a concrete, but still quite general two-scale separation.

Condition 2.3 (Scale separation). Let a scale vector \( S \in \mathbb{R}^D \) be given. The transport- and reaction rates are assumed to obey the scaling laws
\[
q_{ijk} x = \epsilon^{-\mu_i} \tilde{q}_{ijk} S^{-1} x,
\]
\[
u_x (x) = \epsilon^{-\nu_{x}^i} \tilde{\nu}_x (x) = \epsilon^{-\nu_{x}^2} \tilde{\nu}_x (S^{-1} x)
\]
for \( i = 1 \ldots D, (j,k) = 1 \ldots J, \) and \( r = 1 \ldots R. \) For the state variables we define
\[
X_{i,} (t) = S_i \tilde{X}_{i,} (t), \quad S_i = 1 or \epsilon^{-1}.
\]
where \( \nu_r^{(1)} \) is the scaling of the rate (fast/slow) while \( \nu_r^{(2)} \) follows from the number of species involved in transition \( r \) such that \( S_i = \epsilon^{-1} \). Let the complete scaling be

\[
\nu_r = \nu_r^{(1)} + \nu_r^{(2)}.
\]

The dynamics is considered for \( t \in [0, T] \), \( T = O(1) \) with respect to \( \epsilon \). Also, all non-dimensionalized constants and propensities \( \{ \tilde{q}_{ijk}, \tilde{u}_r(\cdot) \} \) are understood to be \( O(1) \) with respect to \( \epsilon \).

It is possible to analyze also the general case where the species scale differently in different voxels, i.e. \( X_{ij} = S_{ij} \tilde{X}_{ij} \). However, this analysis is complicated by the fact that the results then take place in a transient regime, and, in turn, this regime is difficult to generally estimate.

We make a slight abuse of notation by employing \( S \) as if it was the \( D \times D \) matrix \( \text{diag}(S) \). Using a similar convention for \( \nu \) we may write (2.26) in the compact form

\[
(2.28) \quad u(x) = e^{-\epsilon \tilde{u}}(S^{-1}x).
\]

To take a concrete example: the bimolecular reaction \( X + Y \rightarrow \emptyset \) at rate \( kX_2Y \) obeys (2.26) with \( \nu_r = 0 \) for \( k \sim \epsilon \) and one of the species scaling macroscopically as \( \epsilon^{-1} \). If both species are macroscopic, then instead \( \nu_r = 1 \) at the same scaling of the rate \( k \sim \epsilon \).

Following Condition 2.3 we thus divide the species into two disjoint groups, \( G_1 \) and \( G_2 \), with \( |G_1| + |G_2| = D_1 + D_2 = D \). Informally, we suppose that species in low copy numbers are in \( G_1 \) and species in large copy numbers are in \( G_2 \). Under an appropriate enumeration of the species this implies the choice of scaling \( S_i = 1 \) for \( i \in G_1 = \{1 \ldots D_1\} \) and \( = \epsilon^{-1} \) for \( i \in G_2 = \{D_1 + 1 \ldots D\} \) in (2.27). Following this ordering we also write \( w = [w_1; w_2] \) and \( N = [N^{(1)}; N^{(2)}] \), where \( w_i \in \mathbb{R}^{D_i}_{\geq 1} \) and \( N^{(i)} \in \mathbb{R}^{D_i \times R} \) for \( i \in \{1, 2\} \).

We find from (2.3) the governing equation

\[
(2.29) \quad \tilde{X}_{ij}(t) = \tilde{X}_{ij}(0) - \sum_{r=1}^{R} S_{ij}^{-1} N_{r} \Pi_{r} \left( \int_{0}^{t} V_{j} e^{-\nu_r \tilde{u}_r(V_{j}^{-1} \tilde{X}_{ij}(s))} ds \right)
\]

\[
- \sum_{k=1}^{J} S_{ij}^{-1} \Pi_{jk} \left( \int_{0}^{t} e^{-\mu_i \tilde{q}_{ijk} \tilde{X}_{ij}(s)} ds \right)
\]

\[
+ \sum_{k=1}^{J} S_{ik}^{-1} \Pi_{ik} \left( \int_{0}^{t} e^{-\mu_i \tilde{q}_{ik} \tilde{X}_{ik}(s)} ds \right).
\]

For the existence of scale separation it is critical to find conditions such that according to some weight-vector \( l \), \( \| \tilde{X}(t) \|_{L_1} \) for \( t \in [0, T] \) remains \( O(1) \) whenever \( \| \tilde{X}(0) \|_{L_1} \) is \( O(1) \), assuming that \( T \) and \( l \) both are \( O(1) \) with respect to \( \epsilon \). Unfortunately, the assumptions and analysis in §2.3 all concerned the unscaled variable \( X(t) \), which is now assumed to be \( O(\epsilon^{-1}) \). In fact, it is not difficult to see that with, say, \( l := Sw \) replacing \( w \) throughout Assumption 2.2, and requiring that all constants be independent of \( \epsilon \), the results in §2.3 are straightforwardly translated into bounds in terms of the \( L \)-norm of \( \tilde{X}(t) \). Since this is just the \( w \)-norm of \( X(t) \) itself, however, it scales as \( O(\epsilon^{-1}) \). What is additionally required is that the weight-vector \( l \) can be selected independently of \( \epsilon \).
**Assumption 2.4** (Reaction regularity, scaled case). The previous assumption of density dependent propensities (2.10) is assumed to hold. We further assume the existence of a vector \( l \in \mathbb{R}_{\geq 1}^D \), independent of \( \epsilon \), such that

\[
(2.30) \quad -l^T S^{-1} N u(x) \leq A + \alpha \| S^{-1} x \|_l,
\]

\[
(2.31) \quad (-l^T S^{-1} N)^2 u(x)/2 \leq B + \beta_1 \| S^{-1} x \|_l^2 + \beta_2 \| S^{-1} x \|_l^2,
\]

\[
(2.32) \quad |\bar{u}_r(x) - \bar{u}_r(y)| \leq \bar{L}_r(\bar{P}) \| x - y \|, \text{ for } r = 1 \ldots R, \text{ and } \| x \|_l \vee \| y \|_l \leq \bar{P}.
\]

All parameters \( \{A, \alpha, B, \beta_1, \beta_2, L\} \) are assumed to be independent of \( \epsilon \) and non-negative (with negative values allowed for \( \alpha \)).

Equipped with this assumption we revisit the regularity results of §2.3. To this end we consider a version of \( S_F^{p, \text{loc}}(Z^{D \times J}_+) \) scaled with \( S \),

\[
(2.33) \quad S_F^{p, \text{loc}}(S^{-1} Z^{D \times J}_+) = \left\{ \bar{X}(t, \omega): \bar{X}(t) \in S^{-1} Z^{D \times J}_+ \text{ is } F_t\text{-adapted such that } E[\sup_{t \in [0, T]} \| \bar{X} \|_{l, 1}] < \infty \text{ for all } T < \infty \right\},
\]

where the scaled state space is just

\[
(2.34) \quad S^{-1} Z^{D \times J}_+ = [Z^{D_1 \times J}_+; \epsilon Z^{D_2 \times J}_+],
\]

and \( \epsilon Z_+ = \{0, \epsilon, 2\epsilon, \ldots\} \).

**Theorem 2.5** (Regularity, scaled case). Under Condition 2.3, Theorem 2.2 and 2.4 both hold with the new Assumption 2.4 replacing the previous Assumption 2.2 and with \( S_F^{p, \text{loc}}(S^{-1} Z^{D \times J}_+) \) replacing \( S_F^{p, \text{loc}}(Z^{D \times J}_+) \). In particular:

1. The constant \( C \) in Theorem 2.2 can be selected independently of \( \epsilon \).
2. If either \( \beta_2 = 0 \) and \( E[\| \bar{X}(0) \|_{l, 1}^p] \) is \( O(1) \) with respect to \( \epsilon \), or \( \beta_2 > 0 \) and \( E[\| \bar{X}(0) \|_{l, 1}^p] \) is \( O(1) \), then so is \( E[\sup_{t \in [0, T]} \| \bar{X}(s) \|_{l, 1}^p] \) for \( t \in [0, T] \), \( T = O(1) \) with respect to \( \epsilon \).

The proof follows very closely the steps taken to arrive at Theorem 2.4 and is therefore omitted. Theorem 2.5 inherits from Theorem 2.4 the poorer regularity when \( \beta_2 > 0 \). The predicted growth is then \( \exp(t \beta^V_2) \) where, as in (2.15), \( \beta^V_2 := \beta_2 m^{-1} V^{-1}_M \) and is dependent on the mesh.

**2.5. Multiscale splittings.** We shall consider two multiscale splittings: one “exact” in continuous time and one “numerical” in discrete time-steps of length \( h \).

Thus we firstly define \( \bar{Z} \) for \( i \in G_1 \) and using that \( S_i = 1 \),

\[
(2.35) \quad \bar{Z}_{ij}(t) = \bar{Z}_{ij}(0) - \sum_{r=1}^R \Pi_{r,j} \left( \int_0^t V_j e^{-\nu r} \bar{a}_r(V^{-1}_j \bar{Z}_{..,i}(s)) \, ds \right)
- \sum_{k=1}^J \Pi'_{r,kj} \left( \int_0^t e^{-\mu k \bar{q}_{ijk}} \bar{Z}_{ij}(s) \, ds \right)
+ \sum_{k=1}^J \Pi''_{ikj} \left( \int_0^t e^{-\mu k \bar{q}_{ikj}} \bar{Z}_{ik}(s) \, ds \right),
\]
while for $i$ in $G_2$, $S_i = \varepsilon^{-1}$ and the Poisson process is approximated by a deterministic process,

\[
(2.36) \quad \tilde{Z}_{ij}(t) = \tilde{Z}_{ij}(0) - \sum_{r=1}^{R} \varepsilon N_{ir} \left( \int_{0}^{t} V_j \varepsilon^{-\mu r} \tilde{u}_r (V_j^{-1} \tilde{Z}_{.,j}(s)) \, ds \right) \\
- \sum_{k=1}^{J} \varepsilon \left( \int_{0}^{t} \varepsilon^{-\mu k} \tilde{q}_{ijk} \tilde{Z}_{ij}(s) \, ds \right) \\
+ \sum_{k=1}^{J} \varepsilon \left( \int_{0}^{t} \varepsilon^{-\mu k} \tilde{q}_{ik} \tilde{Z}_{ik}(s) \, ds \right).
\]

In general, there is no guarantee that $\tilde{Z}(t)$ remains positive even when $\tilde{X}(t)$ is a conservative chain. For example, the presence of a dimerization reaction, say, $A + A \rightarrow B$ at rate $A(A - 1)$ can reach negative values of $B$ when $A$ is approximated by a continuous variable. In this example one can avoid this problem by reinterpretating the rate as $A(A - 1) \vee 0$. In what follows we will for simplicity assume that all models are conservative and remain in the non-negative orthant, presumably after employing some kind of limiters on the rates.

To see how a result similar to Theorem 2.5 might be obtained for the new process $\tilde{Z}(t)$, we start anew from Dynkin’s formula, appropriately modified for the semi-continuous setting. We find

\[
(2.37) \quad E \left[ \| \tilde{Z}(t) \|_{L^1}^p \right] = E \left[ \| \tilde{Z}(0) \|_{L^1}^p \right] + E \left[ \int_{0}^{t} G(\tilde{Z}(s)) \, ds \right],
\]

where

\[
(2.38) \quad G(Z) = \sum_{j=1}^{J} \sum_{r=1}^{R} V_j u_r (V_j^{-1} S Z_{.,j}) \left( \left( z - t_1 \nu N_r^{(1)} \right)^p - z^p - p t_2 \nu N_r^{(2)} z^{p-1} \right);
\]

and where for brevity $z \equiv \| Z \|_{L^1}$ (compare (2.15)). Using Lemma 2.1 (2.16) we find

\[
(2.39) \quad G(Z) \leq \sum_{j=1}^{J} \sum_{r=1}^{R} V_j u_r (V_j^{-1} S Z_{.,j}) \left[ -p t_1 S^{-1} N_r z^{p-1} + C_p / 2 \left( -t_1 \nu N_r^{(1)} \right)^2 \right] z^{p-2} + C_p^{p-2} \nu N_r^{(1)};
\]

where $C_p$ and $C_{N_r^{(1)}}$ are defined below (2.18). The goal here is to obtain a bound $G(Z) \leq C(1 + z^p)$ (compare (2.18)–(2.19)) and it is not difficult to see what assumption is required.

**Assumption 2.5 (Reaction regularity, semi-continuous case).** In Assumption 2.4, replace (2.31) with

\[
(2.40) \quad \left( -t_1 \nu N_r^{(1)} \right)^2 u(x)/2 \leq B + \beta_1 \| S^{-1} x \|_{L^1} + \beta_2 \| S^{-1} x \|_{L^1}^2.
\]

This assumption can be understood as firstly, a signed bound (2.30) on the drift-part for the fully coupled system, and secondly, the extra assumption due to stochasticity (2.40), which here applies only to $i \in G_1$, that is, to the stochastic part.
Using this in (2.39) we find (compare (2.18))

\[(2.41) \quad G(Z) \leq p(AV_{tot} + \alpha z)z^{p-1} + C_p(BV_{tot} + \beta_1 z + \beta_2 z^2)(z^{n-2} + C^{p-2}_{n(1)}),\]

and following the steps in the proofs of Theorems 2.4 and 2.5 we obtain after some work the following result.

**Theorem 2.6 (Regularity, semi-continuous case).** The statement of Theorem 2.5 applies also to the approximating process $\bar{Z}(t)$ with Assumption 2.5 taking the role of Assumption 2.4. The existence of solutions then concerns the semi-continuous space $S^{p, loc}_{2\nu}(\mathbb{Z}_+^{D_1 \times J}; \mathbb{R}_+^{D_2 \times J})$ and we note that the remark following Theorem 2.5 concerning the dependence on the mesh regularity remains valid since $\beta_2^{p'}$ is present in (2.41).

In practice, a numerical method is required to simulate $\bar{Z}$. The most straightforward way is to evolve the stochastic and deterministic parts in different steps, introducing a new process $\bar{Y}^{(h)}$ which approximates $\bar{Z}$. Following the partition of unity idea in [11] we define the kernel step function

\[(2.42) \quad \sigma_h(t) = 1 - 2(|t/h/2| \mod 2).\]

Then for $i$ in $G_1$,

\[(2.43) \quad \bar{Y}^{(h)}_{ij}(t) = \bar{Y}^{(h)}_{ij}(0) - \sum_{r=1}^{R} N_{rl} \Pi_{rj} \left( \int_0^t (1 + \sigma_h(s)) V_j e^{-\nu r} \bar{u}_r(V_j^{-1} \bar{Y}^{(h)}_{ij}(s)) ds \right)
\]

\[- \sum_{k=1}^{J} \Pi'_{ik} \left( \int_0^t (1 + \sigma_h(s)) e^{-\mu \tilde{q}_{ijk}} \bar{Y}^{(h)}_{ij}(s) ds \right)
\]

\[+ \sum_{k=1}^{J} \Pi_{ik} \left( \int_0^t (1 + \sigma_h(s)) e^{-\mu \tilde{q}_{ikj}} \bar{Y}^{(h)}_{ik}(s) ds \right),\]

and for $i$ in $G_2$,

\[(2.44) \quad \bar{Y}^{(h)}_{ij}(t) = \bar{Y}^{(h)}_{ij}(0) - \sum_{r=1}^{R} e N_{r1} \left( \int_0^t (1 - \sigma_h(s)) V_j e^{-\nu r} \bar{u}_r(V_j^{-1} \bar{Y}^{(h)}_{ij}(s)) ds \right)
\]

\[- \sum_{k=1}^{J} \epsilon \left( \int_0^t (1 - \sigma_h(s)) e^{-\mu \tilde{q}_{ijk}} \bar{Y}^{(h)}_{ij}(s) ds \right)
\]

\[+ \sum_{k=1}^{J} \epsilon \left( \int_0^t (1 - \sigma_h(s)) e^{-\mu \tilde{q}_{ikj}} \bar{Y}^{(h)}_{ik}(s) ds \right),\]

For regularity we start anew from the semi-continuous Dynkin’s formula,

\[(2.45) \quad \mathbb{E} \left[ \left\| \bar{Y}^{(h)}(\tilde{t}) \right\|_{l,1}^p \right] = \mathbb{E} \left[ \left\| \bar{Y}^{(h)}(0) \right\|_{l,1}^p \right] + \mathbb{E} \left[ \int_0^{\tilde{t}} H(\bar{Y}^{(h)}(s), s) ds \right],\]
where this time

\[
H(Y, s) \equiv \sum_{j=1}^{J} \sum_{r=1}^{R} (1 + \sigma_h(s))V_j u_r (V_j^{-1} SY, j) \left[ \left( y - I_1^T N_{1r}^{(1)} \right)^p - y^p \right]
\]

\[
+ \sum_{j=1}^{J} \sum_{r=1}^{R} (1 - \sigma_h(s))V_j u_r (V_j^{-1} SY, j) \left[ -p L_2^T e N_{1r}^{(2)} y^{p-1} \right],
\]

(2.46)

and where as before \( y \equiv \|Y\|_{t,1} \). This leads us to

**Assumption 2.6** *(Reaction regularity, split-step case)*. Besides the modification of Assumption 2.5, in Assumption 2.4, additionally replace (2.30) with

\[
(2.47) \quad \max \left( -I_1^T N^{(1)} u(x), -I_2^T e N^{(2)} u(x) \right) \leq A + \alpha \|S^{-1} x\|_1.
\]

In other words, (2.47) bounds the drift of the stochastic and continuous parts individually, while as before (2.40) is employed to bound the quadratic variation of the stochastic part alone.

Following again the steps in the previous proofs we obtain

**Theorem 2.7** *(Regularity, split-step case)*. Theorem 2.6 applies also to the approximating process \( \bar{Y}^{(h)}(t) \) under Assumption 2.6. The resulting a priori bound is uniform with respect to both \( \epsilon \) and \( h \) provided the initial data is.

The approximation \( \bar{X} \approx \bar{Z} \) gives rise to a *multiscale error*, whereas \( \bar{Z} \approx \bar{Y}^{(h)} \) induces a *splitting error*. Quite generally, any practical numerical method relies on this very structure in \( \bar{X} \approx \bar{Y}^{(h)} \). Insight into the nature of the total error thus follows from a consistent analysis of both approximations. This is the purpose with the next section.

### 3. Error analysis

We present in this section the error analysis of the two approximations (2.35)–(2.36) and, respectively, (2.43)–(2.44). Theorems 2.5, 2.6, and 2.7 assert that all processes are uniformly stable in finite time. By the Lax principle the task has therefore been reduced to an investigation of the degree of consistency of the two approximations. Preliminary lemmas for this are discussed in §3.1, followed by the actual error analysis in §3.2–3.3. In order not to lose the oversight, some material heavily relied upon are developed separately in Appendix A and B.

#### 3.1. Preliminary estimates

Intuitively, the same version of a Poisson process evaluated at two different operational times should enjoy a bounded difference, provided of course the times themselves are bounded in some suitable sense. A precise formulation of this property is related to *Doob’s optional sampling theorem* [20, Theorem 17, Chap. I.2] and has only just recently been investigated [2, 15] for the \( L^1 \)-norm, and in [11] for the \( L^2 \)-norm.

**Lemma 3.1.** Let \( \Pi \) be a unit-rate \( \mathcal{F}_t \)-adapted Poisson process, and let \( T \) be a bounded stopping time. Then

\[
E[\Pi(T)] = E[T],
\]

(3.1)

\[
E[\Pi^2(T)] = 2 E[\Pi(T)T] - E[T^2] + E[T].
\]

(3.2)
Proof. Let $\hat{\Pi}(t) := \Pi(t) - t$ be the compensated process. This is a martingale and the sampling theorem implies $E[\hat{\Pi}(T)] = 0$, which is (3.1). The quadratic variation is $[\hat{\Pi}]_t = \Pi(t)$ and hence $Z(t) := \hat{\Pi}^2(t) - \Pi(t)$ is a local martingale. Since $E[\sup_{s \leq t} Z(s)] < \infty$ for bounded $t$, it is actually a martingale and the sampling theorem now yields $E[Z(T)] = 0$, or,

$$0 = E[\Pi^2(T) - 2\Pi(T)T + T^2 - \Pi(T)],$$

which is (3.2).

□

Lemma 3.2. Let $\Pi$ be a unit-rate $\mathcal{F}_t$-adapted Poisson process, and let $T_1, T_2$ be bounded stopping times. Then

$$E[|\Pi(T_2) - \Pi(T_1)|] = E[|T_2 - T_1|], 
\tag{3.3}
$$

$$E[(\Pi(T_2) - \Pi(T_1))^2] = 2E[|\Pi(T_2) - \Pi(T_1)|(T_1 \lor T_2)] 
- E[|T_2^2 - T_1^2|] + E[|T_2 - T_1|].
\tag{3.4}
$$

Proof. Assume first that $T_2 \geq T_1$. We get from Lemma 3.1 (3.1)

$$E[|\Pi(T_2) - \Pi(T_1)|] = E[|T_2 - T_1|].
$$

For general stopping times $S_1, S_2$, say, not necessarily satisfying $S_2 \geq S_1$, (3.3) now follows upon substituting $T_1 := S_1 \wedge S_2$ and $T_2 := S_1 \lor S_2$ into this equality.

Next put $X := E[|\Pi(T_2) - \Pi(T_1)|^2]$ and assume again that $T_2 \geq T_1$. We get

$$X = E[|\Pi(T_2)|^2 + |\Pi(T_1)|^2 - 2|\Pi(T_1)||\Pi(T_2)|]
= E[|\Pi(T_2)|^2 + |\Pi(T_1)|^2] - 2E[|\Pi(T_1)||\Pi(T_2)|\mathcal{F}_{T_1}].
$$

To evaluate the iterated expectation note that

$$E[\hat{\Pi}(T_2)|\mathcal{F}_{T_1}] = \hat{\Pi}(T_1) \iff E[\Pi(T_2)|\mathcal{F}_{T_1}] = \Pi(T_1) - T_1 + E[T_2|\mathcal{F}_{T_1}].$$

Hence,

$$E[|\Pi(T_1)|E[|\Pi(T_2)|\mathcal{F}_{T_1}]] = E[|\Pi(T_1)|^2] - E[|\Pi(T_1)T_1|] + E[|\Pi(T_1)|T_2],$$

and we thus find that

$$X = E[|\Pi(T_2)|^2 + |\Pi(T_1)|^2] - 2E[|\Pi(T_1)||\Pi(T_2)|].$$

Applying Lemma 3.1 (3.2) twice yields finally

$$X = 2E[|\Pi(T_2) - \Pi(T_1)||T_2] - E[T_2^2 - T_1^2] + E[T_2 - T_1].$$

For general stopping times $S_1, S_2$, (3.4) now follows as before upon substituting $T_1 := S_1 \wedge S_2$ and $T_2 := S_1 \lor S_2$. □

Remark. We will use Lemma 3.2 in the following form. Assuming $T_1 \lor T_2$ has been bounded a priori by some value $B$ we get by combining (3.3) with (3.4) that

$$E[(\Pi(T_2) - \Pi(T_1))^2] \leq (2B^2 + 1)E[|T_2 - T_1|].
\tag{3.5}
$$

Let $\mathcal{F}_t$ be the filtration adapted to $\hat{\Pi}_r, r = 1 \ldots R$. Then for a fixed $t$, $T_r(t) = \int_0^t w_r(X(s))\, ds$ is a stopping time [2, Lemma 3.1] with respect to

$$\mathcal{F}_u = \sigma(\Pi_r(s), s \in [0, u]; \Pi_{k \neq r}(s), s \in [0, \infty])$$

for $1 \leq r \leq R$. Moreover, $T_r(t)$ is a martingale, i.e., $E[T_r(t)|\mathcal{F}_s] = T_r(t)$ for $s \leq t$.
Intuitively, as $X(t) = \sum_r \Pi_r(T_r(t)) N_r$, the event $\{T_r(t) < u\}$ depends on $\Pi_r$ during $[0,u]$ and on all other processes $\{\Pi_k, k \neq r\}$ during $[0,\infty)$. However, as $\Pi_r$, $r = 1\ldots R$ are independent, $\Pi_r(t) - t$ is still a martingale with respect to $\mathcal{F}_u^r$ (and not only with respect to $\mathcal{F}_u = \sigma(\Pi_r(s), s \in [0,u])$). Hence we can apply the stopping time theorems to $T_r(t)$ and the previous lemmas therefore apply. The result stays true for the approximating process $Z$ (and later $Y$). Hence, given the bound

$$
\int_0^t w_r(X(s)) \, ds \vee \int_0^t w_r(Y(s)) \, ds \leq B
$$

we get from (3.5) that

$$
\mathbb{E}\left[ \left( \Pi_r \left( \int_0^t w_r(X(s)) \, ds \right) - \Pi_r \left( \int_0^t w_r(Y(s)) \, ds \right) \right)^2 \right] \leq (2B + 1) \mathbb{E}\left[ \left| \int_0^t w_r(X(s)) \, ds - \int_0^t w_r(Y(s)) \, ds \right| \right].
$$

### 3.2. Multiscale convergence

This section develops a bound for the multiscale error made in the approximation $\hat{X} \approx \hat{Z}$. Throughout §2, a certain weighted norm which greatly simplified the theory was used. However, in the present case of bounding errors we are interested in the more conventional $L^2$-norm,

$$
\|\hat{X}(t)\|_2^2 = \sum_{j=1}^J \|\hat{X}(t)\|_j^2 = \sum_{j=1}^J \sum_{i=1}^D \hat{X}(t)_{ij}^2,
$$

where, for convenience, from now on we shall write $\|\cdot\|$ instead of $\|\cdot\|_2$.

Let $\bar{P} > 0$ and define the joint stopping time

$$
\tau_{\bar{P}} := \inf_s \{ \|\hat{X}(s)\| \vee \|\hat{Z}(s)\| \vee \|\hat{Y}^{(h)}(s)\| > \bar{P} \}, \text{ and put } \hat{t} := \tau_{\bar{P}} \land t.
$$

Recall the stopping time $T_r(t)$ from the remark after Lemma 3.2. Clearly, for any fixed $t$, $T_r(t)$ is still a stopping time.

The first step in the analysis is to split the error in one part which is bounded and one part which is not,

$$
\mathbb{E}\left[ \|\hat{X}(t) - \hat{Z}(t)\|^2 \right] = \mathbb{E}\left[ \|\hat{X}(t) - \hat{Z}(t)\|^2 1_{t > \hat{t}} \right] + \mathbb{E}\left[ \|\hat{X}(t) - \hat{Z}(t)\|^2 1_{t \leq \hat{t}} \right].
$$

The requirement to be able to control the contribution from the non-bounded part motivates the following lemma:

**Lemma 3.3.** For any $p > 1$, there exists a constant $K_p$ independent from $\epsilon$ and $h$ such that

$$
\mathbb{E}\left[ \|\hat{X}\|^2 1_{t > \hat{t}} \right] \leq K_p \bar{P}^{-p/2},
$$

$$
\mathbb{E}\left[ \|\hat{Z}\|^2 1_{t > \hat{t}} \right] \leq K_p \bar{P}^{-p/2},
$$

$$
\mathbb{E}\left[ \|\hat{Y}^{(h)}\|^2 1_{t > \hat{t}} \right] \leq K_p \bar{P}^{-p/2}.
$$

**Proof.** Theorem 2.5 yields

$$
\mathbb{E}\left[ \|\hat{X}(0)\|^4 1_{t \leq \hat{t}} \right] \leq \left( \mathbb{E}\left[ \|\hat{X}(0)\|^4 \right] + 1 \right) \exp(Ct) - 1.
$$

Since $\|\cdot\|$ and $\|\cdot\|_{L,1}$ are equivalent bounds we have an *a priori* bound

$$
\mathbb{E}\left[ \|\hat{X}(t)\|^4 \right] \leq B(t),
$$
with $B(t)$ independent from $\epsilon$. By Cauchy-Schwartz’s inequality,
\[ E[\|X(t)\|^2 1_{t > \tilde{t}}] \leq E[\|X(t)\|^4]^{1/2} P[t > \tilde{t}]^{1/2}. \]
Using that
\[ P[t > \tilde{t}] \leq P[\sup_{s \in [0, t]} \|X_s\| > \bar{P}] + P[\sup_{s \in [0, t]} \|Z_s\| > \bar{P}] + P[\sup_{s \in [0, t]} \|Y_s^{(h)}\| > \bar{P}], \]
we find from Markov’s inequality the bound
\[ P[t > \tilde{t}] \leq E \left[ \left( \sup_{s \in [0, t]} \|X(s)\| \right)^{P} \right] + E \left[ \left( \sup_{s \in [0, t]} \|Y^{(h)}(s)\| \right)^{P} \right]. \]
Using the second part of Theorem 2.5 and the equivalence of norms, it is possible to bound the first term on the right independently from $\epsilon$ and $h$. Reasoning similarly for the terms depending on $Y^{(h)}$ and $Z$ we get the stated result. \qed

To formulate the main result of this section we let
\begin{equation}
R(G_1) := \{ r; \exists i \in G_1 \text{ such that } N_{ri} \neq 0 \},
\end{equation}
and the analogous definition for $R(G_2)$. In words, $R(G_1)$ contains the reactions which affect any species $i \in G_1$. We additionally define the two effective exponents
\begin{align}
(3.14) & \quad u = \min_{r \in R(G_1)} -\nu_r + \min_{i \in G_1} -\mu_i, \\
(3.15) & \quad v = 1 + \min_{r \in R(G_2)} -\nu_r + \min_{i \in G_2} -\mu_i.
\end{align}
Note that, if the transport rates do not scale with $\epsilon$, we generally get $u \leq 0$ and $v \leq 1$.

**Theorem 3.4 (Multiscale error, bounded version).** Under the scale separation Condition 2.3, the regularity Assumptions 2.4 and 2.5, and assuming also that $Z$ and $\bar{X}$ are uniformly bounded with respect to $\epsilon$ by some $\bar{P}$, then whenever $u \geq 0$, $v \geq 0$ it holds that
\begin{equation}
E[\|Z(t) - \bar{X}(t)\|^2] = O(\epsilon^{1+v} + \epsilon^{1/2+u/2+v}).
\end{equation}
\begin{proof}
First notice that, since the processes are uniformly bounded with respect to $\epsilon$, so is $L_\epsilon$. Thus according to Lemma A.1,
\[ E[\|Z(t) - \bar{X}(t)\|^2] \leq C \int_0^t E[\|Z(s) - \bar{X}(s)\|^2] ds + C \int_0^t E[\|Z(s) - \bar{X}(s)\|] ds,
\]
with
\[ A = O(\epsilon^{1+v}), \quad B = O(\epsilon^{2v}), \quad C = O(\epsilon^v). \]
Similarly, according to Lemma A.2,
\[ E[\|\bar{X}(t) - \bar{Z}(t)\|] \leq C D + E \int_0^t E[\|\bar{X}(s) - \bar{Z}(s)\|] ds,
\]
where
\begin{equation}
D = O(\epsilon^{1/2+v/2}), \quad E = O(\epsilon^v).
\end{equation}

\end{proof}
Thus using the Gronwall inequality we find firstly,
\[ \mathbb{E}[\|\hat{X}(t) - \bar{Z}(\hat{t})\|] = O(\epsilon^{1/2+v}/2) . \]

Using this and Gronwall's inequality a second time gives
\[ \mathbb{E}[\|\bar{Z}(\hat{t}) - \bar{X}(\hat{t})\|^2] = O(\epsilon^{1+v} + \epsilon^{1/2+v}/2+u) . \]

Suppose for the moment that \( \bar{Y}^{(h)} \) is uniformly bounded by \( \bar{P} \) with respect to \( \epsilon \) and \( h \). As the processes are bounded by \( \bar{P} \), \( \mathbb{E}[\|\hat{X}(\hat{t}) - \bar{Z}(\hat{t})\|^2] = \mathbb{E}[\|\bar{X}(t) - \bar{Z}(t)\|^2] \) and we get the stated result.

The extra assumption that \( \bar{Y}^{(h)} \) is uniformly bounded can easily be removed by changing the definition of \( \tau \) in (3.9) into
\[ \tau_p := \inf_s \{\|\hat{X}(s)\| \lor \|\bar{Z}(s)\| > \bar{P}\} . \]

The two terms in the error bound can be interpreted as firstly, the error introduced in the macro-species, \( \epsilon^{1+v}/2 \), and secondly, the error made in the meso-species, \( \epsilon^{1/2+v}/2+u \), respectively.

In order to obtain a theorem also in the unbounded case, the growth of the local Lipschitz constants has to be controlled, and so we make the following convenient assumption:

**Assumption 3.1.** There exists \( a_1, \ldots, a_R \geq 0 \) such that \( \bar{L}(\bar{P}) \leq C \bar{P}^{\alpha r} \). Furthermore, we assume \( a_r = 0 \) for each \( r \) such that that \( \nu_r = 0 \). Hence the Lipschitz constants associated with these transitions are bounded independently from \( \bar{P} \).

As in the appendix we use the notation “\( A \leq_C B \)” to indicate that \( A \leq CB \) for some constant \( C > 0 \) which is \( O(1) \) with respect to \( \epsilon, \bar{P}, \) and \( h \).

**Theorem 3.5 (Multiscale error).** Under the scale separation Condition 2.3, and under the regularity Assumptions 2.4, 2.5, and 3.1, and the additional conditions \( u \geq 0, v > 0, \) it holds that
\[ (3.19) \quad \mathbb{E}[\|\bar{Z}(t) - \bar{X}(t)\|^2] = O(\epsilon^{1+v} - \epsilon^{1/2+v}/2+u-). \]

**Proof.** The proof here concerns the case \( u > 0 \). The special case from Assumption 3.1 where \( \nu_r = 0 \) and \( a_r = 0 \) for some \( r \) (and thus \( u = 0 \)) is similar but requires some cumbersome notation and is therefore omitted. Select \( \bar{P} = \epsilon^{-b} \) for some \( b > 0 \) and let \( p > 1, a := \max_r a_r \). Following the same pattern as in the proof of Theorem 3.4 we get
\[ \mathbb{E}[\|\bar{Z}(t) - \bar{X}(t)\|^2] = O(\epsilon^{1+v-(a+1)b} + \epsilon^{1/2+v}/2+u-(3a+1)b/2). \]

Thus using Lemma 3.3,
\[ \mathbb{E}[\|\bar{Z}(t) - \bar{X}(t)\|^2] \leq \mathbb{E}[\|\bar{Z}(t) - \bar{X}(t)\|^21_{t \leq \hat{t}}] + \mathbb{E}[\|\bar{Z}(t) - \bar{X}(t)\|^21_{t > \hat{t}}] \]
\[ \leq \mathbb{E}[\|\bar{Z}(t) - \bar{X}(t)\|^2] + 4K_p\epsilon^{bp/2} \]
\[ = O(\epsilon^{1+v-(a+1)b} + \epsilon^{1/2+v}/2+u-(3a+1)b/2 + \epsilon^{bp/2}). \]

As \( bp/2 \) can be made arbitrarily large while \( (a + 1)b \) and \( (3a + 1)b/2 \) can be made arbitrarily close to 0 (i.e. \( b \to 0 \) and \( p \to \infty \)), we arrive at the stated bound. \( \square \)
Remark. It is possible to get a convergence result for the case \( u = v = 0 \). However, in this case the error bound is of the form \( O(\log(1/\epsilon)^{-\delta}) \) and the dominating part can be traced back to Lemma 3.3.

### 3.3. Splitting convergence

We next consider the error in the approximation \( \bar{Z} \approx Y^{(h)} \), that is, the splitting error. For this part we are able to prove a somewhat weak error bound in the general case, while the situation improves considerably if the processes are assumed to be bounded \textit{a priori}.

**Theorem 3.6 (Splitting error, bounded version).** Under the scale separation Condition 2.3, the regularity Assumptions 2.5, 2.6, and assuming also that \( \bar{X}, \bar{Z} \), and \( Y^{(h)} \) are uniformly bounded with respect to \( h \) and \( \epsilon \) by \( \bar{P} \), then whenever \( u \geq 0, v \geq 0 \) it holds that

\[
\mathbb{E} \left[ \| \bar{Z}(t) - Y^{(h)}(t) \|^2 \right] \leq O \left( \epsilon^2 u + \epsilon^u v \right) + O \left( h^2 \epsilon^2 v \right).
\]

**Proof.** Using Lemma B.2,

\[
\mathbb{E} \left[ \| \bar{Z}(\hat{t}) - Y^{(h)}(\hat{t}) \|^2 \right] = O(\epsilon^u) \int_0^{\hat{t}} \mathbb{E} \left[ \| \bar{Z}(\hat{s}) - Y^{(h)}(\hat{s}) \| \right] ds + O(\epsilon^{2v}) \int_0^{\hat{t}} \mathbb{E} \left[ \| \bar{Z}(\hat{s}) - Y^{(h)}(\hat{s}) \|^2 \right] ds + O(\epsilon^u h) + O(\epsilon^{2v} h^2).
\]

Using Lemma B.3 and the Gronwall inequality, one readily shows that

\[
\mathbb{E} \left[ \| \bar{Z}(\hat{t}) - Y^{(h)}(\hat{t}) \| \right] = O \left( (\epsilon^u + \epsilon^v) h \right).
\]

Taken together we find

\[
\mathbb{E} \left[ \| \bar{Z}(t) - Y^{(h)}(t) \|^2 \right] \leq O \left( (\epsilon^2 u + \epsilon^u v) h \right) + O \left( \epsilon^{2v} h^2 \right)
\]

\[
+ O(\epsilon^{2v}) \int_0^t \mathbb{E} \left[ \| \bar{Z}(s) - Y^{(h)}(s) \|^2 \right] ds.
\]

Hence using the Gronwall inequality anew,

\[
\mathbb{E} \left[ \| \bar{Z}(t) - Y^{(h)}(t) \|^2 \right] = O \left( (\epsilon^2 u + \epsilon^u v) h \right) + O \left( \epsilon^{2v} h^2 \right).
\]

Furthermore, as the processes are bounded by \( \bar{P} \), \( \mathbb{E}[\| \bar{Z}(\hat{t}) - Y^{(h)}(\hat{t}) \|^2] = \mathbb{E}[\| \bar{Z}(t) - Y^{(h)}(t) \|^2] \) and we get the stated result. \( \square \)

As before one can appreciate the two terms of the error as the error made in the meso-species, \( (\epsilon^2 u + \epsilon^u v) h \), and \( \epsilon^{2v} h^2 \), the error introduced in the macro-species.

**Theorem 3.7 (Splitting error).** Under the scale separation Condition 2.3, and under the regularity Assumptions 2.5, 2.6, and 3.1, and the additional conditions \( u > 0, v > 0 \), it holds that

\[
\lim_{h \to 0} \mathbb{E} \left[ \| \bar{Z}(t) - Y^{(h)}(t) \|^2 \right] = 0.
\]

**Proof.** Following the same pattern as in the proof of the bounded version, it is easy to show that for each \( \bar{P} \),

\[
\mathbb{E} \left[ \| \bar{Z}(t) - Y^{(h)}(t) \|^2 \right] \to 0.
\]
We conclude the argument using Lemma 3.3, which implies that

\[ E \left[ \| \bar{Z}(t) - \bar{Y}^{(h)}(t) \|^2_{L^2} \right] \xrightarrow{P \to \infty} 0 \]

uniformly with respect to \( h \).

\( \Box \)

**Remark.** Under the Assumptions of Theorem 3.7, it is possible to get an error bound of the form

\[ E \left[ \| \bar{Z}(t) - \bar{Y}^{(h)}(t) \|^2 \right] \leq O \left( \log(1/h)^{-\delta} \right), \]

for any \( \delta \) greater than some \( \delta_0 \). However, in this case the error can be traced to the unbounded part as covered by Lemma 3.3.

### 4. Numerical examples

We now proceed to illustrate our main findings through some prototypical cases. An all-linear isomerization-type system is investigated in §4.1 and a nonlinear catalytic model in §4.2.

In the experiments below we considered reactions taking place in a one-dimensional geometry \([0,1]\) under periodic boundary conditions. The geometry was discretized into 10 equally spaced segments and a diffusion process implemented via the standard 2nd order finite difference stencil, re-interpreted as linearly dependent transition rates. As for the initial data, we let each segment contain either 10 or 20 molecules for the mesoscopic (discrete) species and 20\( \epsilon^{-1} \) or 10\( \epsilon^{-1} \) for the macroscopic (continuous) species, respectively.

The exact dynamics (2.29) was simulated in an operational time framework. Here we relied on an implementation of the *All Events Method* [5], essentially a spatial extension of the *Common Reaction Path Method* [21] which evolves (2.29) using separate Poisson processes for all events.

The multiscale approximation (2.35)–(2.36) falls under the scope of *Piecewise Deterministic Markov Processes (PDMPs)* for which accurate methods have been proposed [22]. We implemented this through the use of *event-detection* in solvers for Ordinary Differential Equations (ODEs). Notably, this allows for a fully consistent coupling with (2.29) in operational time.

Finally, the split-step approximation (2.43)–(2.44) was implemented. This is quite straightforward via the kernel step function representation and executes very efficiently. The split-step error is much more challenging to determine accurately than the multiscale error is. In fact, on a predetermined grid in time the split-step approximation \( \bar{Y}^{(h)} \) in (2.43) was often found to be exactly equal to the multiscale approximation \( \bar{Z}_{ij} \) in (2.35), thus requiring many realizations for even a very crude estimate.

We make repeated use of the estimator

\[ (4.1) \quad E[(Y - X)(t)]^2 \approx M \equiv \frac{1}{N} \sum_{i=1}^{N} (Y - X)(t; \omega_i)^2, \]
for independent trajectories ($\omega_i$). A basic confidence interval is obtained by computing

$$S^2 \equiv \frac{1}{N-1} \sum_{i=1}^{N} [(Y - X)(t; \omega_i)^2 - M]^2,$$

such that the error in the estimator (4.1) is $\propto S/\sqrt{N}$.

4.1. Isomerization. We first consider the simple linear isomerization reaction pair,

$$A_k A A \rightarrow B B.$$

In order for this example to develop a scale separation, for $A$, the diffusion rate is set to $1/2$ in either direction and per molecule, and for $B$ to 0. By selecting $k_a = 1$ and $k_b = \epsilon$, a scale separation occurs, with $A \sim 10$ and $B \sim 10^{-1} \epsilon$. We may thus evolve the system by the multiscale approximation (2.35)–(2.36), letting $A$ remain discrete while $B$ is approximated with a continuous scaled variable.

Although the unscaled system is closed, from the perspective of scale separation the system scales unfavorably with $\epsilon$ and hence falls under the scope of Theorem 3.5. We have $u = 0$ and $v = 1$ in (3.15)–(3.16) and thus expect a mean square error behaving like $O(\epsilon^2)$ for the macroscopic species and $O(\epsilon)$ for the mesoscopic species. This is verified in Figure 4.1 where the multiscale error for the two components is examined.

Since Theorem 3.6 is formally not applicable, the only result valid is the guaranteed convergence of Theorem 3.7. Nevertheless, in Figure 4.2 the split-step error for the two species have been plotted separately. The different terms of the error estimate in Theorem 3.6 are clearly visible, suggesting that the uniform bounds on the processes, as required by Theorem 3.6, may in fact be relaxed.

Convergence results similar to those of [4] and [18] are here consequences of Theorem 3.5, with the added benefit of an error estimate. Indeed, Theorem 3.5 yields that the difference between $\bar{X}$ and $\bar{Z}$ goes to 0 and the convergence of $\bar{Z}$ is easy to study. Using (2.35) and (2.36) for voxel $j$ yields

$$\bar{Z}_{B,j}(t) = \bar{Z}_{B,j}(0) + \epsilon \int_0^t k_a \bar{Z}_{A,j}(s) ds - \epsilon \int_0^t k_b \epsilon^{-1} \bar{Z}_{B,j}(s) ds \xrightarrow{\epsilon \to 0} \bar{Z}_{B,j}(0),$$

since $(k_a, k_b) = (1, \epsilon)$, and,

$$\bar{Z}_{A,j}(t) \xrightarrow{\epsilon \to 0} \bar{Z}_{A,j}(0) + \Pi_{1,j} (\bar{Z}_{B,j}(0)t) - \Pi_{2,j} \left( \int_0^t \bar{Z}_{A,j}(s) ds \right) + \sum_{k \in \{j-1,j+1\}} \Pi'_{A,k,j} \left( \int_0^t \bar{Z}_{A,k}(s)/2 ds \right) - \Pi'_{A,j,k} \left( \int_0^t \bar{Z}_{A,j}(s)/2 ds \right).$$

Hence for this simple system, the limit $\epsilon \to 0$ for $B$ is trivial.

4.2. Catalytic reactions. We consider the following pair of catalytic reactions:

$$\begin{align*}
A + B & \xrightarrow{k_{AB}} C + B \\
C + D & \xrightarrow{k_{CD}} A + D \\
B & \xrightarrow{k_b B}{k_d D} D
\end{align*}$$

(4.6)
Figure 4.1. Multiscale error (isomerization): the root mean-square (RMS) error as a function of the scale separation $\epsilon$ for the two components $A$ (discrete) and $B$ (continuous).

We assume that species $A$ and $C$ are abundant and $O(\epsilon^{-1})$, and species $B$ and $D$ are $O(1)$. For the diffusion we put $\sigma_{A,C} = \epsilon$ and $\sigma_{B,D} = 1$, and for the rates $k = 0.01$ and $(k_b, k_d) = (1, 0.9)$. The system so defined is closed since there is no coupling from the macro-species to the meso-species (take $l = [1, 1, 1, 1]^T$ in Assumption 2.4). This property carries over to the multiscale and split-step approximations (cf. Assumptions 2.5 and 2.6).

For the scale separation, we get the critical exponents $u = v = 0$ and Theorem 3.4 predicts a slow convergence of $O(\epsilon^{1/4})$ in the RMS sense. However, since the meso-species do not depend on the macro-species the corresponding error is in fact 0. According to the discussion following the proof of Theorem 3.4, the RMS is therefore $O(\epsilon^{1/2})$ and is observed in the macroscopic species only. By the same argument, and from the remark following the proof of Theorem 3.6, we predict that the RMS of the split-step error is $O(h)$.

Experimental results verifying this are shown in Figure 4.3 for the multiscale error ("convergent scaling") and in Figure 4.4 for the split-step error.

Like in the previous example, convergence results similar to those of [4] and [18] are consequences of Theorem 3.4. This time, (2.35) and (2.36) are almost independent of $\epsilon$; only the diffusion for $A$ and $C$ depend on $\epsilon$ and, since $\sigma_{A,C} = \epsilon$,
it vanishes in the limit. For voxel $j$,

$$Z_{A,j}(t) \xrightarrow{\epsilon \to 0} -\int_0^t k Z_{A,j}(s) Z_{B,j}(s) ds + \int_0^t k Z_{C,j}(s) Z_{D,j}(s) ds,$$

$$Z_{B,j}(t) = -\Pi_{1,j} \left( k_b \int_0^t Z_{B,j}(s) ds \right) + \Pi_{2,j} \left( \int_0^t k_a Z_{D,j}(s) ds \right)$$

$$+ \sum_{k \in \{j-1, j+1\}} \Pi'_{B,k,j} \left( \int_0^t Z_{B,k}(s) ds \right) - \Pi'_{B,j,k} \left( \int_0^t Z_{B,j}(s) ds \right).$$

The defining equations for $Z_{C,j}$ and $Z_{D,j}$ are similar. Thus the limit in this case is not a trivial process, stressing that non-trivial models can be described within the framework.

4.3. Catalytic reactions: case of unclear scale separation. It is interesting to turn the scales of the catalytic model around. If we instead let species $A$ and $C$ be $O(1)$, while $B$ and $D$ are $O(\epsilon^{-1})$, the topology does not change and we still have a closed system. We put $k = 0.01 \epsilon^{1/4}$, $(k_b, k_d) = (1, 0.9)$, and use the slow diffusion $\sigma_{A,C} = \sigma_{B,D} = \epsilon$. The critical exponents become $u = -3/4$ and $v = 0$ and thus none of the results apply. Although Figure 4.3 ("divergent scaling") does not strictly exclude the possibility of convergence, the error certainly does not go down convincingly.
\( \varepsilon \sim 10^{-3} \)

\( \varepsilon \sim 10^{-2} \)

\( \varepsilon \sim 10^{-1} \)

\( \varepsilon \sim 10^{0} \)

Figure 4.3. Multiscale errors (catalytic reactions).

\( \varepsilon \sim h^{1/2} \)

\( \propto h \)

Figure 4.4. Split-step error (catalytic reactions). Case of superconvergence of the split-step method.
5. Conclusions

In this paper we have developed a coherent framework for analyzing certain multiscale methods for continuous-time Markov chains of a general spatial structure. Concrete assumptions and conditions have been discovered that enables a multiscale description and a consistent formulation of the approximating methods in operational time. Notably, through explicit \textit{a priori} results, all processes are well-posed and the framework does not rely on any heuristic prior bounds.

The analysis distinguishes between two separate sources of errors, namely \textit{the multiscale error} and \textit{the split-step error}. The first is due to an approximate stochastic/deterministic variable splitting strategy, a kind of stochastic homogenization technique. The second emerges when this approximating process in turn is evolved in discrete time-steps. Notably, we found theoretically how the split-step error is composed of factors remindful of the terms making up the multiscale error, thus connecting the two in a qualitative sense. The behavior of these errors were also examined experimentally via actual implementations of the methods. Although some of the boundary cases are difficult to handle theoretically, in particular when confronted with open systems, the numerical experiments support the sharpness of our theoretical predictions.

The work opens up for some interesting possibilities. Clearly, an ideal implementation should allow the split-step error to be about as large as the multiscale error. The fully discrete approximation is amenable to several efficient algorithms developed for numerical methods for partial differential equations, including for example multigrid techniques. An interesting challenge to which we would like to return is to develop practical procedures for computing accurate error estimates. We believe this is doable following the theory laid out in the paper.

Acknowledgment

Many detailed suggestions by the two referees have helped us to clarify and improve the paper.

The work was supported by ENS Cachan (A. Chevallier) and by the Swedish Research Council within the UPMARC Linnaeus center of Excellence (S. Engblom).

Appendix A. The multiscale error

Below are the statements and proofs of the two critical lemmas used in the proof of Theorem 3.5. Recall the definition of the two effective exponents $u$ and $v$ in (3.15)–(3.16).

Lemma A.1. Define $\bar{X}(t)$ by (2.29) and $\bar{Z}(t)$ by (2.35)–(2.36) with $\bar{X}(0) = \bar{Z}(0)$ almost surely. Then under the stopping time $\hat{t}$ defined in (3.9),

$$E[\|\bar{Z}(\hat{t}) - \bar{X}(\hat{t})\|^2] \leq C A + B \int_0^t E[\|\bar{Z}(\hat{s}) - \bar{X}(\hat{s})\|^2] ds + C \int_0^t E[\|\bar{Z}(\hat{s}) - \bar{X}(\hat{s})\|] ds,$$

where expressions for $A$, $B$, and $C$ are indicated in (A.3) below. These bounds depend on $\epsilon$ and $\bar{P}$ and on the reaction topology $\mathbb{N}$,

$$A = \epsilon^{1+u}[\bar{L}(\bar{P})\bar{P}], \quad B = \epsilon^{2u}[\bar{L}(\bar{P})]^2, \quad C = \epsilon^u\bar{L}(\bar{P})[\epsilon^u\bar{L}(\bar{P})\bar{P} + 1].$$

To improve the readability of the proof, we use the notation “$A \leq C B$” to indicate that $A \leq CB$ for some constant $C > 0$ which is $O(1)$ with respect to...
where in terms of 

\[ \epsilon, \bar{P}, \text{ and } h. \] 

When the processes are assumed to be bounded \textit{a priori}, clearly, 

\[ \bar{L}(c\bar{P}) \leq C_1, \] 

for any constant \( c > 0. \) In the unbounded case, Assumption 3.1 yields similarly \( \bar{L}(c\bar{P}) \leq C \bar{L}(\bar{P}) \) for any constant \( c > 0. \) We additionally let \((c_l, C_l)\) be the constants in the norm equivalence

\[
(A.2) \quad c_l \| X \|_2 \leq \| X \|_t \leq C_l \| X \|_2.
\]

**Proof.** We focus first on a single voxel \( j \) and analyze the errors on species from \( G_1(j) \) and \( G_2(j) \), respectively. For \( i \in G_1(j) \), from (2.29) and (2.35),

\[
(\hat{Z}_{ij}(t) - \bar{X}_{ij}(t))^2 = \left[ -\sum_{r=1}^R N_{ri}(\Pi_{rj}(\cdot) - \Pi_{rj}(\cdot)) - \sum_{k=1}^J (\Pi'_{ijk}(\cdot) - \Pi'_{ijk}(\cdot)) + \sum_{k=1}^J (\Pi''_{ijk}(\cdot) - \Pi''_{ijk}(\cdot)) \right]^2
\]

where we have suppressed the local time arguments of the Poisson processes, available in (2.29) and (2.35).

By Jensen’s inequality and the bound on the mesh connectivity in Definition 2.1 (2.7) we get

\[
(\hat{Z}_{ij}(t) - \bar{X}_{ij}(t))^2 \leq (R + 2M_D)(A_1 + A_2 + A_3),
\]

where in terms of

\[
A_1 = \sum_{r=1}^R N_{rt}^2 \left( \Pi_{rj} \left( \int_0^t \epsilon^{-\nu} V_j \tilde{u}_r(V_{ij}^{-1} \tilde{Z}_{ij}(s)) ds \right) \right)^2
\]

\[
A_2 = \sum_{k=1}^J \left( \Pi'_{ijk} \left( \int_0^t \epsilon^{-\mu} \bar{q}_{ijk} \bar{Z}_{ij}(s) ds \right) - \Pi'_{ijk} \left( \int_0^t \epsilon^{-\mu} \bar{q}_{ijk} \bar{X}_{ij}(s) ds \right) \right)^2
\]

\[
A_3 = \sum_{k=1}^J \left( \Pi''_{ijk} \left( \int_0^t \epsilon^{-\mu} \bar{q}_{ikj} \bar{Z}_{ik}(s) ds \right) - \Pi''_{ijk} \left( \int_0^t \epsilon^{-\mu} \bar{q}_{ikj} \bar{X}_{ik}(s) ds \right) \right)^2
\]

First we need to bound the \( L \)-norm:

\[
\| V_j^{-1} \tilde{Z}_{ij}(s) \|_t \leq C_l \| V_j^{-1} \tilde{Z}_{ij}(s) \|_2 \leq C_l V_j^{-1} \bar{P} \leq C_l m^{-1}_V V^{-1}_M \bar{P}.
\]

Then using the Lipschitz bound (2.32) in Assumption 2.4:

\[
\bar{u}_r(V_j^{-1} \tilde{Z}_{ij}(s)) \leq \bar{u}_r(0) + \bar{L}_r(C_l m^{-1}_V V^{-1}_M \bar{P}) \| V_j^{-1} \tilde{Z}_{ij}(s) \|_t
\]

\[
\leq \bar{u}_r(0) + \bar{L}_r(C_l m^{-1}_V V^{-1}_M \bar{P}) \| V_j^{-1} \tilde{Z}_{ij}(s) \|_2
\]

\[
\leq C_1 + \bar{L}_r(\bar{P}) V_j^{-1} \bar{P}.
\]
Thus,

$$\int_0^t e^{-\nu r} V_j \bar{u}_r (V_j^{-1} \bar{Z}_{-j}(s)) \, ds \leq C \int_0^t e^{-\nu r} (V_j + \bar{L}_r(\bar{P}) \bar{P}) \, ds$$

$$\leq C e^{-\nu r} (MV \bar{V}_M + \bar{L}_r(\bar{P}) \bar{P}) \, ds \leq C e^{-\nu r} (1 + \bar{L}_r(\bar{P}) \bar{P}) \, ds.$$ 

Using the same method for $\bar{X}$, we conclude

$$\int_0^t e^{-\nu r} V_j \bar{u}_r (V_j^{-1} \bar{Z}_{-j}(s)) \, ds \vee \int_0^t e^{-\nu r} V_j \bar{u}_r (V_j^{-1} \bar{X}_{-j}(s)) \, ds$$

$$\leq C e^{-\nu r} (1 + \bar{L}_r(\bar{P}) \bar{P}).$$

Hence using Lemma 3.2 (3.7) and again the Lipschitz bound we get

$$\mathbb{E}[A_1] \leq C \sum_r N_r^2 \int_0^t e^{-\nu r} \bar{L}_r(\bar{P}) \left( e^{-\nu r} (\bar{L}_r(\bar{P}) \bar{P} + 1) \right) \, ds.$$ 

Relying on the same arguments we readily find

$$\mathbb{E}[A_2] \leq C \sum_{k=1}^J \int_0^t \mathbb{E}[\|\bar{Z}(\bar{s}) - \bar{X}(\bar{s})\|] \, ds,$$

and the identical bound for $\mathbb{E}[A_3]$.

For $i \in G_2(j)$, we similarly get

$$\left( \bar{Z}_{ij}(\bar{t}) - \bar{X}_{ij}(\bar{t}) \right)^2 \leq C (R + 2M_D)(A_1' + A_2' + A_3')$$

where

$$A_1' = \sum_{r=1}^R \sum_{i,s} N_r^2 \int_0^t e^{-\nu r} V_j \bar{u}_r (V_j^{-1} Z_{-j}(s)) \, ds - \Pi_{ij} \left( \int_0^t e^{-\nu r} V_j \bar{u}_r (V_j^{-1} X_{-j}(s)) \, ds \right)^2,$$

$$A_2' = \sum_{k=1}^J \left( \int_0^t e^{-\nu r} \bar{q}_{ijk} \bar{Z}_{ij}(s) \, ds - \Pi_{ijk} \left( \int_0^t e^{-\nu r} \bar{q}_{ijk} \bar{X}_{ij}(s) \, ds \right)^2 \right),$$

$$A_3' = \sum_{k=1}^J \left( \int_0^t e^{-\nu r} \bar{q}_{jk} \bar{Z}_{ij}(s) \, ds - \Pi_{ikj} \left( \int_0^t e^{-\nu r} \bar{q}_{ik} \bar{X}_{ik}(s) \, ds \right)^2 \right).$$

The analysis is now slightly different. Species from the second group have a large number of molecules, so $X_{ij}(t)$ is expected to remain close to its mean value. We
thus introduce the centered Poisson processes $\tilde{\Pi}_r$,
\[
A'_1 = \sum_r N_{r_1}^2 \left( \int_0^t \epsilon^{-\nu r} V_j (u_r(V_j^{-1} \tilde{Z}_{-j}(s)) - \tilde{u}_r(V_j^{-1} \tilde{X}_{-j}(s))) \, ds \right. \\
- \tilde{\Pi}_{jr} \left( \int_0^t \epsilon^{-\nu r} V_j \tilde{u}_r(V_j^{-1} \tilde{X}_{-j}(s)) \, ds \right) \right)^2 \\
\leq \sum_r 2N_{r_1}^2 \left( \int_0^t \epsilon^{-\nu r} V_j (u_r(V_j^{-1} \tilde{Z}_{-j}(s)) - \tilde{u}_r(V_j^{-1} \tilde{X}_{-j}(s))) \, ds \right)^2 \\
+ 2N_{r_1}^2 \left( \tilde{\Pi}_{jr} \left( \int_0^t \epsilon^{-\nu r} V_j \tilde{u}_r(V_j^{-1} \tilde{X}_{-j}(s)) \, ds \right) \right)^2.
\]

Using the quadratic variation of $\tilde{\Pi}$ is $[\tilde{\Pi}]_t =\Pi(t)$ and the martingale stopping time theorem we get
\[
\mathbb{E} \left[ \left( \tilde{\Pi}_{jr} \left( \int_0^t \epsilon^{-\nu r} V_j \tilde{u}_r(V_j^{-1} \tilde{X}_{-j}(s)) \, ds \right) \right)^2 \right] \\
= \mathbb{E} \left[ \Pi_{jr} \left( \int_0^t \epsilon^{-\nu r} V_j \tilde{u}_r(V_j^{-1} \tilde{X}_{-j}(s)) \, ds \right) \right] \\
= \mathbb{E} \left[ \int_0^t \epsilon^{-\nu r} V_j \tilde{u}_r(V_j^{-1} \tilde{X}_{-j}(s)) \, ds \right] \leq C \epsilon^{-\nu r} \tilde{L}_r(\tilde{P}) \tilde{P}.
\]

Using Cauchy-Schwartz for the remaining integral part and following the same approach for $A'_2$ and $A'_3$ we get
\[
\mathbb{E}[A'_1] \leq C \sum_{r=1}^R N_{r_1}^2 \epsilon^{-\nu r} \tilde{L}_r(\tilde{P}) \tilde{P} + \sum_{r=1}^R N_{r_1}^2 (\epsilon^{-\nu r} \tilde{L}_r(\tilde{P}))^2 \int_0^t \mathbb{E}[\|\tilde{Z}(\hat{s}) - \tilde{X}(\hat{s})\|^2] \, ds, \\
\mathbb{E}[A'_2] \leq C \sum_{k=1}^j \epsilon^{-\mu k} + \sum_{k=1}^j (\epsilon^{-\mu k})^2 \int_0^t \mathbb{E}[\|\tilde{Z}(\hat{s}) - \tilde{X}(\hat{s})\|^2] \, ds,
\]
as well as an identical bound for $\mathbb{E}[A'_3]$.

We thus get for the $j$th voxel,
\[
\mathbb{E} \left[ \|\tilde{Z}_{-j}(t) - \tilde{X}_{-j}(t)\|^2 \right] \leq C \sum_{i \in G_1(j)} \mathbb{E}[A_1 + A_2 + A_3] + \sum_{i \in G_2(j)} \epsilon^2 \mathbb{E}[A'_1 + A'_2 + A'_3] \tag{A.3}
\]
\[
\leq C A^{(j)} + B^{(j)} \int_0^t \mathbb{E}[\|\tilde{X}(\hat{s}) - \tilde{Z}(\hat{s})\|^2] \, ds + C^{(j)} \int_0^t \mathbb{E}[\|\tilde{X}(\hat{s}) - \tilde{Z}(\hat{s})\|] \, ds.
\]

Summing over $j$ we get the stated result with $A := \sum_j A^{(j)}$, $B := \sum_j B^{(j)}$, and $C := \sum_j C^{(j)}$. \hfill \Box

**Lemma A.2.** Under the same assumptions as in Lemma A.1,
\[
\mathbb{E}[\|\tilde{X}(\hat{t}) - \tilde{Z}(\hat{t})\|] \leq C D + E \int_0^t \mathbb{E}[\|\tilde{X}(\hat{s}) - \tilde{Z}(\hat{s})\|] \, ds,
\]
where explicit expressions for \( D \) and \( E \) are found in (A.5) below and depend on \( c, \bar{P} \), and on the reaction topology \( \mathbb{N} \),

\[
D = c^{1/2 + \nu/2}[\bar{L}(\bar{P})\bar{P}]^{1/2}, \quad E = [c^u + c^v]\bar{L}(\bar{P}).
\]

**Proof.** For voxel \( j \) and for \( i \in G_1(j) \),

\[
|\bar{Z}_{ij}(\bar{t}) - \bar{X}_{ij}(\bar{t})| \leq \sum_{r=1}^R |\mathbb{N}_{ri} (\Pi_{rj}(\cdot) - \Pi_{rj}(\cdot))| + \sum_{k=1}^J |(\Pi'_{ijk}(\cdot) - \Pi'_{ijk}(\cdot))| + \sum_{k=1}^J |(\Pi_{ijk}(\cdot) - \Pi_{ijk}(\cdot))|.
\]

We keep the same notation as in the previous lemma and thus write

\[
|\bar{Z}_{ij}(\bar{t}) - \bar{X}_{ij}(\bar{t})| \leq (A_1 + A_2 + A_3),
\]

where

\[
\mathbb{E}[A_1] = \sum_r |\mathbb{N}_{ri}| \mathbb{E}[|\Pi_{rj}(\cdot) - \Pi_{rj}(\cdot)|]
\]

\[
= \sum_r |\mathbb{N}_{ri}| \mathbb{E} \left[ \int_0^i \epsilon^{-\nu r} V_j \bar{u}_r(V_j^{-1} \bar{Z}_{.j}(s)) ds - \int_0^i \epsilon^{-\nu r} V_j \bar{u}_r(V_j^{-1} \bar{X}_{.j}(s)) ds \right]
\]

\[
\leq C \sum_r |\mathbb{N}_{ri}| \epsilon^{-\nu r} \bar{L}_r(\bar{P}) \int_0^i \mathbb{E} [||\bar{Z}_{.j}(\bar{s}) - \bar{X}_{.j}(\bar{s})||].
\]

In the same spirit we find

\[
\mathbb{E}[A_2] \leq C \sum_{k=1}^J \epsilon^{-\mu_i} \int_0^i \mathbb{E} [||\bar{Z}_{.j}(\bar{s}) - \bar{X}_{.j}(\bar{s})||],
\]

and an identical bound for \( \mathbb{E}[A_3] \).

Continuing with \( i \in G_2(j) \),

\[
(\bar{Z}_{ij}(\bar{t}) - \bar{X}_{ij}(\bar{t})) \leq \epsilon (A_1' + A_2' + A_3'),
\]

where

\[
A_1' = \sum_{r=1}^R |\mathbb{N}_{ri}| \left| \int_0^i V_j \epsilon^{-\nu r} \bar{u}_r(V_j^{-1} \bar{Z}_{.j}(s)) ds - \Pi_{rj} \left( \int_0^i V_j \epsilon^{-\nu r} \bar{u}_r(V_j^{-1} \bar{X}_{.j}(s)) ds \right) \right|,
\]

\[
A_2' = \sum_{k=1}^J \left| \int_0^i \epsilon^{-\mu_i} \bar{q}_{ijk} \bar{Z}_{ij}(s) ds - \Pi'_{ijk} \left( \int_0^i \epsilon^{-\mu_i} \bar{q}_{ijk} \bar{X}_{ij}(s) ds \right) \right|,
\]

\[
A_3' = \sum_{k=1}^J \left| \int_0^i \epsilon^{-\mu_i} \bar{q}_{ikj} \bar{Z}_{ik}(s) ds - \Pi'_{ikj} \left( \int_0^i \epsilon^{-\mu_i} \bar{q}_{ikj} \bar{X}_{ik}(s) ds \right) \right|.
\]
Using the same techniques developed previously we find
\[
\mathbb{E}[A'_1] \leq C \sum_r |N_r| |\epsilon^{-\nu_r} C_t \tilde{L}_r(V_j^{-1}\tilde{P})||\tilde{Z}(s) - \tilde{X}(s)|| \ ds \\
+ |N_r| |\hat{\Pi}_{rj} \left( \int_0^t V_j \epsilon^{-\nu_r} \tilde{u}_r(V_j^{-1}\tilde{X}_{r,j}(s)) \ ds \right)|| \\
\leq C \sum_r |N_r| |\epsilon^{-\nu_r} \tilde{L}_r(\tilde{P})| \int_0^t \mathbb{E}[||\tilde{Z}(s) - \tilde{X}(s)||] \ ds \\
+ |N_r| |\hat{\Pi}_{rj} \left( \int_0^t V_j \epsilon^{-\nu_r} \tilde{u}_r(V_j^{-1}\tilde{X}_{r,j}(s)) \ ds \right)||^{1/2} \\
\leq C \sum_r |N_r| |\epsilon^{-\nu_r} \tilde{L}_r(\tilde{P})| \int_0^t \mathbb{E}[||\tilde{Z}(s) - \tilde{X}(s)||] \ ds \\
+ |N_r| |\epsilon^{-\nu_r} \tilde{L}_r(\tilde{P})| \int_0^t \mathbb{E}[||\tilde{Z}(s) - \tilde{X}(s)||] \ ds + |N_r| |\epsilon^{-\nu_r} \tilde{L}_r(\tilde{P})||^{1/2} \\
\leq C \sum_r |N_r| |\epsilon^{-\nu_r} \tilde{L}_r(\tilde{P})| \int_0^t \mathbb{E}[||\tilde{Z}(s) - \tilde{X}(s)||] \ ds + (\epsilon^{-\nu_r} \tilde{L}_r(\tilde{P})|^{1/2}.
\]

In much the same spirit we get
\[
\mathbb{E}[A'_2] \leq C \sum_j \epsilon^{-\mu_j} |\int_0^t \mathbb{E}[||\tilde{Z}(s) - \tilde{X}(s)||] \ | \ ds + (\epsilon^{-\mu_j})^{1/2},
\]
along with an identical bound for \(\mathbb{E}[A'_3]\).

Combined we thus get for the \(j\)th voxel,
\[
\mathbb{E}[||\tilde{Z}_{r,j}(\tilde{t}) - \tilde{X}_{r,j}(\tilde{t})||] \leq C \sum_{i \in G_1(j)} \mathbb{E}[A_1 + A_2 + A_3] + \sum_{i \in G_2(j)} \epsilon \mathbb{E}[A'_1 + A'_2 + A'_3] \\
(\text{A.5}) \\
\leq C \ D_j + E_j \int_0^t \mathbb{E}[||\tilde{X}(s) - \tilde{Z}(s)||] \ ds.
\]
Summing over \(j\) we get the stated result. \(\square\)

Appendix B. The split-step error

The consistency of the numerical split-step method hinges on the regularity of the kernel function \(\sigma_h(s)\). The following lemma (borrowed from [11, Lemma 3.7]) paired with the strong regularity of the involved processes provides for the order estimate in Theorems 3.6 and 3.7. Note that the result can be thought of as càdlàg-version of the Riemann-Lebesgue lemma.

Lemma B.1. ([11, Lemma 3.7]) Let \(G : \mathbb{R}^D \to \mathbb{R}\) be a globally Lipschitz continuous function with Lipschitz constant \(L\) and let \(f : \mathbb{R} \to \mathbb{R}^D\) be a piecewise constant càdlàg function. Then
\[
\left( \int_0^t \sigma_h(s)G(f(s)) \ ds \right) \leq \frac{h}{2} |G(f(t))| + \frac{h}{2} LV_{[0,1]}(f),
\]

(B.1)
where the total absolute variation may be exchanged with the square root of the quadratic variation \( \left[ f_1^t \right]^{1/2} \). If \( t \) is a multiple of \( h \), then the first term on the right side of (B.1) vanishes.

The proofs of the following two lemmas follow closely the ideas in the proofs of Lemmas A.1 and A.2, but using in addition Lemma B.1 and Theorem 2.7 to bound certain additional terms.

**Lemma B.2.** Define \( \bar{Z}(t) \) by (2.35)–(2.36) and \( \bar{Y}^{(h)}(t) \) by (2.43)–(2.44) with \( \bar{Z}(0) = Y^{(h)}(0) \) almost surely. The under the stopping time \( \hat{t} \) defined in (3.9), for a fixed \( \epsilon \) and \( h \) small enough,

\[
\mathbb{E}[\|\bar{Z}(\hat{t}) - \bar{Y}^{(h)}(\hat{t})\|^2] \leq C A + B \int_0^t \mathbb{E}[\|\bar{Z}(\hat{s}) - \bar{Y}^{(h)}(\hat{s})\|^2] \, ds
\]

where

\[
A = \epsilon^u \bar{L}(\bar{P})[\epsilon^u \bar{L}(\bar{P}) \bar{P} + 1] h + \epsilon^{2\nu}[\bar{L}(\bar{P})]^2 h^2,
\]

\[
B = \epsilon^{2\nu}[\bar{L}(\bar{P})]^2, \quad C = \epsilon^u \bar{L}(\bar{P})[\epsilon^u \bar{L}(\bar{P}) \bar{P} + 1].
\]

**Lemma B.3.** Under the same assumptions as in Lemma B.2,

\[
\mathbb{E}[\|\bar{Z}(\hat{t}) - \bar{Y}^{(h)}(\hat{t})\|] \leq C D + E \int_0^t \mathbb{E}[\|\bar{Z}(\hat{s}) - \bar{Y}^{(h)}(\hat{s})\|] \, ds,
\]

with

\[
D = [\epsilon^u + \epsilon^{2\nu}]\bar{L}(\bar{P}) h, \quad E = [\epsilon^u + \epsilon^{2\nu}]\bar{L}(\bar{P}).
\]

**References**

[1] A. Alfonsi, E. Cancés, G. Turinici, B. D. Ventura, and W. Huisinga. Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems. *ESAIM: Proc.*, 14:1–13, 2005. doi:10.1051/proc:2005001.

[2] D. F. Anderson, A. Ganguly, and T. G. Kurtz. Error analysis of tau-leap simulation methods. *Ann. Appl. Probab.*, 21(6):2226–2262, 2011. doi:10.1214/10-AAP756.

[3] G. Arampatzis, M. Katsoulakis, and P. Plecháč. Parallelization, processor communication and error analysis in lattice kinetic monte carlo. *SIAM J. Numer. Anal.*, 52(3):1156–1182, 2014. doi:10.1137/120889459.

[4] K. Ball, T. G. Kurtz, L. Popovic, and G. Rempala. Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab.*, 16(4):1925–1961, 2006. doi:10.1214/1050516060000000420.

[5] P. Bauer and S. Engblom. Sensitivity estimation and inverse problems in spatial stochastic models of chemical kinetics. In A. Abdulle, S. Deparis, D. Kressner, F. Nobile, and M. Picasso, editors, *Numerical Mathematics and Advanced Applications: ENUMATH 2013*, volume 103 of *Lecture Notes in Computational Science and Engineering*, pages 519–527, Berlin, 2015. Springer. doi:10.1007/978-3-319-10705-9_51.

[6] P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Number 31 in Texts in Applied Mathematics. Springer, New York, 1999.
[7] Y. Cao, D. T. Gillespie, and L. R. Petzold. The slow-scale stochastic simulation algorithm. *J. Chem. Phys.*, 122(1):014116, 2005. doi:10.1063/1.1824902.

[8] B. Drawert, S. Engblom, and A. Hellander. URDME: a modular framework for stochastic simulation of reaction-transport processes in complex geometries. *BMC Syst. Biol.*, 6(76):1–17, 2012. doi:10.1186/1752-0509-6-76.

[9] W. E, D. Liu, and E. Vanden-Eijnden. Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates. *J. Chem. Phys.*, 123(19):194107, 2005. doi:10.1063/1.2109987.

[10] S. Engblom. On the stability of stochastic jump kinetics. *Appl. Math.*, 5(19):3217–3239, 2014. doi:10.4236/am.2014.519300.

[11] S. Engblom. Strong convergence for split-step methods in stochastic jump kinetics. *SIAM J. Numer. Anal.*, 53(6):2655–2676, 2015. doi:10.1137/141000841.

[12] S. Engblom, L. Ferm, A. Hellander, and P. Lötstedt. Simulation of stochastic reaction-diffusion processes on unstructured meshes. *SIAM J. Sci. Comput.*, 31(3):1774–1797, 2009. doi:10.1137/080721388.

[13] S. Engblom. Parallel in time simulation of multiscale stochastic chemical kinetics. *Multiscale Model. Simul.*, 8(1):46–68, 2009. doi:10.1137/080733273.

[14] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley series in Probability and Mathematical Statistics. John Wiley & Sons, New York, 1986.

[15] A. Ganguly, D. Altıntan, and H. Koepfl. Jump-diffusion approximation of stochastic reaction dynamics: Error bounds and algorithms. *Multiscale Model. Simul.*, 13(4):1390–1419, 2015. doi:10.1137/1409843471.

[16] E. L. Haseltine and J. B. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.*, 117(15):6959–6969, 2002. doi:10.1063/1.1505860.

[17] T. Jahnke and M. Kreim. Error bound for piecewise deterministic processes modeling stochastic reaction systems. *Multiscale Model. Simul.*, 10(4):1119–1147, 2012. doi:10.1137/120871894.

[18] H.-W. Kang and T. G. Kurtz. Separation of time-scales and model reduction for stochastic reaction networks. *Ann. Appl. Probab.*, 23(2):529–583, 2013. doi:10.1214/12-AAP841.

[19] S. Plyasunov and A. P. Arkin. Averaging methods for stochastic dynamics of complex reaction networks: Description of multiscale couplings. *Multiscale Model. Simul.*, 5(2):497–513, 2006. doi:10.1137/050633822.

[20] P. E. Protter. *Stochastic Integration and Differential Equations*. Number 21 in Stochastic Modelling and Applied Probability. Springer, Berlin, 2nd edition, 2005. *Version 2.1*.

[21] M. Rathinam, P. W. Sheppard, and M. Khammash. Efficient computation of parameter sensitivities of discrete stochastic chemical reaction networks. *J. Chem. Phys.*, 132(3):034103, 2010. doi:10.1063/1.3280166.

[22] M. G. Riedler. Almost sure convergence of numerical approximations for piecewise deterministic Markov processes. *J. Comput. Appl. Math.*, 239(0):50–71, 2013. doi:10.1016/j.cam.2012.09.021.

[23] H. Salis, V. Sotiropoulos, and Y. N. Kaznessis. Multiscale Hy3S: Hybrid stochastic simulation for supercomputers. *BMC Bioinformatics*, 7(93), 2006. doi:10.1186/1471-2105-7-93.
[24] A. Samant, B. A. Ogunnaike, and D. G. Vlachos. A hybrid multiscale Monte Carlo algorithm (HyMSMC) to cope with disparity in time scales and species populations in intracellular networks. *BMC Bioinformatics*, 8(1):175, 2007. doi:10.1186/1471-2105-8-175.

[25] C. A. Yates and M. B. Flegg. The pseudo-compartment method for coupling partial differential equation and compartment-based models of diffusion. *J. R. Soc. Interface*, 12(106), 2015. doi:10.1098/rsif.2015.0141.

INRIA Sophia Antipolis, 2004 Route des Lucioles, 06902 Valbonne, France.
E-mail address: Augustin.Chevallier@ens-cachan.org

Division of Scientific Computing, Department of Information Technology, Uppsala University, SE-751 05 Uppsala, Sweden.
URL: http://user.it.uu.se/~stefane
E-mail address: stefane@it.uu.se