MULTISCALE MODEL. SIMUL.
Vol. 17, No. 1, pp. 1–30
© 2019 Society for Industrial and Applied Mathematics

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

ON CONSTRAINED LANGEVIN EQUATIONS AND (BIO)CHEMICAL REACTION NETWORKS*

DAVID F. ANDERSON†, DESMOND J. HIGHAM‡, SAUL C. LEITE§, AND RUTH J. WILLIAMS¶

Abstract. Stochastic effects play an important role in modeling the time evolution of chemical reaction systems in fields such as systems biology, where the concentrations of some constituent molecules can be low. The most common stochastic models for these systems are continuous time Markov chains, which track the molecular abundance of each chemical species. Often, these stochastic models are studied by computer simulations, which can quickly become computationally expensive. A common approach to reduce computational effort is to approximate the discrete valued Markov chain by a continuous valued diffusion process. However, existing diffusion approximations either do not respect the constraint that chemical concentrations are never negative (linear noise approximation) or are typically only valid until the concentration of some chemical species first becomes zero (chemical Langevin equation). In this paper, we propose (obliquely) reflected diffusions, which respect the nonnegativity of chemical concentrations, as approximations for Markov chain models of chemical reaction networks. These reflected diffusions satisfy "constrained Langevin equations," in that they behave like solutions of chemical Langevin equations in the interior of the positive orthant and are constrained to the orthant by instantaneous oblique reflection at the boundary. To motivate their form, we first illustrate our constrained Langevin approximations for two simple examples. We then describe the general form of our proposed approximation. We illustrate the performance of our approximations through comparison of their stationary distributions for the two examples with those of the Markov chain model and through simulations of more complex examples.

Key words. density dependent Markov chains, diffusion approximation, Langevin equation, linear noise approximation, chemical reaction networks, stochastic differential equation with reflection, systems biology

AMS subject classifications. Primary, 60J28, 60J60, 65C30, 92C45; Secondary, 60H10, 65C40, 92C40

DOI. 10.1137/18M1190999

1. Introduction.

Reacting chemical species are often modeled by deterministic differential equations representing the time evolution of molecular concentrations. Nonetheless, at a finer scale, chemical reaction systems are fundamentally stochastic in nature. Deterministic models provide a mean field approximation to these systems and are generally good predictors when the abundances of all species are high enough

*Received by the editors May 30, 2018; accepted for publication (in revised form) October 8, 2018; published electronically January 3, 2019. An extended abstract of the current paper is included in *Oberwolfach Rep.*, 14 (2017), pp. 1747–1804.

http://www.siam.org/journals/mms/17-1/M119099.html

Funding: The first author was supported by NSF grant DMS-1318832 and Army Research Office grants W911NF-14-1-0401 and W911NF-18-1-0324. The second author was supported by grant EP/M00158X/1 from the EPSRC/RCUK Digital Economy Programme and by grant EP/P020720/1 from the EPSRC. The third author was supported by CAPES Process 23071.012226/2013-06 and by FAPEMIG Process APQ 00945/14. The fourth author was supported in part by NSF grants DMS-1463657 and DMS-1712974.

†Department of Mathematics, University of Wisconsin - Madison, Madison, WI, 53706 (anderson@math.wisc.edu).

‡Department of Mathematics and Statistics, University of Strathclyde, Glasgow, G1 1XH, UK (d.j.higham@strath.ac.uk).

§Centro de Matemática, Computação e Cognição, Universidade Federal do ABC, Santo André, SP 09210-580, Brazil (saul.leite@ufabc.edu.br).

¶Department of Mathematics, University of California - San Diego, La Jolla, CA, 92093-0112 (williams@math.ucsd.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

to average out the stochastic fluctuations. However, in some applications, such as in systems biology, not every molecular species is present in large numbers. The most common stochastic model of chemical kinetics treats the system as a continuous time Markov chain that tracks the (integer-valued) number of molecules of each chemical species [2, 17, 43]. These Markov chain models are often studied by sample path simulation in order to get Monte Carlo estimates for desired quantities [14, 15, 16]. However, these simulations can quickly become computationally expensive when some reactions are very fast, since every reaction is individually accounted for. When the abundances of the chemical species are large (but not large enough to ignore the influence of stochastic fluctuations) the units can be converted from abundances to concentrations and the solutions to the continuous time Markov chain model can be approximated by solutions of stochastic differential equations (SDEs). The resulting solutions are usually called a diffusion approximation; see, for example, [25]. These diffusion approximations can be simulated by numerical methods for SDEs, where a fixed time step can be set, yielding more efficient simulations in most cases. While the standard continuous time Markov chain model satisfies the natural condition that all abundances remain nonnegative for all time, diffusion approximations do not, in general, respect such a nonnegativity condition.

There are two commonly used diffusion approximations for the Markov chain model, the linear noise approximation [39, 40] and the chemical Langevin equation [18, 19, 25]. The linear noise approximation is obtained by linearizing fluctuations about the deterministic approximation. Although this approximation is well defined for all times, it typically diffuses outside of the positive orthant, predicting negative concentration values. In addition, it is well known [34, 41] that it can fail to capture fluctuations due to nonlinearities in the reaction rate functions. On the other hand, the chemical Langevin equation is known to give better approximations than the linear noise approximation when nonlinearities are present. However, the chemical Langevin equation is usually not defined beyond the first time the boundary of the orthant is reached. In fact, since the diffusion terms of the equation typically involve square roots of the molecular concentrations, the unstopped equation becomes ill posed [28, 36, 42].

For example, consider the following simple reaction system in which a molecule of S_1 can be converted to a molecule of S_2 and vice versa:

$$S_1 \overset{\beta_1}{\to} S_2, \quad S_2 \overset{\beta_2}{\to} S_1, \tag{1}$$

where $\beta_1, \beta_2 > 0$ are the rate constants and we assume the corresponding propensities follow mass action kinetics. When the number of S_1 molecules reaches zero in the Markov chain model, the reaction $S_1 \to S_2$ has zero intensity and cannot proceed until another S_1 molecule is created via the reaction $S_2 \to S_1$. In this manner nonnegativity of the number of S_1 molecules is preserved. Of course, a symmetric argument shows that the number of S_2 molecules remains nonnegative for all time. However, denoting the concentration of S_i at time t by $x_i(t)$, the usual chemical Langevin equation for this model consists of the system of SDEs,

$$dx_1(t) = (-\beta_1 x_1(t) + \beta_2 x_2(t))dt - \frac{1}{\sqrt{r}}\sqrt{\beta_1 x_1(t)}dW_1(t) + \frac{1}{\sqrt{r}}\sqrt{\beta_2 x_2(t)}dW_2(t), \tag{2}$$

$$dx_2(t) = (\beta_1 x_1(t) - \beta_2 x_2(t))dt + \frac{1}{\sqrt{r}}\sqrt{\beta_1 x_1(t)}dW_1(t) - \frac{1}{\sqrt{r}}\sqrt{\beta_2 x_2(t)}dW_2(t), \tag{3}$$

where W_1 and W_2 are independent Brownian motions, the equations are interpreted in the Itô sense, and r is usually taken to be Avogadro's number multiplied by the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

volume of the vessel in which the reactions are taking place. Whenever $x_1(t)$ is near zero,

$$\beta_2 x_2(t)dt + \frac{1}{\sqrt{r}}\sqrt{\beta_2 x_2(t)}dW_2(t) \tag{4}$$

is the dominant term in the right-hand side of (2). Because the term involving W_2 in (4) is as likely to push x_1 in the negative direction as in the positive direction, x_1 can become negative, thereby making (2) nonsensical in our context. Of course, a symmetric argument shows that $x_2(t)$ can become negative due to the stochastic forcing from W_1.

In this paper, we propose a *constrained Langevin approximation* for chemical reaction systems which is a reflected diffusion satisfying a nonnegativity constraint. In order to motivate the approximation, we begin with two simple, but natural, one-dimensional examples and then extend the approximation to the general multidimensional case. For the one-dimensional models, we also show how to compute stationary distributions for the approximation. It is worth noting that because the constrained Langevin approximation is developed via the same first principle arguments used in the development of the standard chemical Langevin equation, solutions to the two models satisfy the same dynamics within the strictly positive orthant. This fact is in contrast to other Langevin-type models developed to fix the negativity problem that perturb the dynamics globally to fix what is inherently a local (to the boundary) problem [42]. We emphasize that the derivation in this paper of the constrained Langevin approximation is only formal. The paper [27] is a rigorous technical complement to this paper. In [27], under mild conditions, the well posedness of the reflected diffusion is proved and it is shown that this diffusion process can be achieved as a weak limit of a sequence of jump-diffusion Markov processes that mimic the Langevin system in the interior of the positive orthant and behave like a scaled version of the Markov chain on the boundary.

In related work, several authors have devised approaches that combine the accuracy and robustness of the Markov chain model with the computational efficiency of the Langevin diffusion or ODE models; see, for example, [4, 10, 13, 20] and the references therein. For example, hybrid models have been proposed that exploit the existence of fast and slow reactions (determined either a priori or dynamically) [13, 20] or blend the jump and diffusion models, depending on the current system state [4, 10]. In a different vein, in [33] it has been proposed to extend the range of solutions for the Langevin equation to the complex numbers. The authors of [33] illustrate their approximation for some unimolecular and bimolecular examples. Although this state representation loses physical meaning, the authors show that this "complex Langevin equation" can be used to give real-valued approximations to moments and first passage times. Our work has a different focus than the references mentioned above. We operate entirely in the diffusion setting and introduce a general strategy to respect nonnegativity and well posedness. This permits simulation of sample paths and avoids the need for ad hoc thresholding, blending parameters or introducing additional state variables, and does not require specialized assumptions on the structure of the reaction system.

The rest of this paper is organized as follows. Section 1.1 gives a short description of the notation which will be used throughout the paper. In section 2, we present the continuous time Markov chain model for chemical reaction networks, beginning with its most common form in section 2.1, where the state represents the number of molecules of each species in the system. Next, in section 2.2, we introduce the scaled

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Markov chain model, where the state representation is converted from abundances to molecular concentrations. The constrained Langevin approximation is presented in section 3. We begin by introducing two motivating one-dimensional examples in section 3.1 and the ideas are then extended to the general multidimensional case in section 3.2. Section 4 is dedicated to numerical results, where we compare the constrained Langevin approximation with the Markov chain model and the linear noise approximation. We begin by comparing the stationary distributions of the one-dimensional examples in section 4.1 and, later, in section 4.2, we present the result of computer simulations for two-dimensional examples. For the benefit of the reader, Appendix A gives a brief summary of some aspects of reflected diffusion processes.

1.1. Notation. For any integer $m \geq 1$, let $\mathbb{Z}^m$ denote the integer lattice and $\mathbb{Z}^m_{\geq 0}$ denote the integer lattice of points with nonnegative components. Let $\mathbb{R}^m$ denote the m-dimensional Euclidean space and let $\mathbb{R}^m_{\geq 0}$ denote the positive orthant in $\mathbb{R}^m$ (i.e., the set of points of $\mathbb{R}^m$ whose components are all nonnegative). When $m = 1$, we write $\mathbb{Z}^1$, $\mathbb{Z}^1_{\geq 0}$, $\mathbb{R}^1$, and $\mathbb{R}^1_{\geq 0}$ as $\mathbb{Z}$, $\mathbb{Z}_{\geq 0}$, $\mathbb{R}$, and $\mathbb{R}_{\geq 0}$, respectively. For a vector $x \in \mathbb{R}^m$, we denote by x' its transpose and for a given set of vectors $\{x_1, \ldots, x_d\} \subset \mathbb{R}^m$, we denote by $\text{span}\{x_1, \ldots, x_d\}$ the set of all linear combinations of its elements. For sets A and B such that $A \subset B$, we denote by $1_A : B \to \mathbb{R}$ the indicator function, where $1_A(x)$ is defined to be 1 when $x \in A$ and 0 otherwise, for all $x \in B$.

2. Markov chain model of chemical reaction systems. We consider a chemical reaction system consisting of a finite set of species $\{S_1, S_2, \ldots, S_m\}$ involved in K possible reactions, where K is a positive integer. For $k \in \{1, \ldots, K\}$, we denote by v_k^- and v_k^+ the vectors in $\mathbb{Z}^m_{\geq 0}$ such that v_{ik}^- and v_{ik}^+ (the ith component of each) give the numbers of molecules of the ith species consumed and produced in the kth reaction, respectively. For example, if the kth reaction in a system consisting of just two species is $2S_1 \to S_2$, then $v_k^- = \left[\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right]$ and $v_k^+ = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$. We denote by $X(t)$ the vector in $\mathbb{Z}^m_{\geq 0}$ whose ith component gives the number of molecules of the ith species at time t. We note that occurrence of the kth reaction at a time t changes the state of the system by addition of the reaction vector $v_k = v_k^+ - v_k^-$; that is,

$$X(t) = X(t-) + v_k.$$

We assume that $v_k \neq 0$ for each $k = 1, \ldots, K$. In the next subsection, we describe the usual continuous time Markov chain model for such systems.

2.1. Continuous time Markov chain model. The standard stochastic model for a chemical system treats the system as a continuous time, discrete state Markov chain [17, 43]. For each reaction there is an associated real-valued function of the state, $\Lambda_k : \mathbb{Z}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$, called the *propensity* or *intensity* function, giving the rate at which the kth reaction occurs. Specifically, it is assumed that for each $k \in \{1, \ldots, K\}$, $x \in \mathbb{Z}^m_{\geq 0}$, and $t \geq 0$,

$$\begin{aligned} &P\{X(t+\Delta t) = x + v_k \mid X(t) = x\} = \Lambda_k(x)\Delta t + o(\Delta t), \\ &P\{X(t+\Delta t) = x \mid X(t) = x\} = \left(1 - \sum_{k=1}^{K} \Lambda_k(x)\Delta t\right) + o(\Delta t), \end{aligned} \tag{5}$$

where $o(\Delta t)/\Delta t \to 0$, as $\Delta t \to 0$. The usual assumption on the intensity functions Λ_k, and the assumption we make throughout, is that they satisfy *stochastic mass*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

action kinetics: for $x \in \mathbb{Z}^m_{\geq 0}$ the rate of the kth reaction is

$$\Lambda_k(x) = \kappa_k \prod_{i=1}^{m} (x_i)_{v_{ik}^-} \tag{6}$$

for some constant $\kappa_k > 0$, where

$$(x_i)_{v_{ik}^-} = x_i(x_i - 1)\dots(x_i - v_{ik}^- + 1) = \frac{x_i!}{(x_i - v_{ik}^-)!}. \tag{7}$$

The constant κ_k is called the *(stochastic) reaction rate constant.*

For example, zeroth order reactions of the form $\emptyset \to S_1$ have constant rate function $\Lambda_k(x) = \kappa_k$, first order reactions of the forms $S_1 \to S_2$ or $S_1 \to \emptyset$ have rate $\Lambda_k(x) = \kappa_k x_1$, and second order reactions of the forms $S_1 + S_2 \to S_3$ and $2S_1 \to S_3$ have respective rates $\Lambda_k(x) = \kappa_k x_1 x_2$ and $\Lambda_k(x) = \kappa_k x_1(x_1 - 1)$. Thus, the rate (6) is proportional to the number of distinct subsets of the molecules present that can form the inputs for the reaction. Intuitively, the mass action assumption reflects the idea that the system is *well stirred* in the sense that all molecules are equally likely to be at any location at any time.

There are different ways to represent the Markov chain model having the properties described in (5), however, we find the following endogenous representation for the Markov chain to be very useful. In this representation, the Markov chain $X(t)$ is given as the solution of the following equation:

$$X(t) = X(0) + \sum_{k=1}^{K} v_k N_k\bigg(\int_0^t \Lambda_k(X(s))ds\bigg), \tag{8}$$

where N_k, for $k \in \{1, \dots, K\}$, are independent unit-rate Poisson processes,[1] and, for each k, the time changed Poisson process $N_k(\int_0^t \Lambda_k(X(s))ds)$ represents the number of times the kth reaction has occurred by time t (for more on this representation see, for example, [2], [12, Chapter 6], or [26]).

2.2. Scaled system. We may convert from abundances to concentrations. To indicate the dependence of quantities such as X on the volume of the vessel in which the reactions are occurring, we let r denote the volume of the vessel multiplied by Avogadro's number and we append a superscript r to X (and other quantities that depend on r). Define $\overline{X}_i^r(\cdot) = \frac{1}{r} X_i^r(\cdot)$ for $i = 1, \dots, m$. Note that $\overline{X}_i^r(t)$ is simply the concentration of the ith species in moles per unit volume at time $t \geq 0$. As the units of the stochastic rate law are in *numbers* of molecules, but the units of $\overline{X}^r$ are moles per unit volume, to be able to write a sensible equation governing the dynamics of $\overline{X}^r$, the rates must also be scaled by r in an appropriate manner. The standard scaling (see, for example, [43, Chapter 6]) is the following: for zeroth order reactions, the stochastic rate constant κ_k^r is equal to rc_k for some $c_k > 0$ that does not depend upon r; for first order reactions, $\kappa_k^r = c_k$; and for second order reactions, $\kappa_k^r = c_k/r$. In general, for jth order reactions, $\kappa_k^r = c_k r^{-j+1}$.

Let Λ_k^r denote the propensity function for the kth reaction that is associated with the system indexed by r, when following stochastic mass action kinetics as in (6), with

[1] Recall, N_k is a unit-rate Poisson process if $N_k(0) = 0$, $N_k(\cdot)$ has independent increments, and $N_k(t+s) - N_k(s)$ has a Poisson distribution with parameter t for all $t, s \geq 0$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

rate constant κ_k^r satisfying the scaling detailed in the previous paragraph. Define

$$\lambda_k(x) = c_k \prod_{i=1}^{m} x_i^{v_{ik}^-}$$

for $x \in \mathbb{R}^m_{\geq 0}$, where we take $0^0 \equiv 1$. This λ_k is the reaction rate function associated with *deterministic mass action kinetics*, with reaction rate constant c_k. It is an exercise to check that for any reaction, i.e., zeroth order, first order, second order, etc.,

$$\Lambda_k^r(X^r(t)) = r\lambda_k(\overline{X}^r(t)) + \epsilon_k^r(\overline{X}^r(t)),$$

where $\epsilon_k^r(x)$ is a multivariate polynomial in the coordinates of x and $1/r$ that is uniformly bounded for all $r \geq 1$ as x varies in a compact set, and is nonzero only if the kth reaction consumes more than one molecule of a particular species. For example, for the second order reaction $S_1 + S_2 \to S_3$ we have

$$\Lambda_k^r(X^r(t)) = \frac{c_k}{r}\left(r\overline{X}_1^r(t)\right)\left(r\overline{X}_2^r(t)\right) = rc_k\overline{X}_1^r(t)\overline{X}_2^r(t) = r\lambda_k(\overline{X}^r(t)),$$

whereas for the second order reaction $2S_1 \to S_3$ we have

$$\begin{aligned}\Lambda_k^r(X^r(t)) &= \frac{c_k}{r} r\overline{X}_1^r(t)\left(r\overline{X}_1^r(t) - 1\right) = rc_k\overline{X}_1^r(t)^2 - c_k\overline{X}_1^r(t) \\ &= r\lambda_k(\overline{X}^r(t)) + \epsilon_k^r(\overline{X}^r(t))\end{aligned}$$

with $\epsilon_k^r(x) = -c_k x_1$.

After performing the above scaling from numbers of molecules to concentrations, and defining

$$\lambda_k^r(x) = \frac{\Lambda_k^r(rx)}{r} = \lambda_k(x) + (1/r)\epsilon_k^r(x), \tag{9}$$

we have that (8) yields

$$\overline{X}^r(t) = \overline{X}^r(0) + \frac{1}{r}\sum_{k=1}^{K} v_k N_k\left(r\int_0^t \lambda_k^r(\overline{X}^r(s))\,ds\right). \tag{10}$$

3. The constrained Langevin approximation. We start by presenting two examples that serve to motivate our constrained Langevin approximation for the model (10). We stress that, as there are no limit theorems given here, the arguments are meant to show the plausibility of the proposed approximation.

3.1. Motivating examples.

Example 1. Consider the chemical reaction system with the two reactions

$$\emptyset \xrightarrow{\alpha} S_1, \quad S_1 \xrightarrow{\beta} \emptyset.$$

The constants $\alpha > 0$ and $\beta > 0$ over the arrows here denote the (deterministic) reaction rate constants c_1 and c_2 for the two reactions. Then, with the scaling of propensities described in section 2.2, the Markov chain $X^r(\cdot)$, which models the stochastic dynamics of the number of molecules of S_1 when the volume measure is r, satisfies

$$X^r(t) = X^r(0) + N_1(r\alpha t) - N_2\left(\int_0^t \beta X^r(s)ds\right), \tag{11}$$

where N_1 and N_2 are independent, unit-rate Poisson processes.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Let $G = [0, \infty)$, $G^\circ = (0, \infty)$, the interior of G, and $G^b = \{0\}$, the boundary of G. We now give an equivalent in distribution representation of X^r, where we separately consider jumps of X^r from the interior of G and from the boundary of G. Specifically, an equivalent in distribution representation for X^r is

$$
\begin{aligned}
X^r(t) = X^r(0) &+ N_1^o\left(r\alpha\int_0^t 1_{\{X^r(s)\in G^\circ\}}ds\right) + N_1^b\left(r\alpha\int_0^t 1_{\{X^r(s)\in G^b\}}ds\right) \\
&- N_2^o\left(\int_0^t \beta X^r(s)1_{\{X^r(s)\in G^\circ\}}ds\right) - N_2^b\left(\int_0^t \beta X^r(s)1_{\{X^r(s)\in G^b\}}ds\right),
\end{aligned} \tag{12}
$$

where $N_1^o, N_1^b, N_2^o, N_2^b$ are independent unit-rate Poisson processes. (The distributional equivalence of the two solution processes in (11) and (12) can be understood informally via the superposition property of Poisson processes. See [3, Chapter 1] for a rigorous argument.)

Recall the definition from section 2.2 of $\overline{X}^r(\cdot) = \frac{1}{r}X^r(\cdot)$, the normalized (concentration-valued) process. For this example, the process satisfies

$$
\begin{aligned}
\overline{X}^r(t) = \overline{X}^r(0) &+ \frac{1}{r}N_1^o\left(r\alpha\int_0^t 1_{\{\overline{X}^r(s)\in G^\circ\}}ds\right) + \frac{1}{r}N_1^b\left(r\alpha\int_0^t 1_{\{\overline{X}^r(s)\in G^b\}}ds\right) \\
&- \frac{1}{r}N_2^o\left(r\int_0^t \beta\overline{X}^r(s)1_{\{\overline{X}^r(s)\in G^\circ\}}ds\right) - \frac{1}{r}N_2^b\left(r\int_0^t \beta\overline{X}^r(s)1_{\{\overline{X}^r(s)\in G^b\}}ds\right).
\end{aligned} \tag{13}
$$

We can center a unit-rate Poisson process N about its mean by defining $\widehat{N}(t) = N(t) - t$ for all $t \geq 0$. Centering N_1^o and N_2^o in this way, collecting terms in an obvious manner, and noting that the last term in (13) is identically equal to zero (since $x1_{\{x=0\}} \equiv 0$), we see that $\overline{X}^r$ satisfies

$$
\begin{aligned}
\overline{X}^r(t) = \overline{X}^r(0) &+ \alpha\int_0^t 1_{\{\overline{X}^r(s)\in G^\circ\}}ds - \int_0^t \beta\overline{X}^r(s)1_{\{\overline{X}^r(s)\in G^\circ\}}ds \\
&+ \frac{1}{\sqrt{r}}\Bigg[\frac{1}{\sqrt{r}}\widehat{N}_1^o\left(r\alpha\int_0^t 1_{\{\overline{X}^r(s)\in G^\circ\}}ds\right) \\
&\qquad - \frac{1}{\sqrt{r}}\widehat{N}_2^o\left(r\int_0^t \beta\overline{X}^r(s)1_{\{\overline{X}^r(s)\in G^\circ\}}ds\right) + \breve{Y}^r(t)\Bigg],
\end{aligned} \tag{14}
$$

where $\breve{Y}^r(t) = \frac{1}{\sqrt{r}}N_1^b(r\alpha\int_0^t 1_{\{\overline{X}^r(s)\in G^b\}}ds)$, $t \geq 0$, defines a nondecreasing jump process, which can only jump at times s for which $\overline{X}^r(s-)$ is at the boundary of G.

Our goal is to propose a diffusion process defined on $[0, \infty)$ that approximates $\overline{X}^r$ for fixed r of moderate size. By the functional central limit theorem for a centered, unit-rate Poisson process $\widehat{N}$, we have that $\frac{1}{\sqrt{r}}\widehat{N}(r\cdot)$ is well approximated in distribution by $\widehat{W}(\cdot)$, a standard one-dimensional Brownian motion.[2] This suggests that for an approximation, we replace $\frac{1}{\sqrt{r}}\widehat{N}_k^\circ(r\cdot)$ by $\widehat{W}_k^\circ(\cdot)$ for $k = 1, 2$, in (14), where $\widehat{W}_1^\circ$ and $\widehat{W}_2^\circ$ are independent standard one-dimensional Brownian motions. Inserting

[2]Indeed, one can even do this in a strong way. One can construct $N(\cdot)$ and $\widehat{W}(\cdot)$ on the same probability space so that $N(t) = t + \widehat{W}(t) + \xi(t)$ for all $t \geq 0$, where $\sup_{t\geq 0}\frac{|\xi(t)|}{\log(2\vee t)}$ is a random variable with a finite exponential moment (see [12, Corollary 5.5, p. 359] and [24]).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

these approximations in (14) and replacing $\overline{X}^r, \breve{Y}^r$ with Z^r, Y^r, respectively, leads us to propose approximating $\overline{X}^r$ by a jump-diffusion process Z^r satisfying

$$Z^r(t) = Z^r(0) + \int_0^t (\alpha - \beta Z^r(s)) 1_{\{Z^r(s)\in G^\circ\}} ds + \frac{1}{\sqrt{r}} \widehat{W}_1^\circ \left(\alpha \int_0^t 1_{\{Z^r(s)\in G^\circ\}} ds \right)$$
$$- \frac{1}{\sqrt{r}} \widehat{W}_2^\circ \left(\int_0^t \beta Z^r(s) 1_{\{Z^r(s)\in G^\circ\}} ds \right) + \frac{1}{\sqrt{r}} Y^r(t), \tag{15}$$

where $Y^r(t) = \frac{1}{\sqrt{r}} N_1^b(r\alpha \int_0^t 1_{\{Z^r(s)\in G^b\}} ds)$. By a martingale representation theorem (see, e.g., [23, Theorem 4.2, p. 170]), the difference of the two processes $\widehat{W}_1^\circ(\alpha \int_0^t 1_{\{Z^r(s)\in G^\circ\}} ds)$ and $\widehat{W}_2^\circ(\int_0^t \beta Z^r(s) 1_{\{Z^r(s)\in G^\circ\}} ds)$ (which are local martingales) can be represented as the single stochastic integral process

$$\int_0^t \sqrt{\alpha + \beta Z^r(s)} 1_{\{Z^r(s)\in G^\circ\}} dW(s), \tag{16}$$

where W is a standard one-dimensional Brownian motion.

Since we seek a diffusion approximation that moves continuously and spends zero time (in the sense of Lebesgue measure) at any particular point in $[0,\infty)$, it is reasonable to suppress the indicator functions in (15) and to replace the jump process Y^r by a continuous nondecreasing process that increases only when our diffusion process is on the boundary. This leads to a proposed reflected diffusion approximation $\widetilde{Z}^r$ for $\overline{X}^r$ that satisfies

$$\widetilde{Z}^r(t) = \widetilde{Z}^r(0) + \int_0^t (\alpha - \beta \widetilde{Z}^r(s)) ds + \frac{1}{\sqrt{r}} \int_0^t \sqrt{\alpha + \beta \widetilde{Z}^r(s)}\, d\widetilde{W}(s) + \frac{1}{\sqrt{r}} \widetilde{Y}^r(t), \tag{17}$$

where $\widetilde{W}$ is a standard one-dimensional Brownian motion, and where $\widetilde{Y}^r$ is a continuous, nondecreasing process that only increases when $\widetilde{Z}^r$ is zero. The process $\frac{1}{\sqrt{r}}\widetilde{Y}^r$ tracks the cumulative amount of pushing at the boundary required to keep $\widetilde{Z}^r$ nonnegative and is usually referred to as the *reflection* or *local time* term. Although the $\frac{1}{\sqrt{r}}$ scale factor could be absorbed into $\widetilde{Y}^r$, we keep it separate here to indicate that this reflection term is expected to be of the same order as the noise term, i.e., of order $\frac{1}{\sqrt{r}}$, to counter the excursions of the stochastic integral term involving $\widetilde{W}$ that try to drive $\widetilde{Z}^r$ negative. It is known[3] that given a pair $(\widetilde{Z}^r(0), \widetilde{W})$, there exists a unique solution $\widetilde{Z}^r$ to (17) that lives in $[0,\infty)$ and is adapted to $\widetilde{Z}^r(0)$ and $\widetilde{W}$. The process $\widetilde{Z}^r$ is a diffusion on $[0,\infty)$ with state dependent drift coefficient $x \mapsto \alpha - \beta x$, dispersion coefficient $x \mapsto \frac{1}{\sqrt{r}}\sqrt{\alpha + \beta x}$, and instantaneous reflection at the origin.

Notice that ignoring terms in (17) of order $\frac{1}{\sqrt{r}}$ leads to the usual deterministic approximation to the scaled model (13). The term $\frac{1}{\sqrt{r}} \int_0^t \sqrt{\alpha + \beta \widetilde{Z}^r(s)}\, d\widetilde{W}(s)$ captures stochastic fluctuations. The term $\frac{1}{\sqrt{r}}\widetilde{Y}^r$ only comes into play when $\widetilde{Z}^r$ is zero, and provides a minimal restoring force to keep $\widetilde{Z}^r$ nonnegative.

[3]This follows from the uniform Lipschitz property of the drift and dispersion coefficient, and the Lipschitz continuity of the so-called Skorokhod map that defines the reflection at the origin (i.e., determines $\widetilde{Y}^r$) in terms of the other parts of the equation. In this case with normal reflection at the boundary, a rigorous justification follows from the work of Tanaka [37].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Example 2. Consider the chemical reaction system given in (1). For $i \in \{1,2\}$, let $X_i^r(t)$ denote the number of molecules of S_i at time t. Let $M^r = X_1^r(0) + X_2^r(0)$, which is a conserved quantity. The process X_1^r can be represented as a solution to

$$X_1^r(t) = X_1^r(0) - N_1\left(\int_0^t \beta_1 X_1^r(s)ds\right) + N_2\left(\int_0^t \beta_2(M^r - X_1^r(s))ds\right), \tag{18}$$

where N_1 and N_2 are independent, unit-rate Poisson processes, and $X_2^r(\cdot) \equiv M^r - X_1^r(\cdot)$.

In a similar manner to that used in Example 1, let $G^r = [0, M^r]$, $G^{\circ,r} = (0, M^r)$, the interior of G^r, and $G^{b,r} = \{0, M^r\}$, the boundary of G^r. A distributionally equivalent way to represent X_1^r is as a solution of

$$\begin{aligned} X_1^r(t) = X_1^r(0) &- N_1^o\left(\int_0^t \beta_1 X_1^r(s)1_{\{X_1^r(s)\in G^{\circ,r}\}}ds\right) \\ &- N_1^b\left(\int_0^t \beta_1 X_1^r(s)1_{\{X_1^r(s)\in G^{b,r}\}}ds\right) \\ &+ N_2^o\left(\int_0^t \beta_2(M^r - X_1^r(s))1_{\{X_1^r(s)\in G^{\circ,r}\}}ds\right) \\ &+ N_2^b\left(\int_0^t \beta_2(M^r - X_1^r(s))1_{\{X_1^r(s)\in G^{b,r}\}}ds\right), \end{aligned} \tag{19}$$

where $N_1^o, N_1^b, N_2^o, N_2^b$ are independent unit-rate Poisson processes.

Now for the normalized process $\overline{X}_1^r(\cdot) = \frac{1}{r}X_1^r(\cdot)$ with $\overline{M}^r = \frac{1}{r}M^r$, the conserved quantity for the normalized process, we have that $\overline{X}_1^r$ satisfies

$$\begin{aligned} \overline{X}_1^r(t) = \overline{X}_1^r(0) &- \frac{1}{r}N_1^o\left(r\int_0^t \beta_1 \overline{X}_1^r(s)1_{\{\overline{X}_1^r(s)\in \widehat{G}^{\circ,r}\}}ds\right) \\ &- \frac{1}{r}N_1^b\left(r\int_0^t \beta_1 \overline{X}_1^r(s)1_{\{\overline{X}_1^r(s)\in \widehat{G}^{b,r}\}}ds\right) \\ &+ \frac{1}{r}N_2^o\left(r\int_0^t \beta_2(\overline{M}^r - \overline{X}_1^r(s))1_{\{\overline{X}_1^r(s)\in \widehat{G}^{\circ,r}\}}ds\right) \\ &+ \frac{1}{r}N_2^b\left(r\int_0^t \beta_2(\overline{M}^r - \overline{X}_1^r(s))1_{\{\overline{X}_1^r(s)\in \widehat{G}^{b,r}\}}ds\right), \end{aligned} \tag{20}$$

where the normalized interior of the state space is $\widehat{G}^{\circ,r} = (0, \overline{M}^r)$ and the normalized boundary is $\widehat{G}^{b,r} = \{0, \overline{M}^r\}$. Centering N_1^o and N_2^o and collecting terms, we see that $\overline{X}_1^r$ satisfies

$$\begin{aligned} \overline{X}_1^r(t) = \overline{X}_1^r(0) &- \int_0^t \beta_1 \overline{X}_1^r(s)1_{\{\overline{X}_1^r(s)\in \widehat{G}^{\circ,r}\}}ds + \int_0^t \beta_2(\overline{M}^r - \overline{X}_1^r(s))1_{\{\overline{X}_1^r(s)\in \widehat{G}^{\circ,r}\}}ds \\ &+ \frac{1}{\sqrt{r}}\Bigg[- \frac{1}{\sqrt{r}}\widehat{N}_1^o\left(r\int_0^t \beta_1 \overline{X}_1^r(s)1_{\{\overline{X}_1^r(s)\in \widehat{G}^{\circ,r}\}}ds\right) \\ &\quad + \frac{1}{\sqrt{r}}\widehat{N}_2^o\left(r\int_0^t \beta_2(\overline{M}^r - \overline{X}_1^r(s))1_{\{\overline{X}_1^r(s)\in \widehat{G}^{\circ,r}\}}ds\right) - \breve{Y}_1^r(t) + \breve{Y}_2^r(t)\Bigg], \end{aligned} \tag{21}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

where $\widehat{N}_1^o$ and $\widehat{N}_2^o$ are centered versions of the Poisson processes N_1^o and N_2^o, respectively, and for $t \geq 0$,

$$\breve{Y}_1^r(t) = \frac{1}{\sqrt{r}} N_1^b \left(r \int_0^t \beta_1 \overline{M}^r 1_{\{\overline{X}_1^r(s) = \overline{M}^r\}} ds \right), \quad \breve{Y}_2^r(t) = \frac{1}{\sqrt{r}} N_2^b \left(r \int_0^t \beta_2 \overline{M}^r 1_{\{\overline{X}_1^r(s) = 0\}} ds \right)$$

are nondecreasing jump processes that only jump at times s for which $\overline{X}^r(s-)$ equals $\overline{M}^r$ or 0, respectively. These processes push $\overline{X}_1^r$ back into $\widehat{G}^{\circ,r}$. Note that some boundary jump terms in the expression for $\overline{X}_1^r(t)$ have been eliminated here due to the facts that $x 1_{\{x=0\}} = 0$ and $(\overline{M}^r - x) 1_{\{x = \overline{M}^r\}} = 0$.

Proceeding in a similar manner to that for Example 1, we approximate $\frac{1}{\sqrt{r}} \widehat{N}_k^o(r\cdot)$ by $\widehat{W}_k^o(\cdot)$ for $k = 1, 2$, where $\widehat{W}_1^o$ and $\widehat{W}_2^o$ are independent standard one-dimensional Brownian motions. This leads us to propose approximating $\overline{X}_1^r$ by a jump-diffusion process Z_1^r that lives in $[0, \overline{M}^r]$ and satisfies

$$\begin{aligned} Z_1^r(t) = Z_1^r(0) &+ \int_0^t \left(-\beta_1 Z_1^r(s) + \beta_2 (\overline{M}^r - Z_1^r(s)) \right) 1_{\{Z_1^r(s) \in \widehat{G}^{\circ,r}\}} ds \\ &- \frac{1}{\sqrt{r}} W_1^\circ \left(\int_0^t \beta_1 Z_1^r(s) 1_{\{Z_1^r(s) \in \widehat{G}^{\circ,r}\}} \right) ds \\ &+ \frac{1}{\sqrt{r}} W_2^\circ \left(\int_0^t \beta_2 (\overline{M}^r - Z_1^r(s)) \, 1_{\{Z_1^r(s) \in \widehat{G}^{\circ,r}\}} ds \right) \\ &- \frac{1}{\sqrt{r}} Y_1^r(t) + \frac{1}{\sqrt{r}} Y_2^r(t), \end{aligned} \tag{22}$$

$$Y_1^r(t) = \frac{1}{\sqrt{r}} N_1^b \left(r \int_0^t \beta_1 \overline{M}^r 1_{\{Z_1^r(s) = \overline{M}^r\}} ds \right), \quad Y_2^r(t) = \frac{1}{\sqrt{r}} N_2^b \left(r \int_0^t \beta_2 \overline{M}^r 1_{\{Z_1^r(s) = 0\}} ds \right).$$

Then using a martingale representation theorem, the difference of the two processes $\widehat{W}_1^\circ(\int_0^t \beta_1 Z_1^r(s) 1_{\{Z_1^r(s) \in \widehat{G}^{\circ,r}\}} ds)$ and $\widehat{W}_2^o(\int_0^t \beta_2 (\overline{M}^r - Z_1^r(s)) 1_{\{Z_1^r(s) \in \widehat{G}^{\circ,r}\}} ds)$ (which are local martingales) can be represented as the single stochastic integral process

$$\int_0^t \sqrt{\beta_1 Z_1^r(s) + \beta_2 (\overline{M}^r - Z_1^r(s))} \, 1_{\{Z_1^r(s) \in \widehat{G}^{\circ,r}\}} dW(s), \tag{23}$$

where W is a standard one-dimensional Brownian motion.

As in Example 1, suppressing the indicator functions in (22) and replacing the jump processes Y_1^r, Y_2^r by continuous nondecreasing processes $\widetilde{Y}_1^r$, $\widetilde{Y}_2^r$ that increase only when $\widetilde{Z}_1^r$ is at $\overline{M}^r$ or 0, respectively, leads to a proposed reflected diffusion approximation $\widetilde{Z}_1^r$ for $\overline{X}_1^r$ that satisfies

$$\begin{aligned} \widetilde{Z}_1^r(t) = \widetilde{Z}_1^r(0) &+ \int_0^t (-\beta_1 \widetilde{Z}_1^r(s) + \beta_2 (\overline{M}^r - \widetilde{Z}_1^r(s))) ds \\ &+ \frac{1}{\sqrt{r}} \int_0^t \sqrt{\beta_1 \widetilde{Z}_1^r(s) + \beta_2 (\overline{M}^r - \widetilde{Z}_1^r(s))} \, d\widetilde{W}(s) \\ &- \frac{1}{\sqrt{r}} \widetilde{Y}_1^r(t) + \frac{1}{\sqrt{r}} \widetilde{Y}_2^r(t), \end{aligned} \tag{24}$$

where $\widetilde{W}$ is a standard, one-dimensional Brownian motion, and $\widetilde{Y}_1^r$,$\widetilde{Y}_2^r$ are continuous, nondecreasing processes that only increase when $\widetilde{Z}_1^r$ is at $\overline{M}^r$ or 0, respectively. It is

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

known[4] that, given the pair $(\widetilde{Z}_1^r(0), \widetilde{W})$, there exists a solution to (24) that lives in $[0, \overline{M}^r]$ and is adapted to $\widetilde{Z}_1^r(0)$ and $\widetilde{W}$.

Notice again that ignoring terms in (24) of order $\frac{1}{\sqrt{r}}$ leads to the usual deterministic approximation to the scaled model (19). The term

$$\frac{1}{\sqrt{r}} \int_0^t \sqrt{\beta_1 \widetilde{Z}_1^r(s) + \beta_2(\overline{M}^r - \widetilde{Z}_1^r(s))}\, d\widetilde{W}(s)$$

captures stochastic fluctuations. The terms $\frac{1}{\sqrt{r}}\widetilde{Y}_1^r$ and $\frac{1}{\sqrt{r}}\widetilde{Y}_2^r$ only come into play when $\widetilde{Z}_1^r$ is on the boundary of the domain of interest and they provide a minimal restoring force to keep $\widetilde{Z}_1^r$ in that domain.

In preparation for the general case, it is convenient to rewrite (24) as

$$\begin{aligned}\widetilde{Z}_1^r(t) = \widetilde{Z}_1^r(0) &+ \int_0^t \mu^r(\widetilde{Z}_1^r(s))\, ds \\ &+ \frac{1}{\sqrt{r}} \int_0^t \sigma^r(\widetilde{Z}_1^r(s))\, d\widetilde{W}(s) + \frac{1}{\sqrt{r}} \int_0^t n^r(\widetilde{Z}_1^r(s))\, d\widetilde{L}^r(s),\end{aligned}$$

where $\mu^r(x) = -\beta_1 x + \beta_2(\overline{M}^r - x)$, $\sigma^r(x) = \sqrt{\beta_1 x + \beta_2(\overline{M}^r - x)}$, $\widetilde{L}^r(t) = \widetilde{Y}_1^r(t) + \widetilde{Y}_2^r(t)$ is continuous and nondecreasing and increases only when $\widetilde{Z}_1^r$ is on the boundary of $\widehat{G}^{\circ,r}$, and $n^r(x) = 1$ if $x = 0$ and $n^r(x) = -1$ if $x = \overline{M}^r$ is the inward unit normal to the boundary of $\widehat{G}^{\circ,r}$. The vector field n^r specifies the "direction" of reflection at the boundary. In this one-dimensional case, there is a unique (up to normalization) inward pointing direction (which is normal to the boundary). In the general case treated in the next subsection, more complicated boundary behavior occurs and the reflection direction is frequently not normal to the boundary.

3.2. The general case. We now consider the general case of a process $\overline{X}^r$ satisfying (10). Let $\mathcal{S} = \text{span}\{v_k, k = 1, \ldots, K\}$ and let $\widehat{G}^r = (\overline{X}^r(0) + \mathcal{S}) \cap \mathbb{R}^m_{\geq 0}$. Considering $\widehat{G}^r$ in $\overline{X}^r(0) + \mathcal{S}$, let $\widehat{G}^{\circ,r}$ and $\widehat{G}^{b,r}$ denote the relative interior and boundary of $\widehat{G}^r$, respectively. Proceeding in a similar manner to that for the two examples given in the previous subsection, an equivalent in distribution representation for $\overline{X}^r$ is given by

$$\begin{aligned}\overline{X}^r(t) = \overline{X}^r(0) &+ \frac{1}{r}\sum_{k=1}^K v_k N_k^o \left(r \int_0^t \lambda_k^r(\overline{X}^r(s)) 1_{\{\overline{X}^r(s) \in \widehat{G}^{\circ,r}\}} ds \right) \\ &+ \frac{1}{r}\sum_{k=1}^K v_k N_k^b \left(r \int_0^t \lambda_k^r(\overline{X}^r(s)) 1_{\{\overline{X}^r(s) \in \widehat{G}^{b,r}\}} ds \right),\end{aligned} \tag{25}$$

where N_k^o, N_k^b, $k = 1, \ldots, K$, are independent unit-rate Poisson processes. Upon centering the Poisson processes N_k^o, $k = 1, \ldots, K$, to obtain $\widehat{N}_k^o, k = 1, \ldots, K$, we

[4]Similarly to what was mentioned with respect to the solution of (17), this also follows from the work of Tanaka [37] by the uniform Lipschitz continuity of the drift, dispersion coefficient, and the Skorokhod map that defines the reflection, since the direction of reflection in this case is normal to the boundary.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

may rewrite the above as

$$\begin{aligned}\overline{X}^r(t) &= \overline{X}^r(0) + \sum_{k=1}^K v_k \int_0^t \lambda_k^r(\overline{X}^r(s))\, 1_{\{\overline{X}^r(s)\in\widehat{G}^{\circ,r}\}}\, ds \\ &\quad + \frac{1}{\sqrt{r}}\sum_{k=1}^K v_k \frac{1}{\sqrt{r}}\widehat{N}_k^o\left(r\int_0^t \lambda_k^r(\overline{X}^r(s))1_{\{\overline{X}^r(s)\in\widehat{G}^{\circ,r}\}}ds\right) \\ &\quad + \frac{1}{\sqrt{r}}\sum_{k=1}^K v_k \breve{Y}_k^r(t),\end{aligned} \tag{26}$$

where $\breve{Y}_k^r(t) = \frac{1}{\sqrt{r}}N_k^b(r\int_0^t \lambda_k^r(\overline{X}^r(s))1_{\{\overline{X}^r(s)\in\widehat{G}^{b,r}\}}ds)$.

Noting the relation (9), it is natural to replace λ_k^r with λ_k to obtain approximate dynamics for $\overline{X}^r$. Also, as in the examples in the previous section, we can approximate $\frac{1}{\sqrt{r}}\widehat{N}_k^o(r\cdot)$ by $\widehat{W}_k^o(\cdot)$ for $k = 1,\dots,K$, where $\widehat{W}_k^o$, $k = 1,\dots,K$, are independent standard one-dimensional Brownian motions. Let

$$\mu(x) = \sum_{k=1}^K v_k\lambda_k(x) \tag{27}$$

for each $x \in \mathbb{R}^m_{\geq 0}$. We suggest approximating $\overline{X}^r$ by a jump diffusion Z^r satisfying $Z^r(0) = \overline{X}^r(0)$ and

$$\begin{aligned}Z^r(t) &= Z^r(0) + \int_0^t \mu(Z^r(s))\, 1_{\{Z^r(s)\in\widehat{G}^{\circ,r}\}}\, ds \\ &\quad + \frac{1}{\sqrt{r}}\sum_{k=1}^K v_k \widehat{W}_k^\circ\left(\int_0^t \lambda_k(Z^r(s))1_{\{Z^r(s)\in\widehat{G}^{\circ,r}\}}ds\right) \\ &\quad + \frac{1}{\sqrt{r}}\sum_{k=1}^K v_k Y_k^r(t),\end{aligned} \tag{28}$$

where $Y_k^r(t) = \frac{1}{\sqrt{r}}N_k^b(r\int_0^t \lambda_k(Z^r(s))1_{\{Z^r(s)\in\widehat{G}^{b,r},\, Z^r(s)\geq v_k^-/r\}}ds)$. Note that in the definition of Y_k^r, the indicator function suppresses jumps from the boundary that would require consumption of a given species when there is an insufficient amount of that species to make the transition possible. This is a small correction needed to account for the fact that the interior diffusion might occasionally bring Z^r to a point x on the boundary of $\widehat{G}^r$ that cannot be reached by the discrete-valued process $\overline{X}^r$ (which lives on a lattice), and where $\lambda_k(x) > 0$, and from which movement by Z^r along the vector $-v_k^-$ would take Z^r outside of the positive orthant. Such occurrences are only possible when more than one component of Z^r is small, that is, whenever the process Z^r is near the intersection of two or more faces of the positive orthant. It is known that such occurrences are rare for some similar reflected diffusion processes.[5] Consequently, we anticipate this correction will likely be a relatively small one.

Using a martingale representation theorem (see, e.g., [23, Theorem 4.2, p. 170]), on a possibly enlarged probability space that accommodates a standard m-dimensional

[5]For some similar nondegenerate reflected diffusion processes (see [31, Theorem 1] and [6, Theorem 7.7]), it is known that the total amount of "pushing" done by the local-time term at the intersection of two or more boundary faces of the positive orthant is almost surely zero. Such local-time terms are approximate measures of the amount of time spent near boundary regions.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Brownian motion W, we can express $\sum_{k=1}^{K} v_k \widehat{W}_k^o(\int_0^t \lambda_k(Z^r(s))\, 1_{\{Z^r(s)\in\widehat{G}^{\circ,r}\}}\, ds)$ as the vector-valued stochastic integral process $\int_0^t \sigma(Z^r(s)) 1_{\{Z^r(s)\in\widehat{G}^{\circ,r}\}}\, dW(s)$, where W is a standard m-dimensional Brownian motion, $\sigma(x) = \sqrt{\Upsilon(x)}$ is the unique[6] positive semidefinite matrix-valued square root of the $m \times m$ matrix

$$\Upsilon(x) = \sum_{k=1}^{K} v_k v_k' \lambda_k(x) \tag{29}$$

for $x \in \mathbb{R}^m_{\geq 0}$, and v_k' is the transpose of v_k.

Remark. The reader may note that the coefficients μ and σ in the above do not depend on r. Example 1 in the previous subsection illustrates this, whereas our Example 2 has coefficients that appear to depend on r. However, the latter dependence occurs because, in that example, we have eliminated one of the variables, effectively projecting down to the concentration of S_1 alone. Indeed, if we had written the approximation $(\widetilde{Z}_1^r, \widetilde{Z}_2^r)$ for the concentrations of both species $(\overline{X}_1^r, \overline{X}_2^r)$ satisfying $\widetilde{Z}_1^r + \widetilde{Z}_2^r = \overline{M}^r$, then in the notation of this subsection, $\mu(x_1, x_2) = (-\beta_1 x_1 + \beta_2 x_2)v$, where $v = (1, -1)'$, and $\sigma(x_1, x_2) = \sqrt{\frac{\beta_1 x_1 + \beta_2 x_2}{2}} vv'$, which leads to an equivalent representation to that given for $\widetilde{Z}_1^r$ in (24).

To obtain an equation for our proposed diffusion approximation, we remove the indicator functions in the first two terms in (28), since the amount of time our diffusion approximation spends on the boundary is zero (in the sense of Lebesgue measure). Furthermore, we want to replace the last term in (28), the boundary term, by a continuous process whose paths are locally of bounded variation and that only changes when the diffusion process is on the boundary $\widehat{G}^{b,r}$. In the examples in the previous section, the diffusion process was one dimensional and there was a unique (up to normalization) direction at each boundary point in which the boundary process would push to keep the diffusion in the state space. In higher dimensions, there is much more freedom in the choice of such a direction. In the following we motivate our choice for this "reflection direction" in the general case.

In our reflected diffusion approximation, the role of the boundary term is to counteract fluctuations of the term driven by white noise that tends to take the diffusion process outside of the orthant. Since the fluctuations are of order $\frac{1}{\sqrt{r}}$, we expect this boundary term to be of order $\frac{1}{\sqrt{r}}$. This leads us to approximate $N_k^b(\cdot)$ in Y_k^r by its deterministic rate process, and to ignore higher order terms, resulting in the following (formal) approximation:

$$\begin{aligned} \sum_{k=1}^{K} v_k Y_k^r(t) &\approx \sum_{k=1}^{K} \frac{v_k}{\sqrt{r}} \int_0^t r\lambda_k(Z^r(s)) 1_{\{Z^r(s)\in\widehat{G}^{b,r},\, Z^r(s)\geq v_k^-/r\}} ds \\ &\approx \int_0^t \gamma(Z^r(s)) dL^r(s), \end{aligned} \tag{30}$$

where

$$\gamma(x) = \frac{\mu(x)}{|\mu(x)|} 1_{\{|\mu(x)|\neq 0\}} \quad \text{for } x \in \widehat{G}^{b,r}, \tag{31}$$

[6]The existence and uniqueness of a (symmetric) positive semidefinite square root for any (symmetric) positive semidefinite matrix is well known. Furthermore, the mapping from the matrix $\Upsilon(x)$ to its square root $\sigma(x)$ is Hölder continuous of order one-half. These results can be found in the book by Bhatia [7], for example.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

for $\mu(x)$ given by (27), and where

$$L^r(t) = \sqrt{r} \int_0^t |\mu(Z^r(s))| 1_{\{Z^r(s) \in \widehat{G}^{b,r}\}} ds. \tag{32}$$

Note here that we have approximated the indicator function in (30) with the simpler indicator function in (32) (ignoring the rare effect mentioned after (28)).

The process L^r is a weighted and scaled version of the amount of time that Z^r spends on the boundary. In our diffusion approximation, we approximate this by a continuous nondecreasing process $\widetilde{L}^r$ that can increase only when our diffusion process is on the boundary $\widehat{G}^{b,r}$. Indeed, in [27], under certain conditions, a more extensive rationale is given for approximating L^r by $\widetilde{L}^r$. This involves showing that if the jump size $\delta = \frac{1}{\sqrt{r}}$ in Y_k^r is sent to zero and at the same time the order of magnitude of the speed of jumping, $\delta^{-2} = r$, is sent to infinity, while keeping the other r dependencies fixed, then L^r converges (weakly) to the process $\widetilde{L}^r$.

This leads us to propose the following equation for our diffusion approximation $\widetilde{Z}^r$ for $\overline{X}^r$:

$$\widetilde{Z}^r(t) = \widetilde{Z}^r(0) + \int_0^t \mu(\widetilde{Z}^r(s))\, ds + \frac{1}{\sqrt{r}} \int_0^t \sigma(\widetilde{Z}^r(s))\, dW(s) + \frac{1}{\sqrt{r}} \int_0^t \gamma(\widetilde{Z}^r(s))\, d\widetilde{L}^r(s), \tag{33}$$

where $\widetilde{Z}^r$ is a continuous process living in $\widehat{G}^r$ and $\widetilde{L}^r$ is a continuous, one-dimensional, increasing process that starts from zero and that can only increase when $\widetilde{Z}^r$ is on the boundary $\widehat{G}^{b,r}$ of $\widehat{G}^r$. The vector field γ defines the "reflection" vector field on the boundary for the process $\widetilde{Z}^r$. This is the direction in which $\widetilde{Z}^r$ is "pushed" to keep it in the set $\widehat{G}^r$. The process $\widetilde{L}^r$ is the cumulative amount of "pushing" done at the boundary. For more detail on reflected diffusion processes, see Appendix A.

In [27], Leite and Williams prove well posedness of (33), under the assumption that the reaction network satisfies a mass-conserving (or mass-dissipating) assumption, augmented by inflows and outflows on all species. The latter means that the reactions

$$\emptyset \longrightarrow S_i, \quad S_i \longrightarrow \emptyset, \tag{34}$$

are part of the set of reactions for each $i = 1, \ldots, m$.[7] Systems without some of these inflow/outflow reactions can be approximated by including such reactions with very small rate constants c_k, so that the reactions rarely occur. If one does not make this assumption, issues regarding existence and uniqueness of the diffusion process can arise. These are related to the fact that σ might only be Hölder continuous near the boundary in some places, the vector field γ on the boundary might degenerate to either become zero or not point strictly into the interior of the state space $\widehat{G}^r$ at some places on $\widehat{G}^{b,r}$. The mass-conserving/mass-dissipating assumption, in combination with outflows on all species, is used to ensure nonexplosion of the diffusion process. These assumptions can sometimes be relaxed, especially when $\widehat{G}^r$ is one dimensional (or effectively so, as in Example 2), and in some cases in higher dimensions, if one can show that problematical boundary regions are not reached by the diffusion and

[7] This assumption ensures that σ is strictly positive definite and locally Lipschitz continuous on $\widehat{G}^r = \mathbb{R}^m_{\geq 0}$, that γ never vanishes on $\widehat{G}^{b,r}$, and it points strictly into the interior of $\widehat{G}^r$. As shown in [27], the conditions given there are sufficient to prove existence, uniqueness, and nonexplosion of the diffusion process $\widetilde{Z}^r$.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

there is no explosion in finite time. However, a systematic treatment of these matters requires new developments for the theory of reflected diffusions in polyhedral domains with degenerate dispersion coefficients and reflection vector fields. Nevertheless, we conjecture that a process $\tilde{Z}^r$ satisfying (33) will be a good approximation to $\overline{X}^r$, whenever the former is well defined. In the next section, we give examples that illustrate how well our diffusion approximation works, despite the informal nature of our derivation. Further examples can be found in [27].[8]

A problem for further investigation is to develop estimates of the error between $\overline{X}^r$ and $\widetilde{Z}^r$, assuming the latter is well defined. While this can be done when $\widetilde{Z}^r$ is one dimensional, a systematic treatment of this is a promising area for future investigation.

4. Examples. We begin this section by showing how the stationary distributions for the constrained Langevin approximation can be computed for the two examples given in section 3.1. The results are then compared with the stationary distributions for the Markov chain model and for the linear noise approximation. Later, in section 4.2, we further illustrate the constrained Langevin approximation by comparing its simulation for some two-dimensional examples with those for the Markov chain model, the linear noise approximation, and the complex Langevin approximation introduced in [33].

4.1. Stationary distributions. The approximations proposed in section 3 are for stochastic processes over compact time intervals. In this subsection we look for insights into the long time behavior by considering stationary distributions for some examples where analytical expressions are available.

Example 1 (revisited).We begin by noting that for a fixed $r > 0$, the stationary distribution of the (scaled) jump model (13) satisfies [1]:

$$\pi(x) = e^{-r\alpha/\beta}\frac{(r\alpha/\beta)^{rx}}{(rx)!}, \quad x \in \left\{0, \frac{1}{r}, \frac{2}{r}, \dots\right\}. \tag{35}$$

Turning to our constrained Langevin approximation, by (17) and Itô's formula [9] for $f \in C_c^2([0,\infty))$ (two times continuously differentiable functions with compact support),

$$\begin{aligned} &f(\widetilde{Z}^r(t)) - f(\widetilde{Z}^r(0)) \\ &\quad = \int_0^t (\alpha - \beta\widetilde{Z}^r(s))f'(\widetilde{Z}^r(s))ds + \frac{1}{\sqrt{r}}\int_0^t \sqrt{\alpha + \beta\widetilde{Z}^r(s)}f'(\widetilde{Z}^r(s))\,d\widetilde{W}(s) \\ &\qquad + \frac{1}{2r}\int_0^t (\alpha + \beta\widetilde{Z}^r(s))f''(\widetilde{Z}^r(s))ds + \frac{1}{\sqrt{r}}\int_0^t f'(\widetilde{Z}^r(s))d\widetilde{Y}^r(s). \end{aligned}$$

Suppose now that $f'(0) = 0$. Then the last term is zero because $\widetilde{Y}^r$ can only increase when $\widetilde{Z}^r$ is at zero. The integral with respect to $d\widetilde{W}$ defines a martingale and so, taking expectations when $\widetilde{Z}^r(0)$ has the stationary distribution π with density ρ, we obtain

$$0 = \int_0^t E_\pi\left[\mathcal{L}f(\widetilde{Z}^r(s))\right]ds \quad \text{for all } t \geq 0,$$

[8] It is also shown in [27], under mild conditions, that a sequence of jump-diffusion processes, in which the jumps at the boundary are allowed to shrink to zero at the same time that the rate of jumping goes to infinity, converges weakly to a solution of (33).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

where

$$(\mathcal{L}f)(x) = (\alpha - \beta x)f'(x) + \frac{1}{2r}(\alpha + \beta x)f''(x)$$
$$= \mu(x)f'(x) + \frac{1}{2r}\Upsilon(x)f''(x)$$

for $\mu(x) = \alpha - \beta x$ and $\Upsilon(x) = \alpha + \beta x$. Hence ρ must satisfy

$$\int_0^\infty (\mathcal{L}f)(x)\rho(x)dx = 0, \qquad \int_0^\infty \rho(x)dx = 1. \tag{36}$$

Integration by parts yields

$$\int_0^\infty (\mathcal{L}f)(x)\rho(x)dx = \int_0^\infty f(x)\left(\frac{1}{2r}\frac{d^2}{dx^2}\left[\Upsilon(x)\rho(x)\right] - \frac{d}{dx}\left[\mu(x)\rho(x)\right]\right)dx$$
$$- f(0)\left(\mu(0)\rho(0) - \frac{1}{2r}\frac{d}{dx}(\Upsilon(x)\rho(x))\big|_{x=0}\right),$$

where we have used the facts that $f'(0) = 0$ and f has compact support in the above calculation. As the above must hold for all $f \in C_c^2([0,\infty))$ with $f'(0) = 0$, we see that ρ must satisfy

$$(\mathcal{L}^*\rho)(x) = -\frac{d}{dx}(\mu(x)\rho(x)) + \frac{1}{2r}\frac{d^2}{dx^2}(\Upsilon(x)\rho(x)) = 0 \quad \text{for all } x \in (0,\infty), \tag{37}$$

where $\mathcal{L}^*$ denotes the adjoint of $\mathcal{L}$, with the boundary condition

$$\left(\mu\rho - \frac{1}{2r}\frac{d}{dx}(\Upsilon\rho)\right)\Big|_{x=0} = 0. \tag{38}$$

Integrating (37) shows that

$$-\mu(x)\rho(x) + \frac{1}{2r}\frac{d}{dx}(\Upsilon(x)\rho(x)) = 0 \quad \text{for all } x \geq 0, \tag{39}$$

where the value of zero on the right-hand side follows from the boundary condition (38). Solving (39), noting that $\Upsilon(x) > 0$ for all $x \geq 0$, yields

$$\rho(x) = \frac{c}{\Upsilon(x)}\exp\left\{\int_0^x \frac{2r\mu(s)}{\Upsilon(s)}ds\right\} \quad \text{for } x \geq 0, \tag{40}$$

where c is a suitable normalizing constant. After substituting for our specific μ and Υ, we obtain

$$\rho(x) = ce^{-2xr}\left(\alpha + \beta x\right)^{(4r\alpha/\beta)-1} \quad \text{for } x \geq 0, \tag{41}$$

where

$$c = \left(\int_0^\infty e^{-2xr}\left(\alpha + \beta x\right)^{(4r\alpha/\beta)-1}dx\right)^{-1} \tag{42}$$

is the normalizing constant.

The linear noise approximation [40] for the Markov chain $\overline{X}_1^r$ has as its stationary distribution, ρ_{LN}, the steady-state distribution for the Ornstein–Uhlenbeck-type process $\hat{Z}_1$ that describes the linearized fluctuations of $\overline{X}_1^r$ about $\bar{x} = \frac{\alpha}{\beta}$, the (stable) steady state for the deterministic reaction rate equation approximation to $\overline{X}_1^r$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

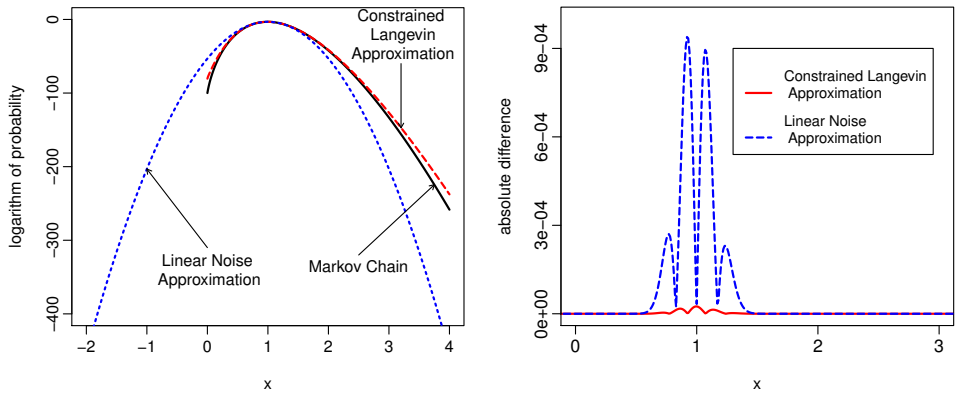


FIG. 1. *(Left) Comparison of* $\log(\pi(x))$, $\log(\pi_{CLA}(x))$, *and* $\log(\pi_{LN}(x))$, *where* $\log(\cdot)$ *denotes the natural logarithm, for Example* 1 *with* $\alpha = \beta = 1$ *and* $r = 100$. *(Right) Absolute difference between the stationary distribution for the Markov chain and those given by the constrained Langevin and linear noise approximations (i.e.,* $|\pi_{CLA}(x) - \pi(x)|$ *and* $|\pi_{LN}(x) - \pi(x)|$*). Values of* x *are taken from the lattice* $(1/r)\mathbb{Z}$ *and linear interpolation is used to connect the values.*

satisfying $\mu(\bar{x}) = 0$. This process $\hat{Z}_1^r$ satisfies

$$\hat{Z}_1^r(t) = \bar{x} + \int_0^t \mu'(\bar{x})(\hat{Z}_1^r(s) - \bar{x})\, ds + \frac{1}{\sqrt{r}} \int_0^t \sigma(\bar{x})\, d\hat{W}(s), \tag{43}$$

where $\mu'(\bar{x}) = -\beta$, $\sigma(\bar{x}) = \sqrt{\Upsilon(\bar{x})} = \sqrt{\alpha + \beta\bar{x}} = \sqrt{2\alpha}$, and $\hat{W}$ is a standard one-dimensional Brownian motion. The stationary distribution for $\hat{Z}_1^r$ is the Gaussian distribution with mean $\frac{\alpha}{\beta}$ and variance $\frac{\alpha}{r\beta}$ [40], so that

$$\rho_{LN}(x) = \sqrt{\frac{r\beta}{2\pi\alpha}} \exp\left(-\frac{r\beta(x - \frac{\alpha}{\beta})^2}{2\alpha}\right), \quad x \in (-\infty, \infty). \tag{44}$$

We now wish to compare the probability mass function in (35) with the densities in (41) and (44). Notice that the probability that a continuous model with strictly positive density function f takes a value in the interval $[x - 1/(2r), x + 1/(2r))$ can be well approximated by $f(x) \cdot r^{-1}$. So in order to compare the density for the stationary distribution of the linear noise approximation with the stationary distribution of the Markov chain, we define $\pi_{LN}(x) = \rho_{LN}(x) \cdot r^{-1}$ for x in the lattice $(1/r)\mathbb{Z}$, with ρ_{LN} as in (44). For the constrained Langevin approximation, we define the discretization $\pi_{CLA}(x) = \rho(x) \cdot r^{-1}$ for $x \in \{1/r, 2/r, \ldots\}$ with ρ as in (41). Since the density of the constrained Langevin approximation has no mass for $x < 0$, we let $\pi_{CLA}(0) = \rho(0) \cdot r^{-1}/2$, which is an approximation of the probability that the model takes a value in the interval $[0, 1/(2r))$. The result is shown in Figure 1 for the system with parameters $\alpha = \beta = 1$ and $r = 100$. Note that π_{LN} is defined for all x, whereas π and π_{CLA} are only defined for $x \geq 0$.

Example 2 (revisited). For fixed $r > 0$, the stationary distribution of the (scaled) Markov chain jump model (20) satisfies

$$\pi(x) = \pi(0) \left(\frac{\beta_2}{\beta_1}\right)^{rx} \frac{(M^r + 1 - rx)(M^r + 2 - rx) \cdots M^r}{(rx)!} \quad \text{for } x \in \left\{\frac{1}{r}, \frac{2}{r}, \ldots, \overline{M}^r\right\}, \tag{45}$$

where $\pi(0)$ is determined so that $\sum_{x=0}^{\overline{M}^r} \pi(x) = 1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Turning to our constrained Langevin approximation, the density function ρ of the stationary distribution for $\widetilde{Z}_1^r$, satisfying (24), is supported on $[0, \overline{M}^r]$ and, similarly to how (36) was derived, must satisfy the following two conditions [22]:

$$\int_0^{\overline{M}^r} (\mathcal{L}f)(x)\rho(x)dx = 0, \qquad \int_0^{\overline{M}^r} \rho(x)dx = 1$$

for all $f \in C^2([0, \overline{M}^r])$ satisfying $f'(0) = f'(\overline{M}^r) = 0$, where

$$\begin{aligned}(\mathcal{L}f)(x) &= \left(\beta_2\overline{M}^r - (\beta_1+\beta_2)x\right) f'(x) + \frac{1}{2r}\left(\beta_2\overline{M}^r + (\beta_1-\beta_2)x\right) f''(x)\\ &= \mu(x)f'(x) + \frac{1}{2r}\Upsilon(x)f''(x)\end{aligned}$$

for $\mu(x) = \beta_2\overline{M}^r - (\beta_1+\beta_2)x$ and $\Upsilon(x) = \beta_2\overline{M}^r + (\beta_1-\beta_2)x$. Here, to simplify notation, we have suppressed the explicit dependence of μ and Υ on r (which occurs through $\overline{M}^r$). Integration by parts gives

$$\begin{aligned}\int_0^{\overline{M}^r} (\mathcal{L}f)(x)\rho(x)dx = &\int_0^{\overline{M}^r} f(x)\left(\frac{1}{2r}\frac{d^2}{dx^2}\left[\Upsilon(x)\rho(x)\right] - \frac{d}{dx}\left[\mu(x)\rho(x)\right]\right)dx\\ &+ \left[f\left(\mu\rho - \frac{1}{2r}\frac{d}{dx}(\Upsilon\rho)\right)\right]_{x=0}^{\overline{M}^r}.\end{aligned}$$

Therefore, as the above must hold for all such f, we must have

$$(\mathcal{L}^*\rho)(x) = -\frac{d}{dx}(\mu(x)\rho(x)) + \frac{1}{2r}\frac{d^2}{dx^2}(\Upsilon(x)\rho(x)) = 0 \quad \text{for all } x \in (0, \overline{M}^r), \tag{46}$$

where $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$, and

$$\left(\mu\rho - \frac{1}{2r}\frac{d}{dx}(\Upsilon\rho)\right)\Big|_{x=0} = \left(\mu\rho - \frac{1}{2r}\frac{d}{dx}(\Upsilon\rho)\right)\Big|_{x=\overline{M}^r} = 0. \tag{47}$$

Integrating (46) gives

$$-\mu(x)\rho(x) + \frac{1}{2r}\frac{d}{dx}(\Upsilon(x)\rho(x)) = 0 \quad \text{for all } x \in [0, \overline{M}^r], \tag{48}$$

where the value of zero on the right-hand side follows from the boundary conditions (47). Solving (48) yields a solution of the form (40), which after substituting for our specific μ and Υ becomes

(49)

$$\rho(x) = \begin{cases} c\exp\left\{-\frac{2(\beta_1+\beta_2)}{\beta_1-\beta_2}rx\right\}\left(\beta_2\overline{M}^r + (\beta_1-\beta_2)x\right)^{-1+\frac{4\overline{M}^r r\beta_1\beta_2}{(\beta_1-\beta_2)^2}} & \text{if } \beta_1 \neq \beta_2,\\ c\exp\left\{-\frac{(x-\frac{1}{2}\overline{M}^r)^2}{\overline{M}^r/2r}\right\} & \text{if } \beta_1 = \beta_2\end{cases}$$

for $0 \le x \le \overline{M}^r$, where c is the normalizing constant chosen so that $\int_0^{\overline{M}^r}\rho(x)dx = 1$. Note that in the case of $\beta_1 = \beta_2$, the stationary distribution is a Gaussian distribution restricted to $[0, \overline{M}^r]$.

The linear noise approximation [40] for the Markov chain $\overline{X}_1^r$ in this example has as its stationary distribution, ρ_{LN}, the steady-state distribution for the Ornstein–Uhlenbeck-type process $\hat{Z}_1^r$ that describes the linearized fluctuations of $\overline{X}_1^r$ about

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

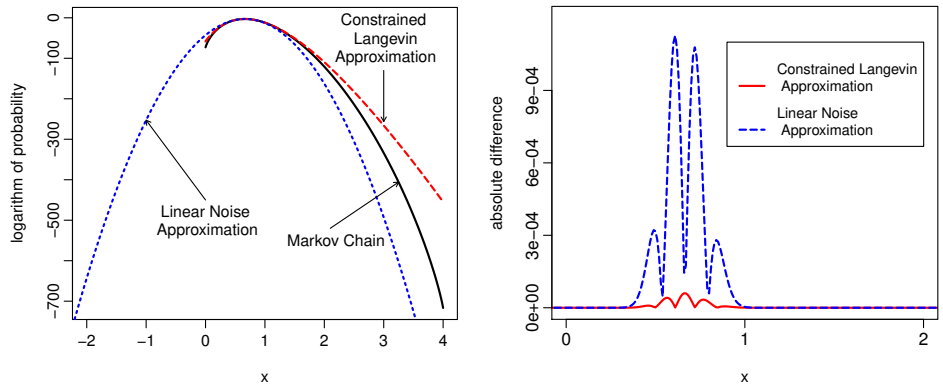


FIG. 2. *(Left) Comparison of* $\log(\pi(x))$, $\log(\pi_{CLA}(x))$, *and* $\log(\pi_{LN}(x))$, *where* $\log(\cdot)$ *denotes the natural logarithm, for Example* 2 *with* $\beta_1 = 5$ *and* $\beta_2 = 1$, $M^r = 400$ *and* $r = 100$. *(Right) Absolute difference between the stationary distribution for the Markov chain and those given by the constrained Langevin and linear noise approximations (i.e.,* $|\pi_{CLA}(x) - \pi(x)|$ *and* $|\pi_{LN}(x) - \pi(x)|$*). Values of* x *are taken from the lattice* $(1/r)\mathbb{Z}$ *and linear interpolation is used to connect the values.*

$\bar{x} = \frac{\beta_2 \overline{M}^r}{\beta_1+\beta_2}$, the (stable) steady-state for the deterministic reaction rate equation approximation to $\overline{X}_1^r$ satisfying $\mu(\bar{x}) = 0$. This process $\hat{Z}_1^r$ satisfies

$$\hat{Z}_1^r(t) = \bar{x} + \int_0^t \mu'(\bar{x})(\hat{Z}_1^r(s) - \bar{x})\, ds + \frac{1}{\sqrt{r}} \int_0^t \sigma(\bar{x})\, d\hat{W}(s), \tag{50}$$

where $\mu'(\bar{x}) = -(\beta_1 + \beta_2)$, $\sigma(\bar{x}) = \sqrt{\Upsilon(\bar{x})} = \sqrt{\beta_2 \overline{M}^r + (\beta_1 - \beta_2)\bar{x}} = \sqrt{\frac{2\beta_1\beta_2\overline{M}^r}{\beta_1+\beta_2}}$, and $\hat{W}$ is a standard one-dimensional Brownian motion. The stationary distribution for $\hat{Z}_1^r$ is the Gaussian distribution with mean $\bar{x} = \frac{\beta_2\overline{M}^r}{\beta_1+\beta_2}$ and variance $\frac{\Upsilon(\bar{x})}{2r|\mu'(\bar{x})|} = \frac{\beta_1\beta_2\overline{M}^r}{r(\beta_1+\beta_2)^2}$ [40], so that

$$\rho_{LN}(x) = \sqrt{\frac{r}{2\pi\beta_1\beta_2\overline{M}^r}}(\beta_1+\beta_2) \exp\left(-\frac{r(\beta_1+\beta_2)^2\left(x - \frac{\beta_2\overline{M}^r}{\beta_1+\beta_2}\right)^2}{2\beta_1\beta_2\overline{M}^r}\right), \quad x \in (-\infty, \infty). \tag{51}$$

In a similar manner to that for the previous example, we want to compare the probability mass function in (45) with the densities in (49) and (51). Again, in order to compare the density of the stationary distribution of the linear noise approximation with the stationary distribution of the Markov chain, we define $\pi_{LN}(x) = \rho_{LN}(x) \cdot r^{-1}$ for x in the lattice $(1/r)\mathbb{Z}$, with ρ_{LN} as in (51). For the constrained Langevin approximation (CLA), we define the discretization $\pi_{CLA}(x) = \rho(x) \cdot r^{-1}$ for $x \in \{1/r, 2/r, \ldots, \overline{M}^r - 1/r\}$, with ρ as in (49) and, since the density of the CLA has no mass for $x < 0$ or for $x > \overline{M}^r$, we let $\pi_{CLA}(0) = \rho(0) \cdot r^{-1}/2$ and $\pi_{CLA}(\overline{M}^r) = \rho(\overline{M}^r) \cdot r^{-1}/2$, which is an approximation of the probability that the model takes a value in the interval $[0, 1/(2r))$ and $[\overline{M}^r - 1/(2r), \overline{M}^r]$, respectively. The result is shown in Figure 2 for the system with parameters $\beta_1 = 5$, $\beta_2 = 1$, $M^r = 400$, and $r = 100$. Note that π_{LN} is defined for all x, whereas π and π_{CLA} are only defined for $0 \le x \le \overline{M}^r$.

For the cases illustrated in Figures 1 and 2, we see that, in addition to having the correct support, the stationary distribution for the CLA captures the behavior of the Markov chain model more accurately than the linear noise approximation.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

4.2. Simulation examples.

Example 3. We now consider a chemical reaction system involving two molecular species S_1 and S_2 with the following set of reactions:

$$2S_1 + S_2 \xrightarrow{\nu} 3S_1, \quad S_1 \xrightarrow{\beta_1} \emptyset, \quad S_2 \xrightarrow{\beta_2} \emptyset, \quad \emptyset \xrightarrow{\alpha_1} S_1, \quad \emptyset \xrightarrow{\alpha_2} S_2. \tag{52}$$

This reaction set is a simple mass-action kinetic system whose reaction rate equation (deterministic model) exhibits a limit cycle [32]. For this example, the CLA, given by (33), has drift vector and diffusion matrix given by

$$\mu(x)=\begin{pmatrix} \nu x_1^2 x_2 - \beta_1 x_1 + \alpha_1 \\ -\nu x_1^2 x_2 - \beta_2 x_2 + \alpha_2 \end{pmatrix}, \quad \Upsilon(x)=\begin{pmatrix} \nu x_1^2 x_2 + \beta_1 x_1 + \alpha_1 & -\nu x_1^2 x_2 \\ -\nu x_1^2 x_2 & \nu x_1^2 x_2 + \beta_2 x_2 + \alpha_2 \end{pmatrix},$$

and the reflection vector field is given by

$$\gamma(x) = \begin{cases} \begin{pmatrix} \alpha_1 \\ \alpha_2 - \beta_2 x_2 \end{pmatrix} \Big/ \sqrt{\alpha_1^2 + (\alpha_2 - \beta_2 x_2)^2} & \text{for } x_1 = 0, \\ \begin{pmatrix} \alpha_1 - \beta_1 x_1 \\ \alpha_2 \end{pmatrix} \Big/ \sqrt{(\alpha_1 - \beta_1 x_1)^2 + \alpha_2^2} & \text{for } x_2 = 0. \end{cases} \tag{53}$$

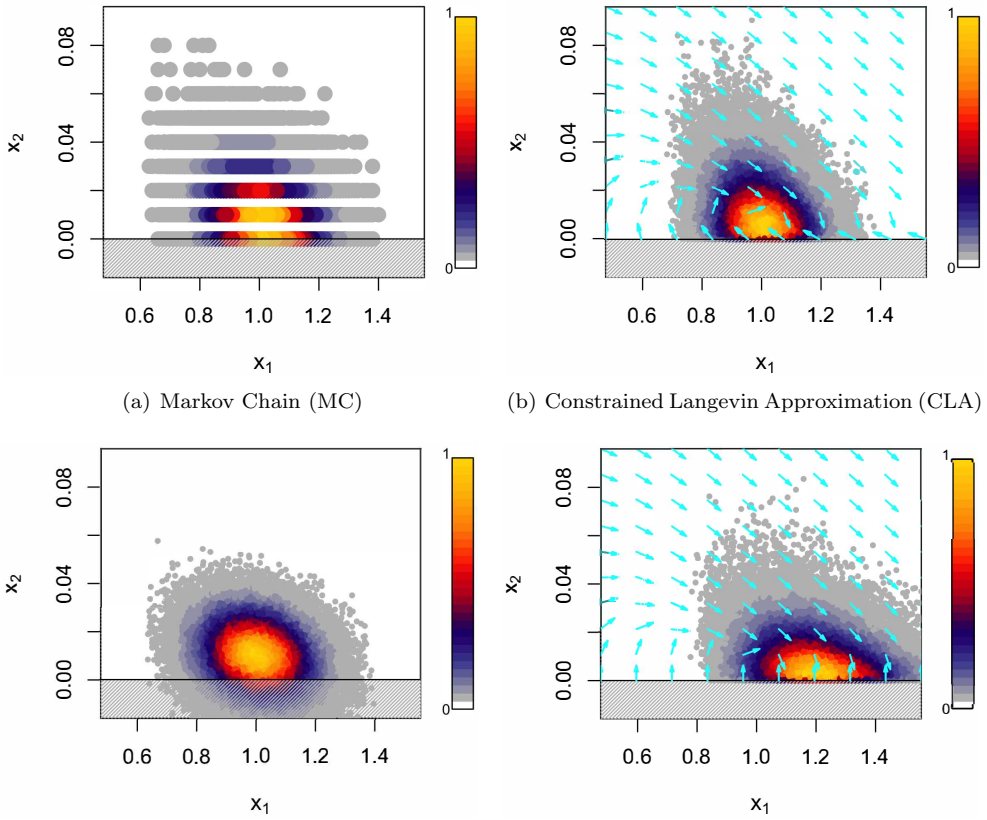
We compare the CLA with the Markov chain (MC) model, the linear noise approximation (LNA), and a Langevin equation with normal reflection at the boundaries (LEN). That is, LEN is given as in (33) with the exception that the direction of reflection is normal to the boundary. The MC model was simulated using Gillespie's algorithm [16]. For the LNA, we used a fourth order Runge–Kutta method for the deterministic part and the Euler–Maruyama method for the stochastic diffusion. For the CLA and LEN, we used Bossy, Gobet, and Talay's algorithm [8], which is a numerical method for simulating obliquely reflected SDEs based on the Euler–Maruyama method. The simulation codes were written in the R programming language [30].

We consider two sets of parameters for this reaction system. First, we set $r = 100$, $\nu = 10$, $\beta_1 = 0.2$, $\beta_2 = 10^{-9}$, and $\alpha_1 = \alpha_2 = 0.1$. For this choice, the reaction rate equation does not exhibit a limit cycle, but it spends most of its time near the boundary $x_2 = 0$. The time step for the numerical methods used for the diffusion approximations and the deterministic reaction rate equation was set to $h = 0.01$ and the simulations were performed up to time $T = 10^4$. The simulations were initialized at the stationary point for the deterministic model. Figure 3 shows the scatter plot of the points generated by the simulations. For the CLA and the LEN, we also display the reflection directions at the boundary and the drift vector field inside the state space, which is normalized to have unit length to improve the display. Notice that the LNA permits negative concentrations. In addition, LEN produces a shift to the right due to the effect of the normal reflection directions. Such a shift is not seen in the CLA simulation, since the reflection on $x_2 = 0$ is oblique, pointing towards the left of the plot.

In order to have a more precise measure of error, we estimate a discrete density for each of the simulations. This density estimation is calculated by dividing the state space into a regular grid of square bins and counting the number of simulation points present in each of these bins. These square bins have a side length of $1/r$ and are centered around each state of the MC model. The total number of points inside each square bin is normalized by the total number of points in the simulation and the area of the square. In order to measure statistical variation among different simulation runs, the experiment was repeated 10 times, using the same data from the previous

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms



(a) Markov Chain (MC)

(b) Constrained Langevin Approximation (CLA)

(c) Linear Noise Approximation (LNA)

(d) Langevin with Normal Reflection (LEN)

FIG. 3. *Scatter plot of the concentrations of S_1 and S_2, given by x_1 and x_2, respectively, generated from simulations of the MC, CLA, LNA, and LEN for the system given by* (52) *with parameters $r = 100$, $\nu = 10$, $\beta_1 = 0.2$, $\beta_2 = 10^{-9}$, and $\alpha_1 = \alpha_2 = 0.1$. For the MC, the plotted points were magnified in order to improve the display, since the values are within the lattice $(1/r)\mathbb{Z}^2$. For* (b) *CLA and* (d) *LEN we also show the directions of reflection at the boundary $x_2 = 0$ and the normalized drift vector field inside the state space. Notice that the distribution of points for* (d) *LEN is shifted to the right from the other plots, due to the normal reflection, and that* (c) *LNA permits negative concentrations for S_2.*

TABLE 1

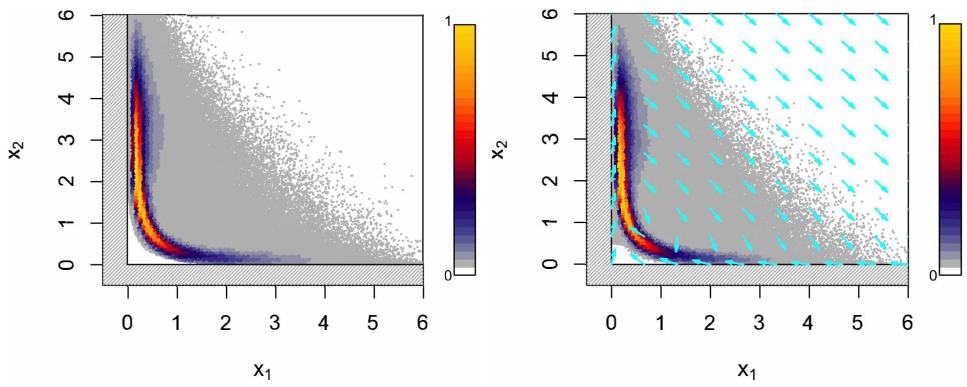
Integral of the absolute difference between the discrete densities calculated for the approximation processes and that for the MC simulation (for parameters as in Figure 3*). The values displayed are averages of* 10 *independent runs. The* 95% *confidence intervals for these averages are also shown.*

	Integral abs. diff.	95% C. I.
Linear Noise Approximation (LNA)	0.2762	(0.2716, 0.2808)
Constrained Langevin Approximation (CLA)	0.2432	(0.2373, 0.2490)
Langevin with Normal Reflection (LEN)	1.1710	(1.1529, 1.1892)

paragraph. Table 1 shows the integral (with respect to Lebesgue measure) of the absolute difference between the discrete density calculated for the MC simulation and those generated by the simulation of each approximation. Notice that the maximum possible value for these integrals is 2.

For the second set of parameters, we let $r = 100$, $\nu = 1$, $\beta_1 = 1$, $\beta_2 = 10^{-9}$, $\alpha_1 = 0.1$, and $\alpha_2 = 0.5$. This time, the reaction rate equation for this system exhibits

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) Markov Chain (MC) (b) Constrained Langevin Approximation (CLA)

FIG. 4. *Scatter plot of the concentrations of S_1 and S_2, given by x_1 and x_2, respectively, generated from simulations of the MC and CLA for the system given by* (52) *with parameters $r = 100$, $\nu = 1$, $\beta_1 = 1$, $\beta_2 = 10^{-9}$, $\alpha_1 = 0.1$, and $\alpha_2 = 0.5$. For* (b)*, for the CLA, we show the reflection directions at the boundaries used by the approximation and the normalized drift vector field. The paths for the LNA and LEN generated by simulation diverge from the limit cycle and are not shown here.*

a limit cycle which lives near the boundaries $x_1 = 0$ and $x_2 = 0$. For this example, we used a time step of $h = 0.1$ for the numerical methods, simulations were performed up to time $T = 10^5$, and the initial condition was set to $x^0 = (1, 1)$. The scatter plots for the simulations of MC and CLA are shown in Figure 4. The paths generated by the simulations of LNA and LEN grow without bound. For this reason, the scatter plots for these simulations are not shown. For the LNA, this divergence occurs despite the fact that the deterministic part of the model exhibits a stable limit cycle. This type of behavior of the LNA is known [34, 38, 41]. Although some corrective measures have been proposed for similar examples (see [29] and references therein), this illustrates the inability of LNA to characterize nonlinear behavior adequately. From Figure 5, we see that the path generated by the simulation of the LEN becomes unstable after it hits a reflection from the boundary $x_2 = 0$, which is normal and pushes the process towards higher concentrations of x_1, as can be seen from the vector field of Figure 4(b). This path is reflected again (perpendicularly) from the $x_2 = 0$ boundary, making the concentration of x_1 increasingly larger, which leads to numerical instability and divergence for the chosen step size.

Similarly to the previous example, a discrete density estimation was calculated for the CLA (it was not possible to perform this calculation for the LEN and the LNA since these simulations diverge). The integral of the absolute difference between the discrete density calculated for the CLA and that for the MC was given by 0.3058 with a 95% confidence interval of $(0.3044, 0.3072)$ among the 10 independent runs.

Example 4. Now we consider a different example in order to compare the CLA proposed here with the complex Langevin approximation introduced in [33]. The examples in [33] involve unimolecular and bimolecular reactions. Here, we consider the following example involving such reactions for species S_1 and S_2:

$$S_1 + S_2 \xrightarrow{\nu} 2S_2, \quad S_1 \xrightarrow{\beta_1} \emptyset, \quad S_2 \xrightarrow{\beta_2} \emptyset, \quad \emptyset \xrightarrow{\alpha_1} S_1, \quad \emptyset \xrightarrow{\alpha_2} S_2. \tag{54}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

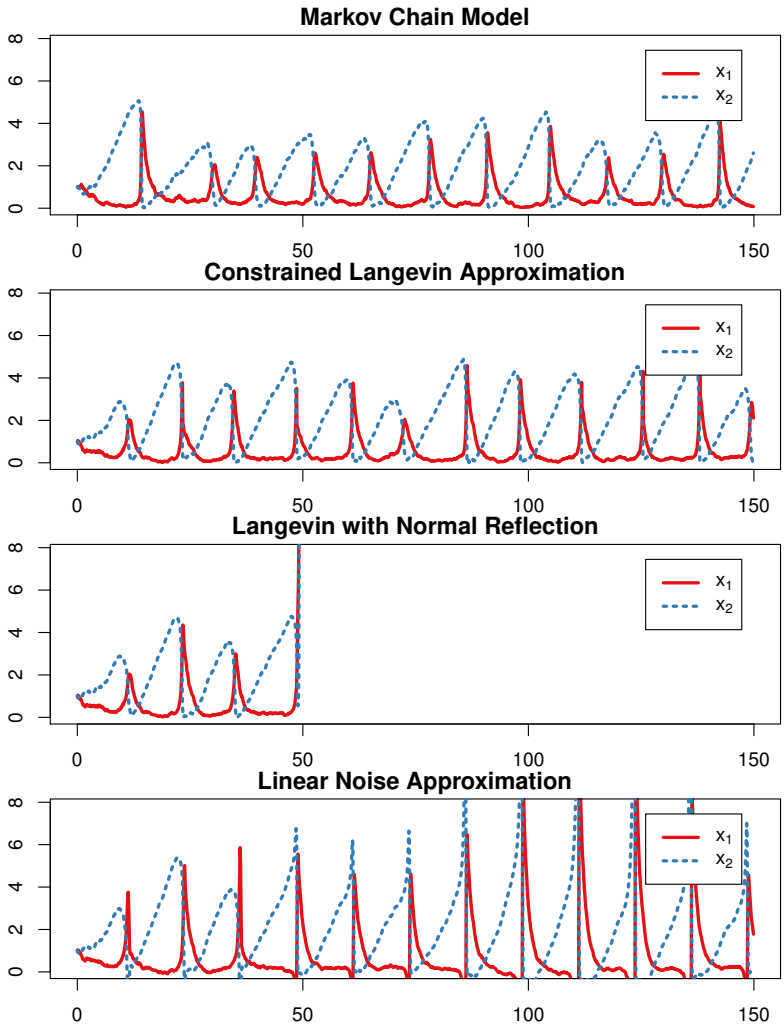


FIG. 5. *Plot of the concentrations of S_1 and of S_2 given by x_1 and x_2, respectively, versus time. The paths were generated from simulations for the system given by* (52) *with parameters $r = 100$, $\nu = 1$, $\beta_1 = 1$, $\beta_2 = 10^{-9}$, $\alpha_1 = 0.1$, and $\alpha_2 = 0.5$. Here we see that the LNA is unstable from early on in the simulation and that the LEN hits an unstable path near $t = 50$, where the normal reflection at the level $x_2 = 0$ pushes the process to take a path with increasingly higher concentrations of x_1, leading the simulation to diverge.*

The drift vector and diffusion matrix of the CLA for this example are given by

$$\mu(x)=\begin{pmatrix} -\nu x_1 x_2 - \beta_1 x_1 + \alpha_1 \\ \nu x_1 x_2 - \beta_2 x_2 + \alpha_2 \end{pmatrix}, \quad \Upsilon(x)=\begin{pmatrix} \nu x_1 x_2 + \beta_1 x_1 + \alpha_1 & -\nu x_1 x_2 \\ -\nu x_1 x_2 & \nu x_1 x_2 + \beta_2 x_2 + \alpha_2 \end{pmatrix}$$

with the reflection vector field at the boundary given by (53), since on the boundaries $x_1 = 0$ and $x_2 = 0$, the reflection vector field for (52) is the same as for (54).

Simulations were performed for the MC model, CLA, the LNA, and the complex Langevin approximation (complex-LE). We used the Euler–Maruyama method for the complex-LE, similarly to what was done in [33], and we used Bossy, Gobet, and Talay's method [8] for the CLA. For the LNA, we used a fourth order Runge–Kutta method for its deterministic part and the Euler–Maruyama method for its diffusion

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

TABLE 2

Means, variances and covariance for the concentration of each molecular species calculated by Markov chain (MC) simulation, the constrained Langevin approximations (CLA), the linear noise approximation (LNA), and the complex Langevin approximation (Complex-LE) (associated with the parameters as in Figure 6). The concentrations of species S_1 and S_2 are represented by x_1 and x_2, respectively. The table displays the average values among 10 independent runs and also the value that should be added to/subtracted from the mean to get the 95% confidence intervals.

	MC	CLA	LNA	Complex-LE
$E[x_1]$	3.854e-02 ±3.3e-05	3.898e-02 ±3.1e-05	3.836e-02 ±5.0e-05	3.835e-02 ±3.3e-05
$E[x_2]$	1.040e+00 ±2.7e-04	1.040e+00 ±3.4e-04	1.040e+00 ±2.9e-04	1.040e+00 ±2.7e-04
$\mathrm{var}(x_1)$	3.839e-04 ±1.1e-06	4.214e-04 ±1.6e-06	4.408e-04 ±1.1e-06	4.413e-04 ±2.2e-06
$\mathrm{var}(x_2)$	1.080e-02 ±3.7e-05	1.070e-02 ±3.9e-05	1.082e-02 ±3.4e-05	1.086e-02 ±3.8e-05
$\mathrm{cov}(x_1, x_2)$	−1.582e-05 ±5.3e-06	−1.924e-05 ±3.4e-06	−1.603e-05 ±4.3e-06	−1.857e-05 ±3.7e-06

part. The simulations were performed up to time $T = 10^5$ and the time steps for the numerical methods were set to $h = 0.01$. The initial condition used for the simulations was set to $x_0 = (1, 1)/r$ with $r = 100$, and the samples from the simulations were collected after an initial time of duration one was completed.

Since the complex-LE predicts real-valued moments, we compare the approximations by calculating mean concentration values, their variances, and covariance. In order to account for statistical variation among different runs and calculate confidence intervals, we repeated the simulations 10 times. Table 2 gives the results for the parameters $r = 100, \nu = \alpha_1 = \alpha_2 = \beta_2 = 1$, and $\beta_1 = 25$. The predicted moments are fairly closely matched for all simulations. In Figure 6, we give the scatter plot for the simulations. For the complex-LE, only the real parts of the simulation points are shown. Although these simulation points appear to be similarly distributed to those in the MC simulation, like the LNA, the real part of the complex Langevin approximation permits values outside of the positive orthant.

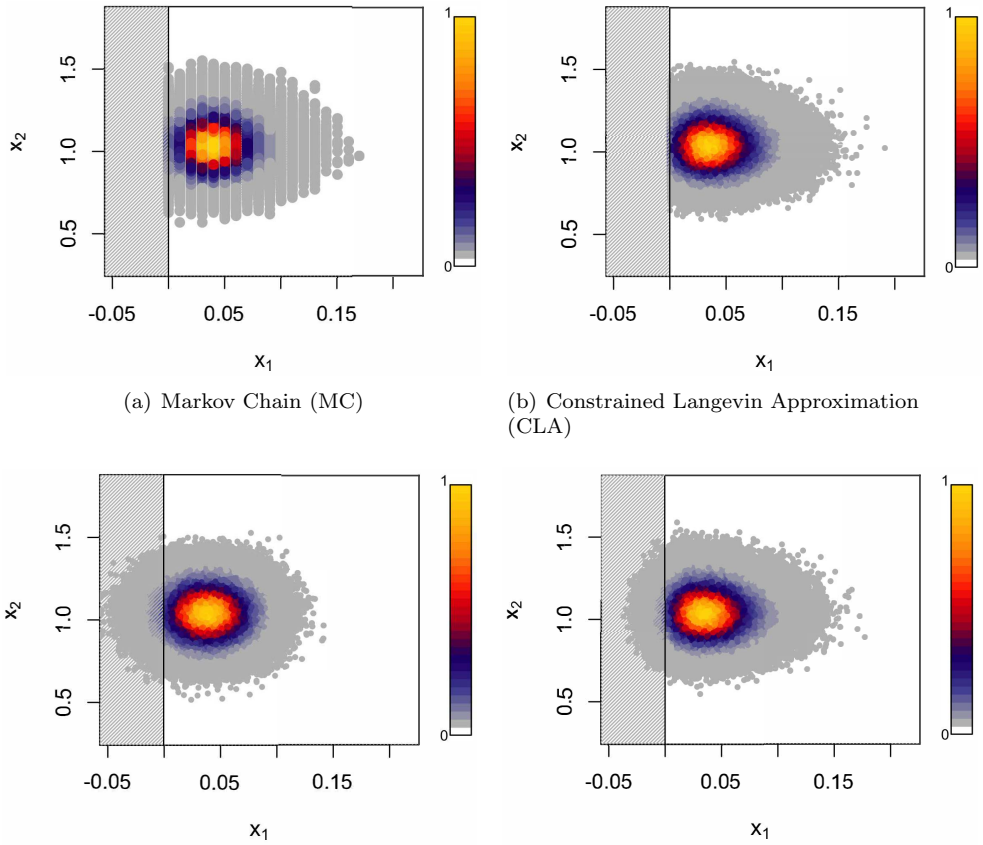
Since the real part of the complex-LE can take values outside of the positive orthant, its behavior depends on the values of the drift and dispersion coefficients there. We found that, for some examples, the drift vector field used by the complex-LE outside of the positive orthant can lead the process to have paths that diverge. One such example is found by considering the same example (54) with the following set of parameters: $r = 100$, $\alpha_1 = \beta_1 = 1$, $\nu = 10$, $\alpha_2 = 0.02$, and $\beta_2 = 5$. For this example, the simulation of the complex-LE using the Euler–Maruyama method diverges even with step sizes as small as $h = 0.001$. In order to illustrate this, we simulated the complex-LE with the time step $h = 0.001$ up to time $T = 10^5$. The simulation hits a divergent path and stops at time $t = 6665.421$. Figure 7 shows the evolution of the real parts of the complex Langevin approximation for the molecular concentrations of species S_1 and S_2, represented by variables x_1 and x_2, respectively, prior to divergence. From this figure, we observe that the process has crossed the $x_2 = 0$ boundary, where the drift vector field pushes the process to higher concentrations of x_1 and negative values of x_2.

The same experiment was repeated for the MC, the CLA, and the LNA, using the time step of $h = 0.01$ for the numerical methods. The scatter plot for a simulation with duration $T = 10^5$ is given in Figure 8. We also calculated the integral of the absolute difference between the discrete density calculated for the MC and those calculated for the CLA and the LNA. The result is given in Table 3.

5. Summary and discussion. It is attractive, both analytically and computationally, to approximate a continuous time, discrete state space Markov chain (MC) by a continuous time, continuous state diffusion process. From a modeling perspective

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms



(a) Markov Chain (MC)

(b) Constrained Langevin Approximation (CLA)

(c) Linear Noise Approximation (LNA)

(d) Complex Langevin Approximation (Complex-LE)

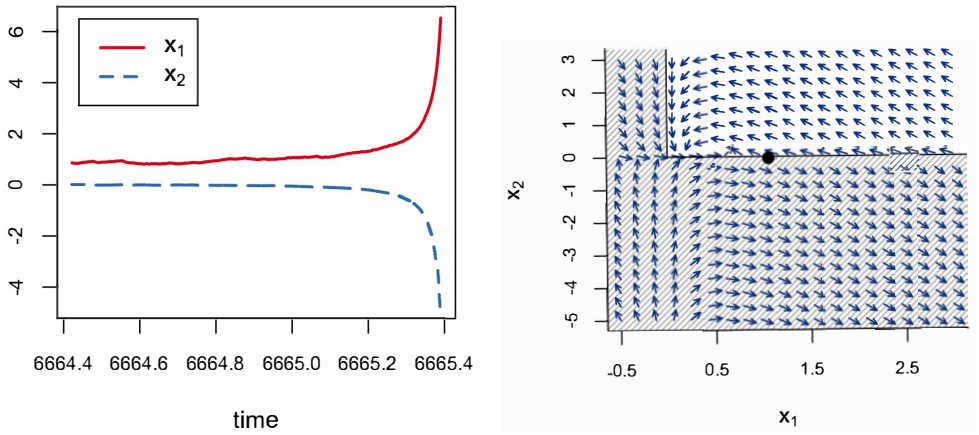
FIG. 6. *Scatter plot of the concentrations of S_1 and S_2, given by x_1 and x_2, respectively, generated from simulations of the MC, CLA, LNA, and complex-LE for the system given by* (54) *with parameters $r = 100$, $\nu = \alpha_1 = \alpha_2 = \beta_2 = 1$, and $\beta_1 = 25$. For the complex-LE, only the real parts of the variables are shown. Notice that LNA and the complex-LE predict negative values for x_1 at some times.*

this step involves replacing integer-valued molecule counts by real-valued concentration levels. It is intuitively clear that this modeling choice is likely to run into difficulties when one or more species has a small molecule count. This issue may manifest itself at a practical level by the solution path taking nonphysical negative values. From a technical perspective, the diffusion process may not remain well defined. Our aim in this work was to address this issue by introducing obliquely reflected diffusions as constrained Langevin approximations. The behavior of these diffusions matches that of solutions to the standard chemical Langevin equation in the interior of the positive orthant and introduces an appropriate minimal perturbation at the boundary. Our formal derivation of the CLA was backed up by analytical and computational examples that illustrate the benefits of the approach. A complementary, more rigorous, derivation of this diffusion approach, which includes existence and uniqueness proofs, has been developed in [27].

A direction for further work that we are pursuing is the development of error estimates for the CLA as an approximation to the underlying MC, both at the transient

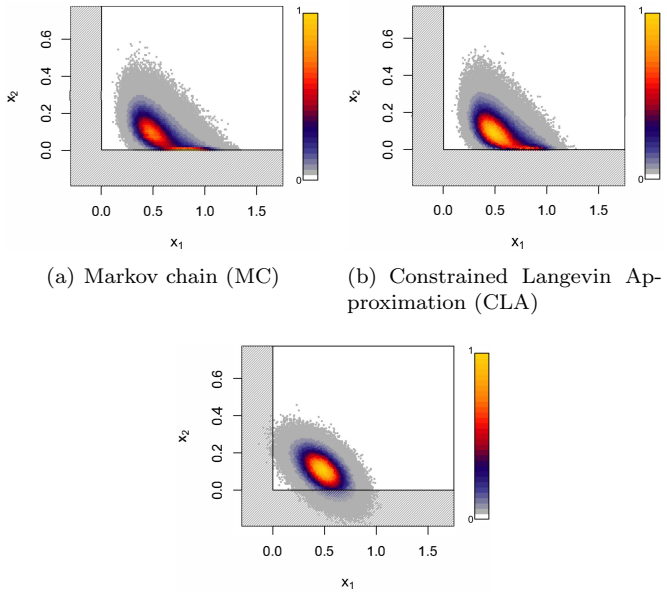
Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) (b)

FIG. 7. *(Left) Real parts (x_1 and x_2, representing concentration for species S_1 and S_2, respectively) for a simulation of the complex-LE of the system given by* (54) *with parameters $r = 100$, $\alpha_1 = \beta_1 = 1$, $\nu = 10$, $\alpha_2 = 0.02$, and $\beta_2 = 5$. (Right) Direction vector field for the drift (normalized to have unit length) of the complex-LE computed for the same simulation at time $t = 6665$, where the real parts of x_1 and x_2 are given by 1.0655 and -0.0483, respectively, (shown as a solid dot) and their imaginary parts are given by -0.0227 and 0.0376, respectively. The vector field shown was computed for values in the real x_1–x_2 plane with the imaginary parts fixed to -0.0227 and 0.0376 for x_1 and x_2, respectively.*



(a) Markov chain (MC) (b) Constrained Langevin Approximation (CLA)

(c) Linear Noise Approximation (LNA)

FIG. 8. *Scatter plot of the concentrations of S_1 and S_2, given by x_1 and x_2, respectively, generated from simulations of the MC, CLA, and LNA for the system given by* (54) *with parameters $r = 100$, $\alpha_1 = \beta_1 = 1$, $\nu = 10$, $\alpha_2 = 0.02$, and $\beta_2 = 5$. The complex-LE is not shown here since the approximation diverges during a long simulation (see Figure 7).*

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

TABLE 3

Integral of the absolute difference between the discrete density calculated for the Markov chain simulation and those calculated for the constrained Langevin approximation and the linear noise approximation (for parameters as in Figure 8). The values displayed are averages of 10 independent runs. The 95% confidence intervals for these averages are also shown.

	Integral abs. diff.	95% C. I.
Constrained Langevin Approximation (CLA)	0.2299	(0.2283, 0.2315)
Constrained Langevin Approximation (LNA)	0.4953	(0.4934, 0.4971)

and steady state level. In another vein, the type of diffusion approximation proposed here is also likely to be of interest for researchers considering other continuous time MCs that live in the positive orthant, e.g., in population genetics and neuroscience. The authors would appreciate hearing from researchers interested in such models and approximations.

Appendix A. Brief introduction to obliquely reflected diffusion processes. In this appendix, for the benefit of the reader, we give a very brief summary of some aspects of obliquely reflected diffusion processes as they pertain to the CLA described in this paper.

Let $\mathcal{G}^o$ be a nonempty domain in $\mathbb{R}^m$, $\gamma : \partial\mathcal{G} \to \mathbb{R}^m$ be a unit length vector field defined on the boundary $\partial\mathcal{G}$ of $\mathcal{G}^o$, and $\mu : \mathcal{G} \to \mathbb{R}^m$ and $\sigma : \mathcal{G} \to \mathbb{S}_+^m$ be continuous functions defined on the closure, $\mathcal{G}$, of $\mathcal{G}^o$. Here $\mathbb{S}_+^m$ denotes the set of (symmetric) $m \times m$ positive semidefinite matrices. Informally, a reflected diffusion associated with the parameters $(\mathcal{G}^o, \gamma, \mu, \sigma)$ is a continuous (strong) Markov process that behaves in the domain $\mathcal{G}^o$ like a solution of the Langevin equation with (state-dependent) drift μ and dispersion σ, and that is constrained to live in the closure $\mathcal{G}$ of $\mathcal{G}^o$ by a control at the boundary which acts in the (state-dependent) direction of the vector field γ. This type of control is often referred to as a singular control because it only acts when the diffusion process is on the boundary, and typically the amount of time that the diffusion process spends on the boundary has zero Lebesgue measure (and so the control acts only at a singular set of times).

One possible way[9] to try to define such a process precisely is to require it to be a continuous process $\mathfrak{Z}$ taking values in $\mathcal{G}$ that is a solution of the following SDE with reflection:

$$\mathfrak{Z}(t) = \mathfrak{Z}(0) + \int_0^t \mu(\mathfrak{Z}(s))\, ds + \int_0^t \sigma(\mathfrak{Z}(s))\, dW(s) + \int_0^t \gamma(\mathfrak{Z}(s))\, d\mathfrak{L}(s), \quad t \geq 0, \tag{55}$$

where W is a standard m-dimensional Brownian motion, the stochastic integral with respect to W is an Itô integral, and $\mathfrak{L}$ is a continuous, nondecreasing, one-dimensional process that satisfies $\mathfrak{L}(0) = 0$ and $\mathfrak{L}$ can only increase when $\mathfrak{Z}$ is on $\partial\mathcal{G}$, that is, $\int_0^\infty 1_{\{\mathfrak{Z}(s) \notin \partial\mathcal{G}\}} d\mathfrak{L}(s) = 0$. Here $\mathfrak{L}(t)$ is the cumulative amount of control (or pushing) exerted at the boundary, in the direction of the vector field γ, up to time t. Note that with $\mathfrak{Z}$, $\mathfrak{L}$ replaced by $\widetilde{Z}^r$, $\frac{1}{\sqrt{r}}\widetilde{L}^r$, and σ replaced by $\frac{1}{\sqrt{r}}\sigma$, (55) has the form of our CLA (33).

For historical reasons, a solution $\mathfrak{Z}$ of (55) is called a reflected diffusion process, although the constraining action at the boundary is more like regulation or control.

[9] An alternative approach is to try to characterize such processes in a distributional sense via submartingale problems, as first introduced by Stroock and Varadhan [35] for reflected diffusions in smooth domains and extended by various authors. See [21] for references on the two approaches and development of the relationship between them.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

The origin of the term reflection lies in the fact that when $m = 1$, $\mathcal{G}^o = (0, \infty)$, $\mu = 0$, $\sigma = 1$, $\gamma = 1$ and $\mathfrak{Z}(0) = 0$, $(\mathfrak{Z}, \mathfrak{L})$ is equivalent in distribution to $(|B|, L)$, where B is a standard one-dimensional Brownian motion, $|B|$ is its reflection about the origin, and $\mathfrak{L}$ is the "local time" of $|B|$ at the origin, which satisfies $L(t) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t 1_{[0,\epsilon)}(|B|(s))ds$ almost surely, and is a normalized measure of the amount of time that $|B|$ spends near the origin. Indeed, by Tanaka's formula [9],

$$|B(t)| = \int_0^t \operatorname{sgn}(B(s))\, dB(s) + L(t), \quad t \geq 0, \tag{56}$$

where $\operatorname{sgn}(x) = +1$ if $x > 0$, $\operatorname{sgn}(x) = -1$ if $x < 0$, and $\operatorname{sgn}(x) = 0$ if $x = 0$; and $\{\int_0^t \operatorname{sgn}(B(s))\, dB(s), t \geq 0\}$ defines another standard one-dimensional Brownian motion. On setting $W(t) = \int_0^t \operatorname{sgn}(B(s))\, dB(s)$ for $t \geq 0$, we see that $(\mathfrak{Z}, \mathfrak{L}) = (|B|, L)$ is a solution of (55) when $(\mathcal{G}^o, \mu, \sigma, \gamma) = ((0, \infty), 0, 1, 1)$. For more details on this reflected Brownian motion case, see [9, Chapters 7 and 8]. Although a mirror reflection construction of solutions of (55) does not generally hold for nonzero μ, state dependent σ, or γ and $\mathcal{G}^o$ in higher dimensions, the term "reflected diffusion" has nevertheless been widely used in the literature for processes of the form (55). We now describe the results relevant to existence and uniqueness of solutions of (55) beyond the simple one-dimensional Brownian motion case just described.

Of course, in general, additional conditions need to be imposed on $\mathcal{G}^o$, γ, μ, and σ in order for (55) to be well posed. For our CLA, $\mathcal{G}^o$ is naturally a polyhedral domain, and in all but one dimension, or in situations that can be reduced to such, the boundary will be nonsmooth, although it will be piecewise smooth. Also, if $\mu \neq 0$ on $\partial\mathcal{G}$, then $\gamma = \mu/|\mu|$ can be extended to a smooth (in fact, C^∞) function in a neighborhood of $\partial\mathcal{G}$.

In [11], Dupuis and Ishii considered the problem of existence and uniqueness of solutions of equations like (55) when the boundary of $\mathcal{G}$ is not smooth. The first of two cases that they treated is relevant to CLAs as it allows for a smooth, state-dependent vector field γ defined on a nonsmooth boundary. For that case, they assume that $\mathcal{G}^o$ is a bounded domain, γ can be extended to a C^2, unit-length vector field on all of $\mathbb{R}^m$, and μ and σ are uniformly Lipschitz continuous on $\mathcal{G}$. They formulate sufficient conditions for the existence and uniqueness of "strong"[10] solutions of (55). The critical condition (3.2) in their paper requires that, at each point on the boundary, the vector field γ points into $\mathcal{G}^o$ in a suitably strong way. Unfortunately, due to topological constraints, such an inward pointing vector field cannot be globally extended to be smooth and of unit length on all $\mathbb{R}^m$ (see [27] for a counterexample). However, as shown in [27, section 5], the existence and uniqueness result of [11] is in fact true with only local extendability of γ. In [27], Leite and Williams further show that this existence and uniqueness result can be extended to where $\mathcal{G}$ is the unbounded positive orthant in $\mathbb{R}^m$, under the assumptions described in our paragraph containing (34). In particular, with the results of Dupuis and Ishii (as extended in [27]) for bounded domains and those in [27] for the orthant, existence and uniqueness of solutions of our CLA (33) for all of the examples considered in this paper, as well as many others, can be obtained. We refer the interested reader to [11, 21, 27] for more details on reflected diffusions, especially in nonsmooth domains.

[10] A strong solution is required to be adapted to the filtration generated by W and the initial condition $\mathfrak{Z}(0)$.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Acknowledgments. The genesis of this work was in informal discussions between DFA, DJH, and RJW at a workshop on Multi-scale Stochastic Modeling of Cell Dynamics held at the Banff International Research Station in 2010. We are grateful to the organizers of that workshop for making this serendipitous connection possible.

REFERENCES

[1] D. F. Anderson, G. Craciun, and T. G. Kurtz, *Product-form stationary distributions for deficiency zero chemical reaction networks*, Bull. Math. Biol., 72 (2010), pp. 1947–1970.
[2] D. F. Anderson and T. G. Kurtz, *Continuous time Markov chain models for chemical reaction networks*, in Design and Analysis of Biomolecular Circuits: Engineering Approaches to Systems and Synthetic Biology, H. Koeppl et al., ed., Springer, New York, 2011, pp. 3–42.
[3] D. F. Anderson and T. G. Kurtz, *Stochastic Analysis of Biochemical Systems*, Springer, Cham, Switzerland, 2015.
[4] A. Angius, G. Balbo, M. Beccuti, E. Bibbona, A. Horvath, and R. Sirovich, *Approximate analysis of biological systems by hybrid switching jump diffusion*, Theoret. Comput. Sci., 587 (2015), pp. 49–72.
[5] E. Baake, T. Kurtz, and C. Wiuf, *Reaction Networks and Population Dynamics*, Oberwolfach Rep., 14 (2017), pp. 1747–1804.
[6] S. Bhardwaj and R. J. Williams, *Diffusion approximation for a heavily loaded multi-user wireless communication system with cooperation*, Queueing Syst., 62 (2009), pp. 345–382.
[7] R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.
[8] M. Bossy, E. Gobet, and D. Talay, *A symmetrized Euler scheme for an efficient approximation of reflected diffusions*, J. Appl. Probab., 41 (2004), pp. 877–889.
[9] K. L. Chung and R. J. Williams, *An Introduction to Stochastic Integration*, 2nd ed., Birkhäuser, Boston, 1990.
[10] A. Duncan, R. Erban, and K. Zygalakis, *Hybrid framework for the simulation of stochastic chemical kinetics*, J. Comput. Phys., 326 (2016), pp. 398–419.
[11] P. Dupuis and H. Ishii, *SDEs with oblique reflection on nonsmooth domains*, Ann. Probab., 21 (1993), pp. 554–580; *Correction*, Ann Probab., 36 (2008), pp. 1992–1997.
[12] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, 2nd ed., Wiley, New York, 2005.
[13] A. Ganguly, D. Altintan, and H. Koeppl, *Jump-diffusion approximation of stochastic reaction dynamics: Error bounds and algorithms*, Multiscale Model. Simul., 13 (2015), pp. 1390–1419.
[14] M. A. Gibson and J. Bruck, *Efficient exact stochastic simulation of chemical systems with many species and many channels*, J. Phys. Chem. A, 104 (2000), pp. 1876–1889.
[15] D. T. Gillespie, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, J. Comput. Phys., 22 (1976), pp. 403–434.
[16] D. T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem., 81 (1977), pp. 2340–2361.
[17] D. T. Gillespie, *Markov Processes: An Introduction for Physical Scientists*, Academic Press, Boston, 1992.
[18] D. T. Gillespie, *The chemical Langevin equation*, J. Comput. Phys., 113 (2000), pp. 297–306.
[19] D. T. Gillespie, *The chemical Langevin and Fokker-Planck equations for the reversible isomerization reaction*, J. Phys. Chem. A, 106 (2002), pp. 5063–5071.
[20] T. Jahnke and M. Kreim, *Error bound for piecewise deterministic processes modeling stochastic reaction systems*, Multiscale Model. Simul., 10 (2012), pp. 1119–1147.
[21] W. Kang and K. Ramanan, *On the submartingale problem for reflected diffusions in domains with piecewise smooth boundaries*, Ann. Probab., 45 (2017), pp. 404–468.
[22] W. N. Kang and K. Ramanan, *Characterization of stationary distributions of reflected diffusions*, Ann. Appl. Probab., 24 (2014), pp. 1329–1374.
[23] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed., Springer, New York, 1991.
[24] J. Kómlos, P. Major, and G. Tusnady, *An approximation of partial sums of independent random variables and the sample distribution function*, Z. Wahrscheinlichkeitstheorie und Verwandte Geb., 32 (1975), pp. 111–131.
[25] T. G. Kurtz, *Strong approximation theorems for density dependent Markov chains*, Stochastic Process. Appl., 6 (1977), pp. 223–240.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[26] T. G. Kurtz, *Approximation of Population Processes*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 36, SIAM, Philadelphia, 1981.
[27] S. C. Leite and R. J. Williams, *A constrained Langevin approximation for chemical reaction networks*, Ann. Appl. Probab., to appear.
[28] T. Manninen, M.-L. Linne, and K. Ruohonen, *Developing Itô stochastic differential equation models for neuronal signal transduction pathways*, Comput. Biol. Chem., 30 (2006), pp. 280–291.
[29] G. Minas and D. A. Rand, *Long-time analytic approximation of large stochastic oscillators: Simulation, analysis and inference*, PLoS Comput. Biol., 13 (2017), pp. 1–23.
[30] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018.
[31] M. I. Reiman and R. J. Williams, *A boundary property of semimartingale reflecting Brownian motions*, Probab. Theory Related Fields, 77 (1988), pp. 87–97. Correction, Probab. Theory Related Fields, 80 (1989), p. 633.
[32] J. Schnakenberg, *Simple chemical reaction systems with limit cycle behaviour*, J. Theoret. Biol., 81 (1979), pp. 389–400.
[33] D. Schnoerr, G. Sanguinetti, and R. Grima, *The complex chemical Langevin equation*, J. Chem. Phys., 141 (2014), 024103.
[34] M. Scott and B. P. Ingalls, *Using the linear noise approximation to characterize molecular noise in reaction pathways*, in Proceedings of the AIChE Conference on Foundations of Systems Biology in Engineering (FOSBE), Santa Barbara, CA, IEE, London, 2005.
[35] D. W. Stroock and S. R. S. Varadhan, *Diffusion processes with boundary conditions*, Comm. Pure Appl. Math., 24 (1971), pp. 147–225.
[36] L. Szpruch and D. J. Higham, *Comparing hitting time behavior of Markov jump processes and their diffusion approximations*, Multiscale Model. Simul., 8 (2010), pp. 605–621.
[37] H. Tanaka, *Stochastic differential equations with reflecting boundary condition in convex regions*, Hiroshima Math. J., 9 (1979), pp. 163–177.
[38] K. Tomita, T. Ohta, and H. Tomita, *Irreversible circulation and orbital revolution—hard mode instability in far-from-equilibrium situation*, Progr. Theoret. Phys., 52 (1974), pp. 1744–1765.
[39] N. G. van Kampen, *A power series expansion of the master equation*, Canad. J. Phys., 39 (1961), pp. 551–567.
[40] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed., Elsevier, Amsterdam, 2007.
[41] E. W. J. Wallace, L. R. Petzold, D. T. Gillespie, and K. R. Sanft, *Linear noise approximation is valid over limited times for any chemical system that is sufficiently large*, IET Syst. Biol., 6 (2012), pp. 102–115.
[42] J. Wilkie and Y. M. Wong, *Positivity preserving chemical Langevin equations*, Chem. Phys., 353 (2008), pp. 132–138.
[43] D. J. Wilkinson, *Stochastic Modelling for Systems Biology*, Chapman and Hall/CRC Press, Boca Raton, FL, 2006.

Downloaded 06/18/21 to 192.41.114.224. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.