The Annals of Applied Probability
2019, Vol. 29, No. 3, 1541–1608
https://doi.org/10.1214/18-AAP1421

A CONSTRAINED LANGEVIN APPROXIMATION FOR CHEMICAL REACTION NETWORKS

BY SAUL C. LEITE[1] AND RUTH J. WILLIAMS[2]

Universidade Federal do ABC and University of California, San Diego

Continuous-time Markov chain models are often used to describe the stochastic dynamics of networks of reacting chemical species, especially in the growing field of systems biology. These Markov chain models are often studied by simulating sample paths in order to generate Monte-Carlo estimates. However, discrete-event stochastic simulation of these models rapidly becomes computationally intensive. Consequently, more tractable diffusion approximations are commonly used in numerical computation, even for modest-sized networks. However, existing 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 (Langevin approximation).

In this paper, we propose an approximation for such Markov chains via reflected diffusion processes that respect the fact that concentrations of chemical species are never negative. We call this a constrained Langevin approximation because it behaves like the Langevin approximation in the interior of the positive orthant, to which it is constrained by instantaneous reflection at the boundary of the orthant. An additional advantage of our approximation is that it can be written down immediately from the chemical reactions. This contrasts with the linear noise approximation, which involves a two-stage procedure—first solve a deterministic reaction rate ordinary differential equation, followed by a stochastic differential equation for fluctuations around those solutions. Our approximation also captures the interaction of nonlinearities in the reaction rate function with the driving noise. In simulations, we have found the computation time for our approximation to be at least comparable to, and often better than, that for the linear noise approximation.

Under mild assumptions, we first prove that our proposed approximation is well defined for all time. Then we prove that it can be obtained as the weak limit of a sequence of jump-diffusion processes that behave like the Langevin approximation in the interior of the positive orthant and like a rescaled version of the Markov chain on the boundary of the orthant. For this limit the-

Received November 2016; revised April 2018.

[1]Supported by CAPES Process No. 23071.012226/2013-06 and by FAPEMIG Process No. APQ 00945/14. Part of this research was conducted while SCL was affiliated with the Universidade Federal de Juiz de Fora, MG, Brazil.

[2]Supported in part by NSF Grants DMS-1206772 and DMS-1712974.

MSC2010 subject classifications. Primary 60J28, 60J60, 60F17, 92C45; secondary 60H10, 92C40.

Key words and phrases. Diffusion approximation, Langevin approximation, chemical reaction networks, stochastic differential equation with reflection, scaling limit, linear noise approximation, density dependent Markov chains, jump-diffusion.

orem, we adapt an invariance principle for reflected diffusions, due to Kang and Williams [*Ann. Appl. Probab.* **17** (2007) 741–779], and modify a result on pathwise uniqueness for reflected diffusions due to Dupuis and Ishii [*Ann. Probab.* **21** (1993) 554–580]. Some numerical examples illustrate the advantages of our approximation over direct simulation of the Markov chain or use of the linear noise approximation.

CONTENTS

1. Introduction. A chemical reaction network is a chemical system composed of a set of molecular species which are involved in multiple reactions. One is usually interested in the time evolution of molecular abundances, or their concentrations, as a series of reactions occurs. The most common model for these networks is the so-called Reaction Rate Equation (RRE) model, which is described by the following ordinary differential equation (written in integrated form):

$$\bar{x}(t) = \bar{x}(0) + \int_0^t \mu(\bar{x}(s))\, ds, \qquad t \geq 0, \tag{1.1}$$

where $\mu(x) = \sum_{k=1}^r v_k \lambda_k(x)$ for $x \in \mathbb{R}_+^d$, $\{\lambda_k\}_{k=1}^r$ are the deterministic (nonnegative) reaction rate functions, and the d-dimensional vectors $\{v_k\}_{k=1}^r$ are such that

$v_k = v_k^+ - v_k^-$ denotes the change in the number of copies of each of the species that occurs during the kth reaction, where v_k^+, v_k^- have nonnegative integer-valued entries. For $i = 1, \dots, d$, and $t \geq 0$, the ith component $\bar{x}_i(t)$ of $\bar{x}(t)$ represents the concentration of molecular species i in the system at time t.

When the system is well stirred and under thermal equilibrium, the most widely used stochastic model for this system is a continuous time Markov chain [10], which tracks the number of molecules of each species and whose forward Kolmogorov equation is commonly known as the chemical master equation. For certain classes of chemical reaction networks, it has been shown [20, 22] that the RRE model is the almost sure limit of a scaled version of this Markov chain model, where the limit is taken as the volume of the vessel containing the molecules tends to infinity. In particular, the RRE model is often adequate for modeling systems containing large quantities of each molecular species.

However, in some application areas, such as molecular biology [27, 37, 43], the number of molecules present in the system, at least for some species, is not large enough to justify the use of the deterministic RRE model, since stochastic fluctuations in the system are significant. In such cases, the Markov chain model is most frequently used. Such a model is often studied by simulating sample paths. These simulations are typically repeated many times to generate Monte Carlo estimates and they can become computationally expensive if performed by direct methods [8, 9]. This can be aggravated in cases where the number of molecules for some species are relatively large, since every reaction is tracked in these simulations. Much effort has been given to finding efficient methods of simulation. One approach to improve computational efficiency is through approximate simulation algorithms, such as the τ-leaping method [12]. Another alternative is via diffusion approximations, which are often seen as reliable and yield efficient computer simulation through commonly known methods for Stochastic Differential Equations (SDEs), such as the Euler–Maruyama method [13, 14, 26, 29]. To set the scene for our proposed approximation, we next briefly describe the existing diffusion approximations. Readers interested in a more expository introduction to this paper are referred to [1], where additional examples can also be found.

1.1. *Existing diffusion approximations.* One diffusion approximation is the so-called *linear noise (or van Kampen) approximation*, which was introduced in [38]. This approximation can be understood as linearizing a scaled version of the Markov chain about the solution of the deterministic RRE model, to capture the effect of stochastic fluctuations. In fact, under certain conditions, it was shown in [22] that one may construct stochastic processes $\bar{X}^V$, D, $\{W_k\}_{k=1}^r$ on the same probability space such that $\bar{X}^V$ is the scaled Markov chain, D is the solution of the following SDE:

$$D(t) = \int_0^t J\mu(\bar{x}(s)) D(s)\, ds + \sum_{k=1}^r \int_0^t v_k \sqrt{\lambda_k(\bar{x}(s))}\, dW_k(s), \qquad t \geq 0,$$

and $\bar{X}^V$ can be approximated by $\tilde{Z} = \bar{x} + V^{-1/2}D$, where the error in the approximation is of order $\xi \log(V)/V$ on a bounded time interval. Here, V is a constant proportional to the volume of the vessel containing the chemical system, ξ is a random variable with a finite exponential moment, $\bar{x}$ is the solution of the RRE and $J\mu(x)$ represents the Jacobian matrix of $\mu(x)$, defined below (1.1). By direct substitution, one can write $\tilde{Z}$ as the solution of the following SDE:

$$\begin{aligned}\tilde{Z}(t) &= \bar{x}(0) + \int_0^t \big(\mu\big(\bar{x}(s)\big) + J\mu\big(\bar{x}(s)\big)\big(\tilde{Z}(s) - \bar{x}(s)\big)\big)\,ds \\ &\quad + \frac{1}{\sqrt{V}} \sum_{k=1}^r \int_0^t v_k \sqrt{\lambda_k\big(\bar{x}(s)\big)}\,dW_k(s), \qquad t \geq 0.\end{aligned} \tag{1.2}$$

Since the linear noise approximation is an unconstrained diffusion process, it typically can take negative values, which may violate the natural nonnegativity of concentrations. In addition, a two-stage procedure is required to use this approximation: one must first solve the RRE for $\bar{x}$ and then solve (1.2). Note that the right-hand side of (1.2) is linear in $\tilde{Z}$. As a consequence, the linear noise approximation can capture local fluctuations near a steady-state of the RRE quite well, but it typically fails to capture global nonlinear behavior well (see the examples in Section 8).

A more frequently used diffusion approximation for chemical reaction networks is the *Langevin approximation* [21, 22] (see also [11]), obtained by solving the (chemical) *Langevin Equation (LE)*. Unfortunately, this approximation is typically only valid until the concentration of some species first reaches zero. It is often preferred in practice because it is a direct approximation to the scaled Markov chain model and it can capture nonlinear effects in the noise. The Langevin equation is given by the following SDE:

$$\begin{aligned}Z(t) &= \bar{X}^V(0) + \int_0^t \mu\big(Z(s)\big)\,ds \\ &\quad + \frac{1}{\sqrt{V}} \sum_{k=1}^r \int_0^t v_k \sqrt{\lambda_k\big(Z(s)\big)}\,dW_k(s), \qquad t \geq 0,\end{aligned} \tag{1.3}$$

where now one approximates $\bar{X}^V$ directly by Z. While there are results on the weak existence and uniqueness of solutions to such equations when the λ_k are defined to be *nonnegative* everywhere on $\mathbb{R}^d$ (see e.g., [6], page 459), where d is the number of species, there are no known general results for existence or uniqueness of solutions of the SDE for the λ_k associated with chemical reaction networks. Indeed, as has been pointed out, for example, in [25, 28, 36, 42], these SDEs may not make sense for all times since they may predict negative molecular concentrations and their dispersion coefficients involve square roots of these quantities.

To illustrate the last point, the Langevin equation proposed as an approximation for the Markov chain model associated with the simple chemical reaction network composed of two species S_1 and S_2 and the reactions:

$$S_1 \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} S_2$$

is written in differential form as

$$dZ_1(t) = \big(\beta Z_2(t) - \alpha Z_1(t)\big)\,dt + \frac{1}{\sqrt{V}}\big(\sqrt{\beta Z_2(t)}\,dW_2(t) - \sqrt{\alpha Z_1(t)}\,dW_1(t)\big),$$

$$dZ_2(t) = \big(\alpha Z_1(t) - \beta Z_2(t)\big)\,dt + \frac{1}{\sqrt{V}}\big(\sqrt{\alpha Z_1(t)}\,dW_1(t) - \sqrt{\beta Z_2(t)}\,dW_2(t)\big).$$

As pointed out in [42], when Z_1 is near zero, the term $\sqrt{\beta Z_2(t)}\,dW_2(t)$ may push Z_1 to take negative values and the equations become ill-posed, because one needs to take the square root of the negative quantity αZ_1. Similarly, there is a problem when Z_2 is near zero.

Consequently, the Langevin approximation for chemical reaction networks, obtained by solving the Langevin equation (1.3), is usually only valid up until Z reaches the boundary of the positive orthant. Sufficient conditions for the validity of such an approximation, which involves stopped diffusions, are given in the paper of Kurtz [21]. Some authors also have considered extending the range of solutions for the Langevin equation to the complex numbers [32]. Here, we focus on an approximation that respects the physical constraints of the positive orthant.

1.2. *Our constrained Langevin approximation*. In this paper, we provide motivation to approximate the scaled Markov chain, $\bar{X}^V$, by a reflected diffusion process that lives in the positive orthant of $\mathbb{R}^d$, and thereby respects the fact that chemical concentrations are never negative. We first propose approximating $\bar{X}^V$ by a jump-diffusion process. This jump-diffusion is a process that has the same behavior as a solution of the Langevin equation in the interior of the positive orthant of $\mathbb{R}^d$ and behaves like the scaled Markov chain $\bar{X}^V$ at the boundary. Then, in order to obtain a continuous diffusion approximation, we let the size of the jumps decrease to zero, while simultaneously increasing their frequency. The resulting limit process Z is a continuous process that lives in the positive orthant $\mathbb{R}^d_+$ and satisfies the following Stochastic Differential Equation with Reflection (SDER):

$$\begin{aligned}Z(t) = Z(0) &+ \int_0^t \mu\big(Z(s)\big)\,ds \\ &+ \frac{1}{\sqrt{V}}\bigg(\int_0^t \sigma\big(Z(s)\big)\,dW(s) + \int_0^t \gamma\big(Z(s)\big)\,dL(s)\bigg), \qquad t \ge 0.\end{aligned} \tag{1.4}$$

Here, $\mu(x)$ is defined below (1.1), $\sigma(x)$ is the symmetric positive definite square root of the diffusion matrix $\Gamma(x) = \sum_{k=1}^r v_k v_k' \lambda_k(x)$, where $'$ denotes transpose, $\gamma(x) = \mu(x)/|\mu(x)|$ defines the reflection vector field on the boundary of $\mathbb{R}^d_+$, and W is a d-dimensional Brownian motion. The process L is a one-dimensional, continuous, increasing process that tracks the cumulative amount of pushing done at the boundary (the process L is sometimes called a boundary "local time"). The integral term involving L is called the reflection term; when Z is on the boundary, L increases instantaneously to push Z in the state-dependent direction given by γ, in the minimal amount needed to keep Z in the positive orthant of $\mathbb{R}^d$. Note that the process Z has the same infinitesimal drift μ and infinitesimal covariance $\frac{1}{V}\Gamma$ as a solution of the Langevin equation (1.3). Consequently, the solutions of (1.3) and (1.4) will have the same distributional behavior in the interior of the positive orthant of d-dimensional Euclidean space.[3]

Assuming that the reaction network satisfies a mass-dissipating assumption, augmented by inflows and outflows on all species, we show that the SDER is well posed. We also prove that our sequence of jump-diffusion processes, with jump-size tending to zero, converges weakly to a solution of the SDER. We call this limit process, satisfying the SDER, the constrained Langevin approximation to the scaled Markov chain.

1.3. *Outline of the rest of the paper.* The paper is structured as follows. We begin in Section 2 by defining the common notation that will be used throughout the paper. In Section 3, we introduce chemical reaction networks, present the associated Markov chain model for systems driven by the law of mass action and introduce the assumptions that will be used throughout the paper. In particular, we consider a class of chemical reaction networks that satisfy a "mass-dissipating" assumption with additional inflow and outflow reactions for every species. In Section 4, we present motivation to approximate the scaled Markov chain model for this class of chemical reaction networks by a jump-diffusion process. We then consider a family of such jump-diffusion processes, in which the size of the jumps from the boundary is allowed to decrease to zero, and the frequency of the jumps tends to infinity. We do this in such a way that the boundary terms stay of the same order as the fluctuation terms, that is, of order $\frac{1}{\sqrt{V}}$, as the jump size goes to zero and the frequency goes to infinity. We will prove that any sequence of these jump-diffusion processes with jump size tending to zero converges weakly to a solution of the SDER, which we call the Constrained Langevin Equation (CLE). In Section 5, we give a precise definition for solutions of the CLE and, in Section 6, we show that the equation satisfies pathwise uniqueness, which also implies weak uniqueness. For the pathwise uniqueness, we modify a proof for reflected diffusions on bounded domains due to Dupuis and Ishii[4] [5]. In Section 7, we give the

[3]One might consider keeping the form of the diffusion term in (1.3) (which involves r independent one-dimensional Brownian motions W_k), for the diffusion term in (1.4). However, under our assumptions, we will have $r > d$, and it is more efficient to combine the r terms into a term driven by a d-dimensional Brownian motion W.

[4]This fixes an issue with the assumptions on the reflection vector field in [5], which as stated (in their global form), are impossible to satisfy.

proof that any sequence of our jump-diffusion processes, with jump size decreasing to zero, converges weakly to a solution of the CLE. For this limit theorem, we adapt an invariance principle for reflected diffusions in piecewise smooth domains, due to Kang and Williams [17], and use the uniqueness in law established in Section 6. Finally, in Section 8, the results of simulation experiments are provided that demonstrate the effectiveness of our proposed constrained Langevin approximation.

2. Preliminaries and notation. For any integer $d \geq 1$, let $\mathbb{Z}_+^d$ denote the d-dimensional lattice of points with nonnegative integer coordinates. We shall often write $\mathbb{Z}_+$ instead of $\mathbb{Z}_+^1$. Let $\mathbb{R}^d$ denote the d-dimensional Euclidean space and $\mathbb{R}_+^d$ denote the set of points in $\mathbb{R}^d$ whose coordinates are all nonnegative. When $d = 1$, we usually write $\mathbb{R}$ and $\mathbb{R}_+$ for $\mathbb{R}^1$ and $\mathbb{R}_+^1$, respectively. The zero vector in $\mathbb{R}^d$ will be simply written as $\mathbb{0}$. Similarly, the vector in $\mathbb{R}^d$ with all components set to 1 is denoted by $\mathbb{1}$. For $x \in \mathbb{R}^d$, let $|x| = (\sum_{i=1}^d |x_i|^2)^{1/2}$ be the usual Euclidean norm, and let $|x|_\infty = \max_{i=1}^d |x_i|$ be the maximum norm. Vectors will usually be column vectors unless indicated otherwise. The transpose of a vector x or matrix A will be denoted by x' or A', respectively. For two vectors $x, y \in \mathbb{R}^d$, we use $\langle x, y\rangle$ to denote the inner product between the two vectors. Inequalities between vectors are to be interpreted componentwise, for example, $x \geq y$ for $x, y \in \mathbb{R}^d$ means $x_i \geq y_i$ for $i = 1, \dots, d$. Let e_i denote the unit vector in the ith coordinate direction in $\mathbb{R}^d$. The set of $d \times l$ matrices with real-valued entries will be denoted by $\mathbb{R}^{d\times l}$. For $A \in \mathbb{R}^{d\times l}$, we denote by $|A| = (\sum_{i=1}^d \sum_{j=1}^l |A_{ij}|^2)^{1/2}$, the Frobenius norm of A. Let $\mathbb{S}^{d\times d}$ denote the set of $d \times d$ symmetric positive definite matrices with real-valued entries. Here, positive definite means strictly positive definite: $\Gamma \in \mathbb{S}^{d\times d}$ if and only if there is a $c > 0$ such that $x'\Gamma x \geq c|x|^2$ for all $x \in \mathbb{R}^d$.

For a set S, we use $|S|$ to denote the number of elements in S. For $S \subset \mathbb{R}^d$, we define $\mathrm{dist}(x, S) \triangleq \inf\{|x - y| : y \in S\}$. Let $\mathbb{B}_r(x)$ denote the closed ball $\{y \in \mathbb{R}^d : |x - y| \leq r\}$ for all $r > 0$ and $x \in \mathbb{R}^d$. In addition, for $S \subset \mathbb{R}^d$, let $C_b(S)$ denote the set of functions $f : S \to \mathbb{R}$ that are continuous and bounded on S; let $C_b^2(S)$ denote the set of twice differentiable functions $f : S \to \mathbb{R}$ that are continuous and bounded together with their first and second partial derivatives; and let $C_c^2(S)$ denote the set of twice continuously differentiable functions $f : S \to \mathbb{R}$ that have compact support.

We denote by $\mathcal{C}$ the set of continuous functions $\omega : [0, \infty) \to \mathbb{R}^d$, equipped with the topology of uniform convergence on compact sets. In addition, let $\mathcal{M}$ denote the Borel σ-algebra associated with $\mathcal{C}$. The measurable space $(\mathcal{C}, \mathcal{M})$ is endowed with the filtration $\{\mathcal{M}_t\}$, where $\mathcal{M}_t \triangleq \sigma\{\omega(s) : 0 \leq s \leq t\}$, which denotes the smallest σ-algebra of subsets of $\mathcal{C}$ that makes the coordinate maps $\omega \to \omega(s)$ from $\mathcal{C}$ to $\mathbb{R}^d$ $\mathcal{M}_t$-measurable for $0 \leq s \leq t$. Denote by $\mathcal{D}$ the set of functions $\omega : [0, \infty) \to \mathbb{R}^d$ that are right continuous and have finite left-hand limits on $(0, \infty)$, equipped with Skorokhod's J_1-topology. We also denote the Borel σ-algebra on $\mathcal{D}$ by $\mathcal{M}$.

Unless indicated otherwise, stochastic processes considered in this paper will have sample paths living in $\mathbb{R}^d$, that are right continuous with finite left limits. Occasionally, we will consider stochastic processes with sample paths that live in $\mathbb{R}^d$ or $\mathbb{R}^d_+$, augmented with an isolated cemetery state, to accommodate possible explosion in finite time. These stochastic processes will have right-continuous paths with finite left limits at times when they are in $\mathbb{R}^d$ or $\mathbb{R}^d_+$, and if they reach the cemetery state, they stay there forever after. Consider a sequence of d-dimensional stochastic processes $\{X^n\}_{n=1}^{\infty}$. This sequence is said to be *tight* if the probability measures induced by the X^n on the measurable space $(\mathcal{D}, \mathcal{M})$ form a tight sequence, that is, the induced probability measures form a weakly relatively compact sequence in the space of probability measures on $(\mathcal{D}, \mathcal{M})$. The sequence of processes $\{X^n\}_{n=1}^{\infty}$ is called $\mathcal{C}$-*tight* if it is tight and if each weakly converging subsequence has a limit point that has sample paths in $\mathcal{C}$ almost surely.

3. Chemical reaction networks. Here, we introduce chemical reaction networks as described in [7]. For integers $d \geq 1$ and $r \geq 1$, a chemical reaction network consists of a triple $(\mathscr{S}, \mathscr{R}, \mathscr{C})$, where $\mathscr{S}$ consists of d molecular species $\mathscr{S} \triangleq \{S_1, \dots, S_d\}$ involved in r possible reactions $\mathscr{R} \triangleq \{R_1, \dots, R_r\}$. For each $1 \leq k \leq r$, let $v_k^- \in \mathbb{Z}^d_+$ be a vector denoting the number of molecules of each species *consumed* by the kth reaction R_k. That is, if the ith component v_{ik}^- of v_k^- equals a, then a molecules of species S_i are consumed by reaction R_k. Similarly, let $v_k^+ \in \mathbb{Z}^d_+$ denote the number of molecules of each species *produced* by reaction R_k. These vectors v_k^- and v_k^+ are generally called complexes and form the set $\mathscr{C}$ (i.e., $\mathscr{C} \triangleq \{v_k^-, v_k^+ : 1 \leq k \leq r\}$). For example, in a chemical reaction network involving three species, namely S_1, S_2 and S_3, a reaction R_k, that consumes one molecule of S_2 and one of S_3 and produces one molecule of S_1 and one of S_3, is usually denoted by

$$S_2 + S_3 \to S_1 + S_3$$

and is associated with vectors: $v_k^- = (0, 1, 1)'$ and $v_k^+ = (1, 0, 1)'$. In general, the kth reaction, R_k, is represented as follows:

$$\sum_{i=1}^{d} v_{ik}^- S_i \to \sum_{i=1}^{d} v_{ik}^+ S_i,$$

for $1 \leq k \leq r$. A reaction may consume zero molecules or not produce any molecules at all. Such reactions are associated with vectors $v_k^- = \mathbb{0}$ or $v_k^+ = \mathbb{0}$, respectively, and are usually denoted by

$$\varnothing \to \sum_{i=1}^{d} v_{ik}^+ S_i \quad \text{or} \quad \sum_{i=1}^{d} v_{ik}^- S_i \to \varnothing.$$

The interpretation for these reactions is that molecules are joining the system from an external source or leaving the system, respectively.

The most common stochastic models for these chemical reaction networks use continuous-time Markov chains that track the number of molecules of each species that are present in the system at each instant of time. These models assume that the times between consecutive reactions (or inter-reaction times) are exponentially distributed. The rate parameters for these exponential distributions are functions of the numbers of molecules of each species that are present in the system and are usually called *propensity* functions in the chemistry and systems biology literature.

For well-mixed systems, these propensity functions are given by the so-called *stochastic law of mass action*. Suppose $x \in \mathbb{Z}_+^d$ is the number of molecules of each species in the system at some time. For $k = 1, \dots, r$,

$$\Lambda_k(x) \triangleq \kappa_k \prod_{i=1}^{d} (x_i)_{v_{ik}^-} \tag{3.1}$$

is the rate of reaction R_k under the stochastic law of mass action when the system is at the state x. Here, $(x_i)_{v_{ik}^-}$ denotes the falling factorial: $(x_i)_{v_{ik}^-} \triangleq x_i(x_i - 1)\cdots(x_i - v_{ik}^- + 1)$, with the interpretation that $(x_i)_0 = 1$. Also, $\kappa_k > 0$ is a given constant, that is usually called the *reaction rate constant*.

Reactions are typically ranked according to the number of molecules consumed. For instance:

- *Zero-order reactions* correspond to inflow of molecules from an outside source. The propensities are given by $\Lambda_k(x) = \kappa_k$, where κ_k is a positive constant. An example for this type of reaction is the following:

$$\varnothing \to S_1,$$

which corresponds to a molecule of S_1 joining the system from an external source.

- *First-order reactions* consume one molecule of a certain species. The propensities are given by $\Lambda_k(x) = \kappa_k x_i$, for some positive constant κ_k and $i \in \{1, \dots, d\}$. An example for such a reaction is

$$S_1 \to S_2,$$

where one molecule of S_1 is consumed in order to produce a molecule of S_2.

- *Second-order reactions* consume two molecules of one or more species. The propensities are given by

$$\Lambda_k(x) = \begin{cases} \kappa_k x_i x_j & \text{for } i \neq j, \\ \kappa_k x_i(x_i - 1) & \text{otherwise} \end{cases}$$

for some $i, j \in \{1, \dots, d\}$ and positive constant κ_k. An example of this type of reaction is the following:

$$S_1 + S_2 \to \varnothing,$$

which corresponds to a molecule of S_1 and one of S_2 leaving the system together.

The order of the reaction R_k is given by $[v_k^-] \triangleq \sum_{i=1}^d v_{ik}^-$. Reactions of third order and higher are considered unlikely to occur in nature, since it would require more than two molecules to collide at the same instant [10]. Generally, a reaction involving more than two molecules may be regarded as the combination of two or more reactions of second order. However, it is not uncommon to find higher order reactions in the literature. These are often employed as a simplifying device to reduce the number of species tracked in the system. In this paper, to allow the most generality, we do not limit the order of reactions that we consider.

Under the assumption that the inter-reaction times are exponentially distributed and the rates satisfy the stochastic law of mass action, the number of occurrences of reaction R_k by time t may be represented by

$$U_k(t) \triangleq N_k\left(\int_0^t \Lambda_k(X(s))\,ds\right),$$

where N_k is a unit rate Poisson process and $X(t)$ is a random vector taking values in $\mathbb{Z}_+^d$ that tracks the number of molecules of each species that are present in the system at time t. The Poisson processes $\{N_k\}_{k=1}^r$ are assumed to be mutually independent. For notational convenience, it is sometimes useful to define the so-called stoichiometric matrix which has columns $v_k \triangleq v_k^+ - v_k^-$, $1 \le k \le r$, that is,

$$S \triangleq (v_1, \dots, v_r) \in \mathbb{Z}^{d\times r},$$

where $\mathbb{Z}^{d\times r}$ denotes the vector space of $d \times r$ matrices with integer-valued entries. In addition, we define the vector valued process $U \triangleq (U_1, \dots, U_r)'$.

The stochastic process $X = \{X(t)\}_{t\ge 0}$ is a Markov chain and, provided it does not explode in finite time, it can be represented as the solution of the following equation:

$$X(t) = X(0) + SU(t), \tag{3.2}$$

for each $t \ge 0$, where the random vector $X(0)$ taking values in $\mathbb{Z}_+^d$ is the initial configuration of the system, which is assumed to be independent of the driving Poisson processes. If X explodes in finite time, (3.2) holds prior to the random explosion time ζ. The function $\Lambda_k(X(\cdot))$ is the stochastic intensity of the point process U_k, and we henceforth refer to $\Lambda_k(X(\cdot))$ as the *intensity* of reaction R_k. Given $X(0), N_1, \dots, N_r$, there is a unique solution X of (3.2) up until ζ. For more on this type of representation, see Anderson and Kurtz [2].

3.1. *Assumptions on the chemical reaction network.* Throughout this paper, we consider the continuous-time Markov chain model X for the number of molecules of each species in the chemical reaction network $(\mathscr{S}, \mathscr{R}, \mathscr{C})$, where the propensity function Λ_k, for each $1 \le k \le r$, is given by the stochastic law of mass action (3.1). We suppose in addition that Assumptions 3.1 and 3.2, given below, hold. We begin this subsection with the following definition. For this, recall that the change in the state of X associated with reaction R_k is given by $v_k = v_k^+ - v_k^-$.

DEFINITION 3.1. Let $D \subset \{1, \ldots, r\}$ be nonempty. The set of reactions $\{R_k : k \in D\}$ is said to be:

(a) a *mass-dissipating* set if there exists a vector $u \in \mathbb{R}^d$ such that $u \geq \mathbb{1}$ and $\langle u, v_k \rangle \leq 0$ for all $k \in D$;
(b) an *external input* set if $v_k^- = \mathbb{0}$ for all $k \in D$;
(c) an *external output* set if $v_k^+ = \mathbb{0}$ for all $k \in D$,

where $\mathbb{1}$ is the vector of all ones and $\mathbb{0}$ is the zero vector in $\mathbb{R}^d$.

The following is a fairly standard assumption in chemical reaction network theory.

ASSUMPTION 3.1. The set of reactions $\mathscr{R}$ can be partitioned into sets $\mathscr{R}_1$, $\mathscr{R}_2$, and $\mathscr{R}_3$ such that $\mathscr{R}_1$ is a mass-dissipating set, $\mathscr{R}_2$ is an external input set, and $\mathscr{R}_3$ is an external output set.

REMARK 3.1. The partition of $\mathscr{R}$ in Assumption 3.1 need not be unique, since the external output reactions are also mass-dissipating for any vector $u \geq \mathbb{1}$. In this paper, however, we do not need this partition to be unique as long as it also satisfies Assumption 3.2 below. Clearly, the partition can be made unique if we required no external output reaction to be present in the mass-dissipating set $\mathscr{R}_1$.

Notice that Assumption 3.1 implies that the system does not produce extra "mass" as it evolves in time, other than that which is received from an external input source. In other words, if the effect of external input sources is ignored, the process X can always be found within the half-space given by $\{x \in \mathbb{R}^d : \langle x, u \rangle \leq \langle u, X(0) \rangle\}$. With this assumption, we have the following result (see the Appendix for the proof).

LEMMA 3.1. *Suppose that $(\mathscr{S}, \mathscr{R}, \mathscr{C})$ satisfies Assumption 3.1. Then a.s., X does not explode in finite time.*

The following assumption requires each molecular species to have an input reaction and an output reaction involving only itself. These assumptions are needed for technical reasons in our proofs.

ASSUMPTION 3.2. Let $(\mathscr{S}, \mathscr{R}, \mathscr{C})$ be a chemical reaction network satisfying Assumption 3.1. Suppose that for each $1 \leq i \leq d$:

(a) the set of input reactions $\mathscr{R}_2$ contains a reaction $R_{k_i^\uparrow}$ such that $v_{k_i^\uparrow}^+ = e_i$;
(b) the set of output reactions $\mathscr{R}_3$ contains a reaction $R_{k_i^\downarrow}$ such that $v_{k_i^\downarrow}^- = e_i$.

Recall that e_i denotes the ith element of the canonical basis for the d-dimensional Euclidean space.

Assumption 3.2 is quite often assumed for continuous-flow stirred-tank reactors used in chemical engineering. For biochemical systems, Assumption 3.2(b) is often assumed, since all species eventually degrade or are diluted away. Even when these assumptions do not hold, in practice, one might artificially introduce the missing input and output reactions with very, very small reaction rate constants, so that Assumption 3.2 is satisfied and the added reactions occur so rarely that they have a negligible effect on the dynamics of the system.

From a technical standpoint, Assumption 3.2(a) is used to show that the CLE has well-defined solutions (at least prior to explosion) and to prove tightness of our sequence of jump-diffusion processes. In particular, this assumption implies that the reflection vector field at any point on the boundary of $\mathbb{R}_+^d$ points into the interior of $\mathbb{R}_+^d$. It also ensures that the diffusion matrix Γ will be positive definite. Assumption 3.2(b) prevents explosion of solutions of the CLE, and is used in showing that our sequence of jump-diffusion processes satisfies a compact containment condition. In particular, we use it to show that the drift in the CLE will be negative when dotted with the vector u, given by Definition 3.1(a), whenever the solution of the CLE is far enough from the origin. This further implies that the reflection vector field at the boundary of $\mathbb{R}_+^d$ will point toward the origin, outside of a sufficiently large compact set.

3.2. *The scaled system.* When the number of molecules in the system is high, it is common to represent the state of the system in molecular concentrations rather than the number of molecules. This is often done in order to approximate the system by models whose state descriptor varies continuously with time.

For chemical reaction networks, the concentration of molecules of each species is measured by dividing the number of molecules in the system by Avogadro's number times the volume of the vessel in which the molecules are contained. Let us denote Avogadro's number times the volume of the vessel by V. With no loss of generality, we suppose throughout the paper that $V \geq 1$. The vector of concentrations of each of the d species in the system at time $t \geq 0$ is then given by

$$\bar{X}^V(t) \triangleq \frac{X^V(t)}{V},$$

where X^V is the stochastic process satisfying (3.2), which is now indexed with the superscript V to indicate the volume in which the molecules are contained. Then, by (3.2), the molecular concentrations in the network at time $t \geq 0$ satisfy the following equation:

$$\bar{X}^V(t) = \bar{X}^V(0) + \frac{1}{V} S U^V(t), \tag{3.3}$$

where $\bar{X}^V(0) \triangleq X(0)/V$, U^V is the vector valued stochastic process U, which is now indexed by the superscript V, and its components U_k^V, $1 \le k \le r$, are given by

$$U_k^V(t) \triangleq N_k\left(\int_0^t \Lambda_k^V\big(V\bar{X}^V(s)\big)\,ds\right), \qquad t \ge 0, \tag{3.4}$$

where Λ_k^V denotes the propensity function for reaction R_k, given by (3.1), and the superscript V was added to indicate its dependency on the volume of the vessel in which the reactions occur. This dependency is via the reaction rate constant κ_k^V.

It is useful to scale the propensity functions Λ_k^V, $1 \le k \le r$, so that they represent changes in concentrations rather than the number of molecules. This is done by scaling the reaction rate constants κ_k^V. The standard scaling for κ_k^V, which is associated with reaction R_k of order $[v_k^-] \triangleq \sum_{i=1}^d v_{ik}^-$, is such that $\kappa_k^V = c_k V^{1-[v_k^-]}$, where $c_k > 0$ is a constant independent of V (see for instance [11, 20], [43], Chapter 6, and [2]). We call this constant c_k the *normalized reaction rate constant*. Let us define the lattice $G^V \triangleq \{y/V : y \in \mathbb{Z}_+^d\}$. Then, for $x \in G^V$ and κ_k^V scaled as above, we have that

$$\Lambda_k^V(Vx) = c_k V^{1-[v_k^-]} \prod_{i=1}^d (Vx_i)_{v_{ik}^-} = c_k V \prod_{i=1}^d \prod_{l=0}^{v_{ik}^- -1} (x_i - l/V),$$

with the interpretation that $\prod_{l=0}^{-1} = 1$. Therefore,

$$\Lambda_k^V(Vx) = V\left[c_k \prod_{i=1}^d x_i^{v_{ik}^-} + V^{-1}\epsilon_k^V(x)\right],$$

where $y^0 = 1$ for any $y \ge 0$ and $\epsilon_k^V(x)$ is a multivariate polynomial in the coordinates of x and $1/V$ that is uniformly bounded on compact sets in G^V (as $V \ge 1$ varies), for each $1 \le k \le r$. Notice that the terms $\epsilon_k^V(x)$, $1 \le k \le r$, are nonzero only when more than one molecule of the same species is consumed by reaction k. For instance, if reaction R_k is a second-order reaction consuming two molecules of S_i, we have

$$\Lambda_k^V(Vx) = c_k V^{-1} V x_i (V x_i - 1) = V\big(c_k x_i^2 - V^{-1} c_k x_i\big), \tag{3.5}$$

where the term $\epsilon_k^V(x)$ is given by $-c_k x_i$.

So far, $\Lambda^V(V\cdot)$ has only been defined on G^V, where it is nonnegative. Since $\epsilon_k^V(x)$ is a multivariate polynomial in x and $1/V$, it is well defined for $x \in \mathbb{R}_+^d$ and so we can define $\Lambda_k^V(Vx)$ for all such x. We then define the *scaled propensity function* $\lambda_k^V : \mathbb{R}_+^d \to \mathbb{R}$ by

$$\lambda_k^V(x) \triangleq V^{-1}\Lambda_k^V(Vx) = c_k \prod_{i=1}^d x_i^{v_{ik}^-} + V^{-1}\epsilon_k^V(x), \tag{3.6}$$

for each $x \in \mathbb{R}^d_+$. For large V, the contribution of the term $V^{-1}\epsilon^V_k(x)$ is small. Furthermore, when $\bar{X}^V$ is approximated by a process Z that takes values in $\mathbb{R}^d_+$, the scaled propensity functions $\lambda^V_k(Z)$ may become negative for small values of Z and thereby lose their meaning as rates. For instance, for some $x \in \mathbb{R}^d_+$ with ith component given by $x_i = 1/(2V)$, the value of $\lambda^V_k(x)$, with corresponding $\Lambda^V_k(Vx)$ given by (3.5), is $-c_k/(4V^2)$. Thus, in practice, when approximating $\bar{X}^V$ by a process taking values in all of $\mathbb{R}^d_+$, λ^V_k is usually replaced by the *deterministic rate function* for reaction R_k, $\lambda_k : \mathbb{R}^d_+ \to \mathbb{R}^d_+$, which is defined by

$$\lambda_k(x) \triangleq c_k \prod_{i=1}^{d} x_i^{v_{ik}^-} \qquad \text{for } x \in \mathbb{R}^d_+. \tag{3.7}$$

The name is a reference to the rate in the reaction rate equation (1.1), which is a deterministic dynamical system describing an approximation to $\bar{X}^V$, obtained by letting $V \to \infty$.

4. Motivation and preparation for the constrained Langevin approximation. Let $G \triangleq \mathbb{R}^d_+$ denote the positive orthant of $\mathbb{R}^d$ and let G° and G^b denote its interior and boundary, respectively. The Markov chain $\bar{X}^V$, represented by (3.3), is equivalent in distribution[5] to a Markov chain $\bar{X}^V$ satisfying

$$\bar{X}^V(t) = \bar{X}^V(0) + \frac{1}{V} S\big(U^{V,\circ}(t) + U^{V,b}(t)\big), \tag{4.1}$$

where, for each $\alpha = \circ, b$, $U^{V,\alpha}$ is an r-dimensional stochastic process with components $U^{V,\alpha}_k$, $1 \le k \le r$, given by

$$U^{V,\alpha}_k(t) = N^\alpha_k\left(V \int_0^t \lambda^V_k\big(\bar{X}^V(s)\big)\mathbf{1}_{\{\bar{X}^V(s)\in G^\alpha\}}\, ds\right), \tag{4.2}$$

where $\{N^\alpha_k; 1 \le k \le r, \alpha = \circ, b\}$ are independent unit rate Poisson processes, and λ^V_k is the scaled propensity function given by (3.6). Notice that $U^{V,\circ}_k$ and $U^{V,b}_k$, for $1 \le k \le r$, are equivalent in distribution to the contributions of the process U^V_k, given by (3.4), within the interior of the domain G° and at the boundary G^b, respectively. By defining the centered d-dimensional process $\hat{U}^{V,\circ}$ with components $\hat{U}^{V,\circ}_k$, $1 \le k \le r$, given by

$$\hat{U}^{V,\circ}_k(t) \triangleq U^{V,\circ}_k(t) - V \int_0^t \lambda^V_k\big(\bar{X}^V(s)\big)\mathbf{1}_{\{\bar{X}^V(s)\in G^\circ\}}\, ds, \qquad t \ge 0,$$

[5]Markov chains satisfying (3.3) and (4.1) are equivalent in distribution since they have the same infinitesimal generator and so solve the same local martingale problem, which has a unique solution by Theorem 4.1 of [6], Chapter 6, with the β_l there defined to be zero outside of G^V.

and the vector-valued function $\lambda^V \triangleq (\lambda_1^V, \dots, \lambda_r^V)'$, we may further rewrite (4.1) as follows:

$$\bar{X}^V(t) = \bar{X}^V(0) + \int_0^t S\lambda^V(\bar{X}^V(s))\mathbf{1}_{\{\bar{X}^V(s)\in G^\circ\}}\,ds + \frac{1}{\sqrt{V}}\left(\frac{1}{\sqrt{V}}S\hat{U}^{V,\circ}(t) + \frac{1}{\sqrt{V}}SU^{V,b}(t)\right). \tag{4.3}$$

For V fixed, but large, we seek a diffusion approximation for $\bar{X}^V$. When $\lambda_k^V = \lambda_k$ for all k, Theorem 3.13 of Kurtz [21] provides such an approximation *up until the time of first exit* from a bounded domain whose closure is contained in the interior of G. Under our assumptions, this approximation can be extended until the first time that the boundary of G is reached. We first review that diffusion approximation. We then propose a way to extend this approximation beyond the first time that the boundary of G is reached, using a jump-diffusion process.

4.1. *Approximation inside G.* If $\mathcal{E} \subset G^\circ$ is a bounded domain whose closure is contained in G°, then the closure of $\mathcal{E}$ is a positive distance from the complement of G°, and on $\mathcal{E}$, λ_k is bounded and uniformly Lipschitz continuous, and $\sqrt{\lambda_k}$ is uniformly Lipschitz continuous. Supposing that λ_k^V is replaced by λ_k in the definition of $\bar{X}^V$ (stopped at the boundary of $\mathcal{E}$), then Theorem 3.13 and Lemma 3.7 of Kurtz [21] provide that the Poisson processes used in defining $X^V(\cdot)$ can be defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as independent one-dimensional Brownian motions $\breve{W}_1, \dots, \breve{W}_r$ such that $\bar{X}^V(\cdot \wedge \zeta^{\bar{X}^V})$, where $\zeta^{\bar{X}^V} \triangleq \inf\{t \geq 0 : \bar{X}^V(t) \notin \mathcal{E}\}$, is well approximated by a stopped diffusion process $\breve{Z}$ that satisfies the following for $t \geq 0$:

$$\breve{Z}(t) = \bar{X}^V(0) + \int_0^{t\wedge\zeta^{\breve{Z}}} \sum_{k=1}^r v_k\lambda_k(\breve{Z}(s))\,ds + \frac{1}{\sqrt{V}}\sum_{k=1}^r v_k \int_0^{t\wedge\zeta^{\breve{Z}}} \sqrt{\lambda_k(\breve{Z}(s))}\,d\breve{W}_k(s), \tag{4.4}$$

where $\zeta^{\breve{Z}} \triangleq \inf\{t \geq 0 : \breve{Z}(t) \notin \mathcal{E}\}$ and $\breve{Z}$ is adapted to the filtration generated by $\bar{X}^V(0)$ and $\breve{W}_1, \dots, \breve{W}_r$. This approximation is good in the sense that for each $T > 0$, $V > e$, $\breve{Z}$ and $\bar{X}^V$ can be realized on the same probability space such that

$$\left|\bar{X}^V(t) - \breve{Z}(t)\right| \leq \breve{\Upsilon}\frac{\log V}{V} \qquad \text{for all } 0 \leq t \leq \zeta^{\breve{Z}} \wedge \zeta^{\bar{X}^V}, \tag{4.5}$$

where $\breve{\Upsilon}$ is a nonnegative random variable whose distribution depends on the $\{\lambda_k, k = 1, \dots, r\}$, $\mathcal{E}$, T and $\mathbb{E}[\exp(\breve{c}\breve{\Upsilon})] < \infty$ for some $\breve{c} > 0$. This result of Kurtz is justified using an implied version of a well known strong approximation theorem of Komlós, Major and Tusnády [19], combined with the fact that a continuous local martingale can be time changed to a Brownian motion (see Lemma 3.7 of Kurtz [21]).

By Theorem 4.2 of [18], page 170, the local martingale stochastic integral terms in (4.4) can be combined into a local martingale term having the same covariation process, so that, on a possibly enlarged probability space, the process $\breve{Z}$ is the solution of the following stochastic differential equation for all $t \geq 0$:

$$\breve{Z}(t) = \bar{X}^V(0) + \int_0^{t\wedge\zeta^{\breve{Z}}} \mu(\breve{Z}(s))\,ds + \frac{1}{\sqrt{V}}\int_0^{t\wedge\zeta^{\breve{Z}}} \sigma(\breve{Z}(s))\,dW(s), \tag{4.6}$$

where W is a standard d-dimensional Brownian motion, $\breve{Z}$ is adapted to the filtration generated by $\bar{X}^V(0)$ and W,

$$\mu(x) \triangleq \sum_{k=1}^{r} v_k\lambda_k(x), \qquad x \in G, \tag{4.7}$$

and the dispersion coefficient $\sigma : G \to \mathbb{S}^{d\times d}$ is the symmetric, positive definite square root of Γ defined by

$$\Gamma(x) \triangleq \sum_{k=1}^{r} v_k v_k'\lambda_k(x), \qquad x \in G, \tag{4.8}$$

which satisfies

$$\Bigl(\min_{i=1}^{d} c_{k_i^\uparrow}\Bigr)|\theta|^2 \leq \sum_{i=1}^{d}\theta_i^2 c_{k_i^\uparrow} \leq \langle\theta, \Gamma(x)\theta\rangle \leq K(x)|\theta|^2, \tag{4.9}$$

for all $\theta \in \mathbb{R}^d$, $x \in G$, where $k_i^\uparrow$ is given by Assumption 3.2(a) and $K(x) \triangleq \sum_{k=1}^{r}|v_k|^2\lambda_k(x)$. The function σ is well defined and continuously differentiable, since Γ is continuously differentiable, being defined in terms of multivariate polynomials in the coordinates of x, and the map $A \to A^{1/2}$ is analytic on the set of positive definite $d \times d$ symmetric matrices, by Lemma 5.2.1 of [35], page 131.

Under our assumptions (and with λ_k in place of λ_k^V), by using a sequence of stopping times and showing nonexplosion of the approximating diffusion process, one can extend the aforementioned approximation (with a relaxed error estimate) to where the domain $\mathcal{E}$ is all of G^o. (The requisite nonexplosion is a consequence of estimates that we give in Section 4.4.) However, this only gives an approximation to $\bar{X}^V$ up until the first time that the boundary G^b is reached. In the next subsection, we propose an extension of this approximation, beyond the first hitting time of G^b, by using a jump-diffusion process.

REMARK 4.1. The book of Ethier and Kurtz (see [6], page 459), is often cited in the applied literature to justify a Langevin-type diffusion approximation to a chemical reaction network model. However, that book assumes that the analogues of our λ_k (which are called β_l in [6]) are nonnegative everywhere on all of d-dimensional Euclidean space, not just on the positive orthant, and a Lipschitz continuity condition is usually imposed on $\sqrt{\lambda_k}$ to ensure uniqueness of solutions of the Langevin equation. We have cited the paper [21] here because its statement of the diffusion approximation uses stopping times to restrict to a domain where the requisite nonnegativity and Lipschitz conditions are satisfied.

REMARK 4.2. The process $\breve{Z}$ in (4.6) can also be obtained as a weak limit of a sequence of processes $\{\breve{Z}^{\delta_n}\}$ where δ_n tends to zero as $n \to \infty$ and for any $0 < \delta \le \frac{1}{\sqrt{V}}$, $\breve{Z}^\delta$ satisfies for all $t \ge 0$,

$$\begin{aligned}\breve{Z}^\delta(t) &= \bar{X}^V(0) + \int_0^{t\wedge\zeta^{\breve{Z}^\delta}} \mu(\breve{Z}^\delta(s))\,ds \\ &\quad + \frac{1}{\sqrt{V}}\left(\delta \sum_{k=1}^r v_k \hat{N}_k^\circ\left(\delta^{-2}\int_0^{t\wedge\zeta^{\breve{Z}^\delta}} \lambda_k(\breve{Z}^\delta(s))\,ds\right)\right)\end{aligned}$$

and $\hat{N}_k^\circ(t) = N_k^\circ(t) - t$, $\zeta^{\breve{Z}^\delta} = \inf\{t \ge 0 : \breve{Z}^\delta(s) \notin \mathcal{E}\}$. By comparison with (4.3), we see that when λ_k^V is replaced by λ_k, the equation for $\bar{X}^V(\cdot \wedge \zeta^{\bar{X}^V})$ is the same as that for $\breve{Z}^\delta(\cdot)$ when $\delta = \frac{1}{\sqrt{V}}$. Thus, one can view $\breve{Z}$ as a limit of stopped approximations to $\bar{X}^V$ in which the jump size and jump rate in the fluctuation term in $\bar{X}^V$ have been sent to zero and infinity, respectively, in a manner such that the limit is a diffusion process with fluctuations of order $\frac{1}{\sqrt{V}}$. A similar type of limit, involving rescaling of a boundary process, will be used in proposing a reflected diffusion approximation for the unstopped process $\bar{X}^V$.

4.2. *Jump-diffusion approximation.* Our proposed *jump-diffusion approximation* to $\bar{X}^V$ is a d-dimensional, adapted process $\hat{Z}$ that is defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$, on which there is also defined a standard d-dimensional Brownian motion W that is a martingale with respect to $\{\mathcal{F}_t\}$, a random variable $\bar{X}^V(0)$ that is $\mathcal{F}_0$-measurable, and r independent Poisson processes $\{N_k^b; 1 \le k \le r\}$ that are independent of W and $\bar{X}^V(0)$, such that $\hat{Z}(t) \in G$ for all $t \ge 0$,

$$\begin{aligned}\hat{Z}(t) &= \bar{X}^V(0) + \int_0^t \mu(\hat{Z}(s))\mathbf{1}_{\{\hat{Z}(s)\in G^\circ\}}\,ds \\ &\quad + \frac{1}{\sqrt{V}}\left(\int_0^t \sigma(\hat{Z}(s))\mathbf{1}_{\{\hat{Z}(s)\in G^\circ\}}\,dW(s) + \tilde{Y}(t)\right), \qquad t \ge 0,\end{aligned} \tag{4.10}$$

where

$$\tilde{Y}(t) \triangleq \frac{1}{\sqrt{V}} S\tilde{U}^b(t),$$
$$\tilde{U}_k^b(t) \triangleq N_k^b\left(V \int_0^t \lambda_k(\hat{Z}(s)) \mathbf{1}_{\{\hat{Z}(s)\in G^b, \hat{Z}_i(s)\geq \frac{v_{ik}^-}{V} \text{ for } i=1,\dots,d\}} ds\right),$$

for $1 \leq k \leq r$, and

$$\hat{U}_k^b(t) \triangleq \tilde{U}_k^b(t) - V \int_0^t \lambda_k(\hat{Z}(s)) \mathbf{1}_{\{\hat{Z}(s)\in G^b, \hat{Z}_i(s)\geq \frac{v_{ik}^-}{V} \text{ for } i=1,\dots,d\}} ds,$$
$$\hat{Y}(t) \triangleq \tilde{Y}(t) - \sqrt{V} \sum_{k=1}^r v_k \int_0^t \lambda_k(\hat{Z}(s)) \mathbf{1}_{\{\hat{Z}(s)\in G^b, \hat{Z}_i(s)\geq \frac{v_{ik}^-}{V} \text{ for } i=1,\dots,d\}} ds, \tag{4.11}$$

are local martingales with respect to the filtration $\{\mathcal{F}_t\}$.

Notice that the process $\tilde{Y}$ is defined analogously to the jump process $Y^V(\cdot) \triangleq \frac{1}{\sqrt{V}} SU^{V,b}(\cdot)$ in (4.3), but where λ_k^V is replaced by λ_k and the indicator function of the set $\{\hat{Z}_i(s) \geq \frac{v_{ik}^-}{V}$ for $i = 1, \dots, d\}$ is included in the definition of $\tilde{U}_k^b$. The reason for the latter is that this prevents the process $\hat{Z}$ from jumping to a point that is outside of G. Such a constraint was not needed for $\bar{X}^V$ as it lives on a lattice in G and the form of the rates λ_k^V prevents $\bar{X}^V$ from jumping outside of G from these lattice points. However, $\hat{Z}$ diffuses in G° and so it may reach parts of the boundary where more than one component of $\hat{Z}$ is so small that certain jumps from there would take it outside of G.

Since $\tilde{Y}$ cannot increase until $\hat{Z}$ first reaches the boundary of G, the equation for $\hat{Z}$ and the equation (4.6) for $\breve{Z}$ agree until the boundary of G is first reached. Because of the uniqueness of solutions, it follows that $\hat{Z}$ agrees with $\breve{Z}$ until this time. After this first hitting time, $\hat{Z}$ continues by jumping in an analogous manner to how $\bar{X}^V$ does at the boundary (with λ_k^V replaced by λ_k) and then as soon as $\hat{Z}$ jumps back into the interior of G, it continues on in the same manner as if it had started from there. This defines a strong Markov process $\hat{Z}$ that behaves like $\breve{Z}$ in the interior of G and jumps from the boundary in a similar manner to $\bar{X}^V$, except for a small perturbation in the dynamics when more than one component of $\hat{Z}$ is near zero.

We emphasize that $\hat{Z}$ is a formal approximation that we propose for $\bar{X}^V$, that is based on the idea of extending the approximation of Kurtz beyond the first time that the boundary of G is reached, where the extension is in a manner that is consistent with the original behavior of $\bar{X}^V$ on the boundary of G. We believe this jump-diffusion process $\hat{Z}$ should be a good approximation to $\bar{X}^V$. It would be interesting to have error estimates to confirm this. However, the focus of our attention in this paper is on proposing a continuous reflected diffusion approximation for $\bar{X}^V$. To obtain this, we shall modify the boundary process $\tilde{Y}$ so that the jump sizes are of order δ rather than $1/\sqrt{V}$ and the rate of jumping is of order δ^{-2} rather than V. This is similar to the kind of rescaling described in Remark 4.2 that leads to the approximation of $\bar{X}^V$ by a diffusion in the interior of G. We now define our family of jump-diffusion processes, indexed by the parameter $\delta \in (0, \frac{1}{\sqrt{V}}]$. A version of the process $\hat{Z}$, just described, can be realized by setting $\delta = \frac{1}{\sqrt{V}}$. We will obtain our candidate reflected diffusion approximation for $\bar{X}^V$ by letting $\delta \to 0$.

DEFINITION 4.1 (Family of jump-diffusion processes). For each $\delta \in (0, \frac{1}{\sqrt{V}}]$, we suppose that there is a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ on which there is defined a standard d-dimensional Brownian motion W that is a martingale with respect to $\{\mathcal{F}_t\}$, an $\mathcal{F}_0$-measurable random variable $\bar{X}^V(0)$, independent Poisson processes $\{N_k^b; 1 \le k \le r\}$ that are independent of W and $\bar{X}^V(0)$, and an adapted d-dimensional process Z^δ satisfying for all $t \ge 0$, $Z^\delta(t) \in G$ and

$$\begin{aligned} Z^\delta(t) = {} & \bar{X}^V(0) + \int_0^t \mu\big(Z^\delta(s)\big)\mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\,ds \\ & + \frac{1}{\sqrt{V}}\bigg(\int_0^t \sigma\big(Z^\delta(s)\big)\mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\,dW(s) + \tilde{Y}^\delta(t)\bigg), \end{aligned} \tag{4.12}$$

where

$$\tilde{Y}^\delta(t) \triangleq \delta S \tilde{U}^{\delta,b}(t), \tag{4.13}$$

with $\tilde{U}^{\delta,b} = (\tilde{U}_1^{\delta,b}, \dots, \tilde{U}_r^{\delta,b})'$, where each component $\tilde{U}_k^{\delta,b}$, $1 \le k \le r$, is given by

$$\tilde{U}_k^{\delta,b}(t) \triangleq N_k^b\bigg(\delta^{-2}\int_0^t \tilde{\lambda}_k^\delta\big(Z^\delta(s)\big)\mathbf{1}_{\{Z^\delta(s)\in G^b\}}\,ds\bigg), \tag{4.14}$$

the rates $\tilde{\lambda}_k^\delta$, $1 \le k \le r$, are defined by

$$\tilde{\lambda}_k^\delta(x) \triangleq \lambda_k(x)\mathbf{1}_{\{x_i \ge \frac{\delta v_{ik}^-}{\sqrt{V}} \text{ for } i=1,\dots,d\}}, \qquad x \in G, \tag{4.15}$$

and the compensated processes $\hat{U}^{\delta,b} = (\hat{U}_1^{\delta,b}, \dots, \hat{U}_r^{\delta,b})'$ and $\hat{Y}^\delta = (\hat{Y}_1^\delta, \dots, \hat{Y}_r^\delta)'$ defined for $1 \le k \le r$, $t \ge 0$, by

$$\hat{U}_k^{\delta,b}(t) \triangleq \tilde{U}_k^b(t) - \delta^{-2}\int_0^t \tilde{\lambda}_k^\delta\big(Z^\delta(s)\big)\mathbf{1}_{\{Z^\delta(s)\in G^b\}}\,ds, \tag{4.16}$$

$$\begin{aligned} \hat{Y}^\delta(t) & \triangleq \delta \sum_{k=1}^r v_k \hat{U}_k^{\delta,b}(t) \\ & = \tilde{Y}^\delta(t) - \delta^{-1}\sum_{k=1}^r v_k \int_0^t \tilde{\lambda}_k^\delta\big(Z^\delta(s)\big)1_{\{Z^\delta(s)\in G^b\}}\,ds, \end{aligned} \tag{4.17}$$

are local martingales with respect to the filtration $\{\mathcal{F}_t\}$.

REMARK 4.3. We note, by the right continuity of the sample paths of $\hat{U}^{\delta,b}$, $\hat{Y}^{\delta}$, and W, that the (local)-martingale properties also hold with $\mathcal{F}_t$ replaced by $\mathcal{F}_{t+} \triangleq \bigcap_{s>t} \mathcal{F}_s$, so we may assume that the filtration $\{\mathcal{F}_t\}$ is right continuous. Now, a random variable $\zeta : \Omega \to [0, \infty]$, is a stopping time relative to $\{\mathcal{F}_{t+}\}$ if and only if $\{\zeta < t\} \in \mathcal{F}_t$ for all $t \geq 0$. In particular, the first time that Z^{δ} is in an open set is an $\{\mathcal{F}_{t+}\}$-stopping time.

The process Z^{δ} is well defined as a strong Markov process up until the explosion time,

$$\zeta_{\infty}^{\delta} = \lim_{M\to\infty} \zeta_M^{\delta}, \tag{4.18}$$

where $\zeta_M^{\delta} = \inf\{t \geq 0 : \langle u, Z^{\delta}(t)\rangle > M\}$. (We shall shortly show that this explosion time is $+\infty$ almost surely, and so Z^{δ} is well defined for all time, almost surely.) Indeed, up until the explosion time and prior to hitting the boundary G^b, Z^{δ} can be constructed from a given standard d-dimensional Brownian motion W and initial condition, as a strong solution of a stochastic differential equation driven by W with drift coefficient μ and dispersion coefficient σ. The stochastic differential equation is uniquely solvable until this time because the drift and dispersion coefficients are continuously differentiable on G and the dispersion coefficient is positive definite on G. Once Z^{δ} has reached the boundary G^b, at some point x say, it waits there for an exponentially distributed amount of time with parameter $\delta^{-2}\sum_{k=1}^r \tilde{\lambda}_k^{\delta}(x)$ before jumping to another state. For $k = 1, \dots, r$, with probability

$$p_k^{\delta}(x) = \frac{\tilde{\lambda}_k^{\delta}(x)}{\sum_{\ell=1}^r \tilde{\lambda}_{\ell}^{\delta}(x)}, \tag{4.19}$$

Z^{δ} jumps to $x + \frac{\delta}{\sqrt{V}} v_k$. Then Z^{δ} continues on from there, as a strong Markov process, as if it had started there. Note that Z^{δ} does not get stuck on the boundary and the jump times do not accumulate because $\tilde{\lambda}_k^{\delta}(x) = c_k > 0$ for all $k \in \mathcal{I}$ and $x \in G^b$, where $\mathcal{I}$ denotes the set of external input reactions. By the construction, the strong Markov property of Brownian motion and the exponential holding times on the boundary, Z^{δ} is well defined as a strong Markov process up until the explosion time ζ_{∞}^{δ}. To make Z^{δ} well defined for all time, we define $Z^{\delta}(t) = \partial$ for all $t \geq \zeta_{\infty}^{\delta}$, where ∂ is a "cemetery state" that is disjoint and isolated from G. Let G^{∂} denote G augmented with this isolated state, endowed with the usual augmented topology where $\{\partial\}$ is an open (and closed) set.

The space where the paths of Z^{δ} lie is

$$\begin{aligned} \mathcal{D}^{\partial} = \big\{\omega : [0,\infty) \to G^{\partial}, \omega \text{ is right continuous with finite left limits} \\ \text{at } t \in [0, \zeta_{\infty}), \text{ and } \omega(t) = \partial \text{ for all } t \geq \zeta_{\infty}\big\}, \end{aligned} \tag{4.20}$$

where $\zeta_{\infty} \triangleq \inf\{s \geq 0 : \omega(s) = \partial\}$. We endow this space with the σ-algebra

$$\mathcal{M}^{\partial} = \sigma\{\omega(s) : 0 \leq s < \infty, \omega \in \mathcal{D}^{\partial}\},$$

and filtration

$$\mathcal{M}_t^\partial = \sigma\{\omega(s) : 0 \le s \le t, \omega \in \mathcal{D}^\partial\}, \qquad t \ge 0. \tag{4.21}$$

Although we have defined Z^δ above with initial condition $\bar{X}^V(0)$, for some developments below, we shall need to consider arbitrary starting states $z \in G$ for Z^δ. Thus, for each $z \in G$, if Z^δ starts from $z \in G$, we let P_z^δ denote the probability measure induced on $(\mathcal{D}^\partial, \mathcal{M}^\partial)$ by this process:

$$P_z^\delta(A) = \mathbb{P}(Z^\delta \in A) \qquad \text{for all } A \in \mathcal{M}^\partial. \tag{4.22}$$

We let $\omega(\cdot)$ denote the canonical process on $(\mathcal{D}^\partial, \mathcal{M}^\partial)$. The function $z \to P_z^\delta(A)$ can be shown to be Borel measurable on G for each set $A \in \mathcal{M}^\partial$.

4.3. *Inequalities for μ and $\tilde{\mu}^\delta$.* In this subsection, we obtain some important inequalities involving μ and $\tilde{\mu}^\delta$, where the latter function is defined analogously to μ but with $\tilde{\lambda}_k^\delta$ replacing λ_k. That is, for each $\delta \in (0, \frac{1}{\sqrt{V}}]$, $\tilde{\mu}^\delta : G \to \mathbb{R}^d$ is defined as follows:

$$\tilde{\mu}^\delta(x) \triangleq \sum_{k=1}^r v_k \tilde{\lambda}_k^\delta(x), \qquad x \in G. \tag{4.23}$$

Recall that $\tilde{\lambda}_k^\delta$, $1 \le k \le r$, are defined by (4.15). These inequalities enable us to show nonexplosion of Z^δ, almost surely, and they are also used in the proof of tightness in a later section.

For $1 \le j \le d$, let n_j be the unit normal vector on the face

$$F_j \triangleq \{x \in G : x_j = 0\}. \tag{4.24}$$

For $x \in G^b$, define $I(x) \triangleq \{1 \le j \le d : x \in F_j\}$. The proofs of Lemmas 4.1 and 4.2 below are given in the Appendix.

LEMMA 4.1. *Let $\mathcal{I}$ be the set of indices of external input reactions* [*as in Assumption* 3.2(a)] *and let M^* be the constant defined by*

$$M^* \triangleq \frac{\sum_{k\in\mathcal{I}} c_k |\langle v_k, u\rangle|}{\min_{i=1}^d c_{k_i^\downarrow}} + 1. \tag{4.25}$$

There exists an $\alpha > 0$ such that:

(i) $\langle \mu(x), n_j\rangle \ge \alpha$, *for all $j \in I(x)$ and $x \in G^b$*;
(ii) $\langle \mu(x), -u\rangle \ge \alpha$ *for all $x \in G$ such that $\langle x, u\rangle > M^*$*,

where u is the vector given in Definition 3.1(a).

LEMMA 4.2. *Let u be as in Definition 3.1(a) and define the constant*

$$\tilde{M}^* \triangleq M^* + \frac{1}{V}\sum_{i=1}^{d} u_i, \tag{4.26}$$

where M^* is defined by (4.25). Then, there exists a constant $\alpha > 0$ such that for all $\delta \in (0, \frac{1}{\sqrt{V}}]$:

(i) $\langle \tilde{\mu}^\delta(x), n_j \rangle \geq \alpha$, *for all $j \in I(x)$ and $x \in G^b$;*
(ii) $\langle \tilde{\mu}^\delta(x), -u \rangle \geq \alpha$ *for all $x \in G$ such that $\langle x, u \rangle > \tilde{M}^*$.*

4.4. *Estimates for hitting times by Z^δ.* In the following, the stopping times for Z^δ will be with respect to the right-continuous filtration $\{\mathcal{F}_t\}$ and for $\omega(\cdot)$ will be with respect to the right-continuous filtration $\{\mathcal{M}^\partial_{t+}\}$.

In the next lemma, we develop an estimate which implies that almost surely Z^δ does not explode in finite time for any $\delta \in (0, \frac{1}{\sqrt{V}}]$. For $M > 0$, recall the definition of ζ^δ_M from (4.18). We also define on the canonical space $(\mathcal{D}^\partial, \mathcal{M}^\partial)$, the following stopping time:

$$\zeta_M = \inf\{t \geq 0 : \langle u, \omega(t) \rangle > M\}, \tag{4.27}$$

where $\zeta_M = 0$ if $\omega(t) = \partial$ for all $t \geq 0$. The proof of the lemma is deferred to the Appendix.

LEMMA 4.3. *For each $\delta \in (0, \frac{1}{\sqrt{V}}]$, $t > 0$ and $m > 0$,*

$$\sup_{z \in G : \langle u, z \rangle \leq m} P^\delta_z(\zeta_M \leq t) \to 0 \qquad \text{as } M \to \infty. \tag{4.28}$$

REMARK 4.4. Since the law of Z^δ given $Z^\delta(0) = z$ is P^δ_z for each $z \in G$, it follows immediately from the lemma that almost surely, Z^δ, with initial state $\bar{X}^V(0)$, does not explode in finite time. Consequently, without loss of generality, we may and do assume henceforth that Z^δ surely does not explode in finite time.

The following result will be used in a subsequent section to establish a compact containment condition as part of a proof of tightness for a sequence of processes $\{Z^{\delta_n}\}_{n=1}^\infty$, where $\delta_n \to 0$ as $n \to \infty$. In preparation for the lemma, we make the following definitions. Let

$$\rho = 1 \vee \max_{k=1}^{r} |\langle u, v_k \rangle|. \tag{4.29}$$

Note that, since $\delta \in (0, \frac{1}{\sqrt{V}}]$, $V \geq 1$, and the jumps of Z^δ are of the form $\frac{\delta}{\sqrt{V}} v_k$, the sizes of the jumps of $\langle u, Z^\delta \rangle$ are bounded by $\frac{1}{V}\rho \leq \rho$. We also define for each $m > 0$,

$$\tau^\delta_m = \inf\{t \geq 0 : \langle u, Z^\delta(t) \rangle < m\} \tag{4.30}$$

and the set

$$\Theta_m = \{z \in G : m \le \langle u, z\rangle \le m + \rho\}. \tag{4.31}$$

The proof of the following lemma can be found in the Appendix.

LEMMA 4.4. *For each fixed $m \ge \tilde{M}^* + \rho$ and $t \ge 0$, there is $\delta_0 \in (0, \frac{1}{\sqrt{V}}]$ (depending on m, t), such that*

$$\sup_{0<\delta\le\delta_0} \sup_{z\in\Theta_m} P_z^\delta(\zeta_M \le t) \to 0 \qquad \text{as } M \to \infty. \tag{4.32}$$

5. Definition of the constrained Langevin equation (CLE). Recall the definitions of μ and σ from Section 4.1. Let $\mathcal{S}_1$ denote the unit sphere in $\mathbb{R}^d$ centered at 0, that is $\mathcal{S}_1 \triangleq \{x \in \mathbb{R}^d : |x| = 1\}$. Define a reflection field on the boundary of G, $\gamma : G^b \to \mathcal{S}_1$, by

$$\gamma(x) \triangleq \frac{\mu(x)}{|\mu(x)|} \qquad \text{for } x \in G^b, \tag{5.1}$$

where $|\mu(x)| \ne 0$ since we have that for $x \in G^b$, $|\mu(x)| \ge |\langle \mu(x), n_j\rangle| \ge \alpha$ for $j \in I(x)$, by Lemma 4.1.

In the following, we define the notion of a (weak) solution to the constrained Langevin equation. This is similar to the definition of a (weak) solution of a stochastic differential equation with reflection (SDER) used in [16].

DEFINITION 5.1 (Solution of the CLE). Given a Borel probability measure ϑ on G, a solution of the Constrained Langevin Equation (CLE), with initial distribution ϑ, is a d-dimensional $\{\mathcal{F}_t\}$-adapted process Z defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}^\vartheta)$, on which there are also defined auxiliary processes W and L, such that:

(i) $Z(t) \in G$ for all $t \ge 0$, $\mathbb{P}^\vartheta$-a.s., and the distribution of $Z(0)$ is given by ϑ;

(ii) under $\mathbb{P}^\vartheta$, W is a standard d-dimensional Brownian motion that is a martingale with respect to $\{\mathcal{F}_t\}$;

(iii) L is a continuous, $\{\mathcal{F}_t\}$-adapted, one-dimensional process that is nondecreasing, $\mathbb{P}^\vartheta$-a.s., and satisfies

$$L(t) = \int_0^t \mathbf{1}_{\{Z(s)\in G^b\}}\, dL(s) \qquad \text{for all } t \ge 0, \mathbb{P}^\vartheta\text{-a.s.},$$

that is, $\mathbb{P}^\vartheta$-a.s., L only increases when Z is on G^b;

(iv) the triple (Z, W, L) satisfies equation (1.4) $\mathbb{P}^\vartheta$-a.s.

REMARK 5.1. In addition to the conditions stated in Definition 5.1, it is usually required for solutions of an SDER that the following condition holds:

$$\mathbb{P}^\vartheta\left(\int_0^t |\mu(Z(s))| + |\sigma(Z(s))|^2\, ds < \infty\right) = 1$$

for every $0 \le t < \infty$. However, since the paths of Z are continuous $\mathbb{P}^{\vartheta}$-a.s. and μ, σ are continuous, the above condition is automatically satisfied.

REMARK 5.2. The auxiliary process L is sometimes referred to as a local time process associated with Z. We shall occasionally use this terminology to refer to this auxiliary process. We shall refer to the triple (Z, W, L) as a solution of the CLE along with its auxiliary processes.

Our definition of a solution of the CLE uses a stochastic differential equation with reflection (SDER) approach. An alternative approach to describing such reflected diffusion processes is via submartingale problems. In the recent work [16], Kang and Ramanan give sufficient conditions (which are satisfied for the situation considered in this paper) for the SDER and submartingale problem approaches to be equivalent, in that they both characterize the law of the process given the initial condition. While we do not use the submartingale problem approach here, it can be useful, for example, for characterizing the stationary distribution of a solution of an SDER, as seen in [16].

6. Pathwise and weak uniqueness for the CLE. In this section, we first prove pathwise uniqueness for the CLE. Our proof relies on the work of Dupuis and Ishii [5], which gives sufficient conditions for pathwise uniqueness, and then for strong existence and uniqueness of solutions, for certain SDERs on bounded domains. (In fact, we use a slight modification of the pathwise uniqueness argument in [5], as we explain in the proof of our Theorem 6.1 below.) In order to use the work of [5], which is for bounded domains, we let

$$G_M \triangleq \{x \in G : \langle x, u\rangle \le M\}, \tag{6.1}$$

for each $M > M^*$, where M^* is defined in (4.25), and u is the vector associated with the mass-dissipating reactions of Assumption 3.1. The vector n_j is the unit normal vector on the face $F_j^M \triangleq \{x \in G_M : x_j = 0\}$, for $j = 1, \dots, d$. Let us define in addition $n_{d+1} \triangleq -u/|u|$ to be the inward unit normal vector on the face $F_{d+1}^M \triangleq \{x \in G_M : \langle x, u\rangle = M\}$. Let G_M^b denote the boundary of G_M and let $I_M(x) \triangleq \{j : 1 \le j \le d+1, x \in F_j^M\}$, for $x \in G_M^b$. Then we have the following result.

LEMMA 6.1. *Fix $M > M^*$. Then there are functions $\mu_M : \mathbb{R}^d \to \mathbb{R}^d$, $\Gamma_M : \mathbb{R}^d \to \mathbb{S}^{d\times d}$ that coincide with μ, Γ, respectively, on G_M, and $\gamma_M : \mathbb{R}^d \to \mathbb{R}^d$ that coincides with γ on $\tilde{G}_M^b \triangleq \{x \in G_M^b : x \notin G^\circ\}$, such that*:

(i) *$\langle \theta, \Gamma_M(x)\theta\rangle \ge k_M|\theta|^2$, for all $x, \theta \in \mathbb{R}^d$, for some constant $k_M > 0$;*

(ii) *$|\mu_M(x) - \mu_M(y)| \vee |\sigma_M(x) - \sigma_M(y)| \le K_M|x - y|$, for all $x, y \in \mathbb{R}^d$, for some constant $K_M > 0$, where $\sigma_M(x)$ is the symmetric positive definite square root of $\Gamma_M(x)$ for $x \in \mathbb{R}^d$;*

(iii) *γ_M is twice continuously differentiable, $|\gamma_M| = 1$ in a neighborhood of G^b_M, γ_M is a constant vector outside of a compact set, and there is a $\beta \in (0, 1)$ such that*

$$\bigcup_{0\le t\le\beta} \mathbb{B}_{t\beta}\big(x - t\gamma_M(x)\big) \subset \big(G^\circ_M\big)^c \qquad \text{for } x \in G^b_M,$$

where $\mathbb{B}_p(x) \triangleq \{y \in \mathbb{R}^d : |x - y| \le p\}$ for $x \in \mathbb{R}^d$ and $p \in \mathbb{R}_+$, $(G^\circ_M)^c = \mathbb{R}^d \setminus G^\circ_M$, and G°_M is the interior of G_M.

The proof of Lemma 6.1 is given in the Appendix. We now present the main theorem of this section, which states that pathwise uniqueness holds for the CLE. As we will show in a subsequent corollary, this implies weak uniqueness for solutions of the CLE, which is actually the result we shall use in the next section.

THEOREM 6.1. *Fix a Borel probability measure ϑ on G. Suppose that $(Z^{(1)}, W, L^{(1)})$ and $(Z^{(2)}, W, L^{(2)})$ are two solutions to the CLE (together with their associated auxiliary processes) defined on the same filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}^\vartheta)$, with the same initial condition, so that $Z^{(1)}(0) = Z^{(2)}(0) = \xi$ $\mathbb{P}^\vartheta$-a.s., where ξ is a random variable taking values in G with distribution ϑ. (Note that the Brownian motion W is common to the two solutions.) Then $(Z^{(1)}, L^{(1)})$ and $(Z^{(2)}, L^{(2)})$ are indistinguishable under $\mathbb{P}^\vartheta$, that is,*

$$\mathbb{P}^\vartheta\big(\big(Z^{(1)}, L^{(1)}\big)(t) = \big(Z^{(2)}, L^{(2)}\big)(t) \text{ for all } t \ge 0\big) = 1. \tag{6.2}$$

PROOF. For a fixed $M > M^*$, let μ_M, Γ_M, σ_M and γ_M be the functions described in Lemma 6.1. Define

$$\check{\tau}_M \triangleq \inf\{t \ge 0 : \langle u, Z^{(1)}(t)\rangle > M \text{ or } \langle u, Z^{(2)}(t)\rangle > M\}.$$

Then, since μ, Γ, σ coincide with μ_M, Γ_M, σ_M on G_M, and γ coincides with γ_M on $\tilde{G}^b_M$, $\mathbb{P}^\vartheta$-a.s. on $\{\xi \in G_M\}$, up until the time $\check{\tau}_M$, $(Z^{(1)}, W, L^{(1)})$ and $(Z^{(2)}, W, L^{(2)})$ solve an SDER on G_M, with drift and dispersion coefficients given by μ_M and $\sigma_M/\sqrt{V}$, respectively, with reflection vector field given by γ_M on the boundary G^b_M of G_M, and with initial condition given by ξ. In particular, if we let $\check{Z}^{(1)}, \check{L}^{(1)}, \check{Z}^{(2)}, \check{L}^{(2)}, \check{W}, \check{\xi}$ be defined for $t \ge 0$ by

$$\check{Z}^{(1)}(t) \triangleq Z^{(1)}(t \wedge \check{\tau}_M)\mathbf{1}_{\{\xi\in G_M\}}, \qquad \check{L}^{(1)}(t) \triangleq L^{(1)}(t \wedge \check{\tau}_M)\mathbf{1}_{\{\xi\in G_M\}},$$
$$\check{Z}^{(2)}(t) \triangleq Z^{(2)}(t \wedge \check{\tau}_M)\mathbf{1}_{\{\xi\in G_M\}}, \qquad \check{L}^{(2)}(t) \triangleq L^{(2)}(t \wedge \check{\tau}_M)\mathbf{1}_{\{\xi\in G_M\}},$$
$$\check{W}(t) \triangleq W(t \wedge \check{\tau}_M)\mathbf{1}_{\{\xi\in G_M\}}, \qquad \check{\xi} \triangleq \xi\mathbf{1}_{\{\xi\in G_M\}},$$

then equation (5.3) in [5] holds $\mathbb{P}^\vartheta$-a.s. with $(\check{Z}^{(1)}, \check{Z}^{(1)}, \check{W}, \frac{1}{\sqrt{V}}\check{L}^{(1)}, \check{\xi}, t \wedge \check{\tau}_M, \mu_M, \frac{\sigma_M}{\sqrt{V}}, \gamma(\check{Z}^{(1)}))$ [resp., $(\check{Z}^{(2)}, \check{Z}^{(2)}, \check{W}, \frac{1}{\sqrt{V}}\check{L}^{(2)}, \check{\xi}, t \wedge \check{\tau}_M, \mu_M, \frac{\sigma_M}{\sqrt{V}}, \gamma(\check{Z}^{(2)}))$] in place of $(Y, X, M, |k|, x, t, b, \sigma, \gamma)$ there.

We would like to use the result of Theorem 5.1 in [5], to establish the desired uniqueness. It turns out that we need to make a slight modification to the argument in [5], in order to obtain the result. The reason is revealed when one attempts to check the conditions in [5]. First, we can complete the probability space and augment the filtration so that the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}^{\vartheta})$ satisfies the usual conditions assumed in Section 5 of [5]. Furthermore, by Lemma 6.1, the drift, dispersion and domain, μ_M, σ_M, and G_M°, satisfy the conditions of Section 5 and Case 1 in [5]. An issue arises with the vector field γ_M, which satisfies the conditions of Case 1 of [5], except that it is only of unit length in a neighborhood of G_M^b, whereas the paper [5], and the paper [4] on which [5] relies for the existence of the functions f_ϵ and h in Theorem 3.2 (Case 1) of [5], assumes γ_M is of unit length *everywhere* on $\mathbb{R}^d$. (In fact, due to homotopy considerations,[6] one cannot extend γ_M continuously from G_M^b to $\mathbb{R}^d$ such that it is nonzero everywhere on $\mathbb{R}^d$.) However, this issue can be resolved as follows. (We are grateful to Paul Dupuis for a private communication related to resolving this issue.)

Scrutiny of the proof of Case 1 of Theorem 5.1 in [5] reveals what is needed for this proof to apply is the existence of functions f_ϵ and h satisfying Case 1 of Theorem 3.2 in [5], where (3.15) there only need hold for $x \in G_M^b$ and (3.16) there only need hold for $y \in G_M^b$. In fact, the function h is readily seen to exist since it can be made zero off a neighborhood of G_M^b. The function f_ϵ can be defined using a localization of the procedure used to define the function w_ϵ on pages 1136–1137 of [4]. In particular, we can define $f_\epsilon = \epsilon \tilde{h}(x, \frac{x-y}{\epsilon})$, for

$$\tilde{h}(x,p) = 1 + \psi_{\eta/2}(x)\left(\frac{g(p,\gamma_M(x))}{\theta}\right) + \big(1-\psi_{\eta/2}(x)\big)\langle p,p\rangle \tag{6.3}$$
$$\text{for } x \in G,\ p \in \mathbb{R}^d,$$

where $\psi_{\eta/2}$ is as in our construction of γ_M [see (A.61)] in our proof of Lemma 6.1 in the Appendix, and g, θ are as defined in Lemma 4.4 of [4] (with a suitable value of δ there related to θ in Theorem 3.2 of [5]). The function $g(p,\xi)$ is a function of $p \in \mathbb{R}^d$ and ξ lying on the unit sphere in $\mathbb{R}^d$. The second term in the right member of (6.3) is well defined since it is defined to be zero where $\psi_{\eta/2}$ is zero, and on the support of $\psi_{\eta/2}$ [which is contained in the set $G_M^{b,\eta/2}$ defined by (A.59)], the vector field γ_M has unit length. We note that $\psi_{\eta/2}$ is equal to one on $G_M^{b,\eta/4}$. The function f_ϵ satisfies the conditions (3.13)–(3.19) in Case 1 of Theorem 3.2 and can be used in the proof of Theorem 5.1 in [5]. (We note in passing that this same line of reasoning shows that the results for Case 1 in [5] only require the vector field to be of unit length on the boundary of the domain.)

[6]The simplest example illustrating the problem occurs when G is the unit interval $(0,1)$ and the associated vector field γ is such that $\gamma(0) = 1$ and $\gamma(1) = -1$; any continuous extension of γ to $\mathbb{R}$ must cross 0 at some point $x \in G$.

With the above modification, it follows by the proof of Theorem 5.1 (Case 1) in [5], that for each $T \geq 0$, there is a constant $C_T > 0$ such that for $0 \leq t \leq T$,

$$\begin{aligned} &\mathbb{E}^{\vartheta}\Big[\sup_{0\leq s\leq t}\big|\check{Z}^{(1)}(s) - \check{Z}^{(2)}(s)\big|^2\Big] \\ &\quad \leq C_T \int_0^t \mathbb{E}^{\vartheta}\Big[\sup_{0\leq \upsilon\leq s}\big|\check{Z}^{(1)}(\upsilon) - \check{Z}^{(2)}(\upsilon)\big|^2\Big]\,ds, \end{aligned} \tag{6.4}$$

where $\mathbb{E}^{\vartheta}$ denotes expectation with respect to $\mathbb{P}^{\vartheta}$ and we used the fact that $\check{Z}^{(1)}(0) = \check{Z}^{(2)}(0) = \check{\xi}$, $\mathbb{P}^{\vartheta}$-a.s. (although the stochastic processes in Theorem 5.1 of [5] are not stopped and the initial condition is constant, rather than random as we have here, Theorem 5.1 in [5] still applies. Indeed, the proof of Theorem 5.1 on page 575 of [5] still holds with t there replaced by $t \wedge \check{\tau}_M$ and, since the initial condition is almost surely the same for $\check{Z}^{(1)}$ and $\check{Z}^{(2)}$, these conditions cancel one another.). Using Gronwall's inequality, and because $T \geq 0$ was arbitrary, we conclude from (6.4) that

$$\mathbb{P}^{\vartheta}\big(\check{Z}^{(1)}(t) = \check{Z}^{(2)}(t) \text{ for all } t \geq 0\big) = 1.$$

Rewriting this in terms of the original processes $Z^{(1)}$, $Z^{(2)}$, and using the fact that on $\{\xi \notin G_M\}$ we have $\check{\tau}_M = 0$ and $Z^{(1)}(0) = Z^{(2)}(0)$, $\mathbb{P}^{\vartheta}$-a.s., we see that

$$\mathbb{P}^{\vartheta}\big(Z^{(1)}(t \wedge \check{\tau}_M) = Z^{(2)}(t \wedge \check{\tau}_M) \text{ for all } t \geq 0\big) = 1. \tag{6.5}$$

Since $Z^{(1)}$ and $Z^{(2)}$ have continuous paths in G for all time, it follows that $\check{\tau}_M \to \infty$ $\mathbb{P}^{\vartheta}$-a.s. as $M \to \infty$. Therefore, letting $M \to \infty$ in (6.5) yields

$$\mathbb{P}^{\vartheta}\big(Z^{(1)}(t) = Z^{(2)}(t) \text{ for all } t \geq 0\big) = 1. \tag{6.6}$$

It follows from this and the CLE satisfied by $Z^{(1)}$ (resp., $Z^{(2)}$) that we also have $\mathbb{P}^{\vartheta}$-a.s., for all $t \geq 0$,

$$\mathcal{V}^{(1)}(t) \triangleq \int_0^t \gamma\big(Z^{(1)}(s)\big)\,dL^{(1)}(s) = \int_0^t \gamma\big(Z^{(2)}\big)(s)\,dL^{(2)}(s) \triangleq \mathcal{V}^{(2)}(t). \tag{6.7}$$

Combining (6.6)–(6.7) with the facts that γ is a unit length vector field on G^b, and $L^{(1)}$ (resp., $L^{(2)}$) can increase only when $Z^{(1)}$ (resp., $Z^{(2)}$) is on G^b, we have $\mathbb{P}^{\vartheta}$-a.s., for all $t \geq 0$,

$$\begin{aligned} L^{(1)}(t) &= \int_0^t \big\langle\gamma\big(Z^{(1)}(s)\big), \gamma\big(Z^{(1)}(s)\big)\big\rangle\,dL^{(1)}(s) = \int_0^t \gamma\big(Z^{(1)}(s)\big)\,d\mathcal{V}^{(1)}(s) \\ &= \int_0^t \gamma\big(Z^{(2)}(s)\big)\,d\mathcal{V}^{(2)}(s) = \int_0^t \big\langle\gamma\big(Z^{(2)}(s)\big), \gamma\big(Z^{(2)}(s)\big)\big\rangle\,dL^{(2)}(s) \\ &= L^{(2)}(t). \end{aligned} \tag{6.8}$$

The desired result (6.2) then follows from (6.6) and (6.8). $\square$

COROLLARY 6.1. *Fix a Borel probability measure ϑ on G. Suppose that (Z, W, L) is a solution of the CLE (together with its auxiliary processes) with initial distribution ϑ. Then the law of (Z, L) is unique.*

PROOF. Since we have proven pathwise uniqueness in the previous theorem and the corollary assumes the existence of a solution with initial distribution ϑ, the result follows by a standard argument due to Yamada and Watanabe, which usually is summarized as "pathwise uniqueness implies uniqueness in law." Although the original argument of Yamada and Watanabe was given for stochastic differential equations without reflection, the argument is quite general and easily extends to the situation with reflection. See, for example, Step 1 of the proof of Theorem V.17.1 in [31]. In that proof, one replaces (X, W) there with our (Z, L, W). □

7. Weak convergence of a sequence of jump-diffusion processes. This section is devoted to showing that the jump-diffusion process Z^δ of Definition 4.1 converges in distribution to a solution of the constrained Langevin equation as $\delta \downarrow 0$ through a monotonic decreasing sequence $\{\delta_n\}_{n=1}^\infty$ of positive real numbers. We begin by showing, in Section 7.1, that such a sequence $\{Z^{\delta_n}\}$ of jump-diffusion processes, along with certain interior processes and boundary processes, is $\mathcal{C}$-tight. The results are based on the conditions for $\mathcal{C}$-tightness given by Kang and Williams [17] for approximately reflected processes in domains with piecewise smooth boundaries. Next, in Section 7.2, we prove that any weak limit point of the $\mathcal{C}$-tight sequence yields a solution of the CLE. Using uniqueness in law for solutions of the CLE, we then conclude that $\{Z^{\delta_n}\}$ converges weakly to a solution of the CLE.

7.1. *Tightness*. Let $\{\delta_n\}_{n=1}^\infty$ be a monotonic decreasing sequence of positive real numbers such that $\delta_1 \le 1/\sqrt{V}$ and $\delta_n \to 0$ as $n \to \infty$. Recall the definition of $\tilde{\lambda}_k^{\delta_n}$, for $1 \le k \le r$, given by (4.15) and that of $\tilde{\mu}_k^{\delta_n}$ given by (4.23). For notational convenience, we write $\tilde{\lambda}_k^n$ instead of $\tilde{\lambda}_k^{\delta_n}$ and also write $\tilde{\mu}^n$ instead of $\tilde{\mu}^{\delta_n}$. For later reference, we define the vector field $\tilde{\gamma}^n : G^b \to \mathcal{S}_1$ by

$$\tilde{\gamma}^n(x) \triangleq \frac{\tilde{\mu}^n(x)}{|\tilde{\mu}^n(x)|} \qquad \text{for } x \in G^b. \tag{7.1}$$

Notice that $\tilde{\gamma}^n(x)$ is well defined for $x \in G^b$, since $|\tilde{\mu}^n(x)| > 0$ for all such x, by Lemma 4.2. The following lemma establishes a property involving the vector fields $\tilde{\gamma}^n$ and γ, where $\gamma : G^b \to \mathcal{S}_1$ is given by (5.1). The proof is in the Appendix.

LEMMA 7.1. *For each $M > 1$, the functions $\gamma, \tilde{\gamma}^n : G^b \to \mathcal{S}_1$ satisfy the following condition: for each $\varepsilon > 0$ there is a $n_\varepsilon > 1$, which may depend on M, such* *that*

$$\left|\gamma(x) - \tilde{\gamma}^n(x)\right| < \varepsilon,$$

for all $n \geq n_\varepsilon$ *and* $x \in G^b$ *such that* $|x| \leq M$.

Recall the definition of the stochastic processes Z^{δ_n}, $\hat{U}^{\delta_n,b}$ and $\hat{Y}^{\delta_n}$, from Definition 4.1. For notational convenience, we shall simply denote these by $Z^n \equiv Z^{\delta_n}$, $\hat{U}^{n,b} \equiv \hat{U}^{\delta_n,b}$ and $\hat{Y}^n \equiv \hat{Y}^{\delta_n}$, where the superscript δ_n was replaced by n. The family $\{(Z^n, \hat{U}^{n,b}, \hat{Y}^n)\}_{n=1}^{\infty}$ of processes can be defined on a single probability space $(\Omega, \mathcal{F}, \mathbb{P})$, on which the driving Brownian motion W and Poisson processes N_k^b, $k = 1, \dots, r$ are defined. For each n, let $\{\mathcal{F}_t^n\}$ be the filtration defined by $\mathcal{F}_t^n = \mathcal{G}_{t+}^n$ where $\mathcal{G}_t^n \triangleq \sigma\{Z^n(s), W(s), \hat{U}^{n,b}(s), \hat{Y}^n(s), 0 \leq s \leq t\}$ for $t \geq 0$. The processes W, $\hat{U}^{n,b}$ and $\hat{Y}^n$ are local martingales with respect to $\{\mathcal{F}_t^n\}$, for each n.

We now rewrite the equation (4.12) satisfied by Z^n, using (4.12)–(4.17), (4.23) and (7.1), as follows:

$$Z^n(t) = \mathcal{X}^n(t) + \mathcal{Y}^n(t) + \frac{1}{\sqrt{V}} \int_0^t \tilde{\gamma}^n\big(Z^n(s)\big)\, dL^n(s), \qquad t \geq 0, \tag{7.2}$$

where for $t \geq 0$,

$$\begin{aligned} \mathcal{X}^n(t) \triangleq{}& \bar{X}^V(0) + \int_0^t \mu\big(Z^n(s)\big)\mathbf{1}_{\{Z^n(s) \in G^\circ\}}\, ds \\ &+ \frac{1}{\sqrt{V}} \int_0^t \sigma\big(Z^n(s)\big)\mathbf{1}_{\{Z^n(s) \in G^\circ\}}\, dW(s), \end{aligned} \tag{7.3}$$

$$\mathcal{Y}^n(t) = \frac{\delta_n}{\sqrt{V}} \sum_{k=1}^r v_k \hat{U}_k^{n,b}(t), \tag{7.4}$$

$$\begin{aligned} \hat{U}_k^{n,b}(t) ={}& N_k^b\left(\delta_n^{-2} \int_0^t \tilde{\lambda}_k^n\big(Z^n(s)\big)\mathbf{1}_{\{Z^n(s) \in G^b\}}\, ds\right) \\ &- \delta_n^{-2} \int_0^t \tilde{\lambda}_k^n\big(Z^n(s)\big)\mathbf{1}_{\{Z^n(s) \in G^b\}}\, ds, \end{aligned} \tag{7.5}$$

and

$$L^n(t) \triangleq \frac{1}{\delta_n} \int_0^t \left|\tilde{\mu}^n\big(Z^n(s)\big)\right| \mathbf{1}_{\{Z^n(s) \in G^b\}}\, ds. \tag{7.6}$$

For completeness, we include the following proposition, which will be used to verify that a sequence of processes is $\mathcal{C}$-tight.

PROPOSITION 7.1. *Suppose* $\ell \geq 1$ *and that for each* n, ς^n *is an* ℓ*-dimensional stochastic process defined on* $(\Omega, \mathcal{F}, \mathbb{P})$. *The sequence* $\{\varsigma^n\}$ *is* $\mathcal{C}$*-tight if and only if the following two conditions hold*:

(i) *for each* $T > 0$, $\epsilon > 0$, *there are* $K > 0$ *and* $n_0 > 0$ *such that*

$$\mathbb{P}\Big(\sup_{0\le t\le T}\big|\varsigma^n(t)\big| > K\Big) \le \epsilon \qquad \textit{for all } n \ge n_0,$$

(ii) *for each* $T > 0$, $\epsilon > 0$ *and* $\eta > 0$, *there are* $\nu \in (0, T)$ *and* $n_0 > 0$ *such that*

$$\mathbb{P}\big(w_T(\varsigma^n, \nu) > \eta\big) \le \epsilon \qquad \textit{for all } n \ge n_0,$$

where for $\omega \in \mathcal{D}$,

$$w_T(\omega, \nu) \triangleq \sup\Big\{\sup_{\upsilon_1,\upsilon_2\in[t,t+\nu]}\big|\omega(\upsilon_2) - \omega(\upsilon_1)\big| : 0 \le t < t + \nu \le T\Big\}.$$

PROOF. See, for example, Proposition 3.26 of Chapter VI in [15], page 351. □

REMARK 7.1. The first condition in Proposition 7.1 is usually referred to as a compact containment condition.

LEMMA 7.2. *The sequence of processes* $\{Z^n\}$ *satisfies the compact containment condition* (i) *of Proposition* 7.1.

PROOF. Fix $T > 0$ and $\epsilon > 0$. Recall the vector u from Assumption 3.1. Since $\bar{X}^V(0)$ is a finite random variable, there is $K_\epsilon > 0$ such that

$$\mathbb{P}\big(\langle u, \bar{X}^V(0)\rangle > K_\epsilon\big) < \epsilon/2.$$

Note that since $u \ge \mathbb{1}$, for $z \in G$, $\langle u, z\rangle \ge \sum_{i=1}^d z_i \ge |z|$. Thus, if $|Z^n(t)| > K$ for some $K > 0$, then $\langle u, Z^n(t)\rangle > K$ as well. Now suppose $K > m + \rho$ where $m = \max(K_\epsilon, \tilde{M}^* + \rho)$ and ρ is given by (4.29). We use the notation of Section 4 with the superscript n in place of δ_n; in particular, $\zeta_m^n = \inf\{s \ge 0 : \langle u, Z^n(s)\rangle > m\}$. Note that, on $\{\langle u, \bar{X}^V(0)\rangle \le K_\epsilon\}$, if $|Z^n|$ exceeds the level K by time T, then Z^n must reach $\Theta_m = \{x \in G : m \le \langle u, x\rangle \le m + \rho\}$ by time T and the supremum of $\{\langle u, Z^n(\zeta_m^n + t)\rangle, 0 \le t \le T\}$ will exceed K. Consequently, using the strong Markov property of Z^n, we have

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{0\le t\le T}\big|Z^n(t)\big| > K\Big) \\
&\quad\le \mathbb{P}\big(\langle u, \bar{X}^V(0)\rangle > K_\epsilon\big) + \mathbb{P}\Big(\langle u, \bar{X}^V(0)\rangle \le K_\epsilon, \zeta_m^n \le T, \sup_{0\le t\le T}\big|Z^n(t)\big| > K\Big) \\
&\quad\le \frac{\epsilon}{2} + \mathbb{P}\Big(\langle u, \bar{X}^V(0)\rangle \le K_\epsilon, \zeta_m^n \le T, \sup_{0\le t\le T}\langle u, Z^n(\zeta_m^n + t)\rangle > K\Big) \qquad (7.7)\\
&\quad\le \frac{\epsilon}{2} + \mathbb{E}\big[\mathbf{1}_{\{\langle u, \bar{X}^V(0)\rangle\le K_\epsilon, \zeta_m^n\le T\}} P^n_{Z^n(\zeta_m^n)}(\zeta_K \le T)\big].
\end{aligned}$$

Let n_0 be such that $\delta_n \le \delta_0$ for all $n \ge n_0$, where δ_0 (depending on m and T) is defined in Lemma 4.4. Then, by that lemma, we can choose K sufficiently large that

$$\sup_{z\in\Theta_m} P_z^n(\zeta_K \le T) < \frac{\epsilon}{2} \tag{7.8}$$

for all $n \ge n_0$. Since $Z^n(\zeta_m^n) \in \Theta_m$ on $\{\langle u, \bar{X}^V(0)\rangle \le K_\epsilon\}$, it then follows from (7.7) and (7.8) that

$$\mathbb{P}\Big(\sup_{0\le t\le T} |Z^n(t)| > K\Big) < \epsilon$$

for all $n \ge n_0$. $\square$

THEOREM 7.1. *For each n, let $\varsigma^n = (Z^n, \mathcal{X}^n, \mathcal{Y}^n, L^n)$. The sequence of $(3d+1)$-dimensional processes $\{\varsigma^n\}$ is $\mathcal{C}$-tight.*

PROOF. For $T > 0$, $\epsilon > 0$, $\eta > 0$, $\nu > 0$, $M > 0$, $K > 0$, recalling the definition of G_M from (6.1), we have

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{0\le t\le T} |\varsigma^n(t)| > K\Big)\\
&\quad\le \mathbb{P}\Big(\sup_{0\le t\le T} \langle u, Z^n(t)\rangle > M\Big) + \mathbb{P}\Big(\sup_{0\le t\le T} \langle u, Z^n(t)\rangle \le M, \sup_{0\le t\le T} |\varsigma^n(t)| > K\Big)\\
&\quad\le \mathbb{P}\Big(\sup_{0\le t\le T} |Z^n(t)| > \frac{M}{|u|}\Big) + \mathbb{P}\Big(\zeta_M^n \ge T, \sup_{0\le t\le T} |\varsigma^n(t\wedge \zeta_M^n)| > K\Big)\\
&\quad\le \mathbb{P}\Big(\sup_{0\le t\le T} |Z^n(t)| > \frac{M}{|u|}\Big) + \mathbb{P}\Big(Z^n(0) \in G_M, \sup_{0\le t\le T} |\varsigma^n(t\wedge \zeta_M^n)| > K\Big)
\end{aligned}$$

and similarly,

$$\begin{aligned}
&\mathbb{P}(w_T(\varsigma^n, \nu) > \eta)\\
&\qquad\le \mathbb{P}\Big(\sup_{0\le t\le T} |Z^n(t)| > \frac{M}{|u|}\Big) + \mathbb{P}\big(Z^n(0) \in G_M, w_T(\varsigma^n(\cdot\wedge \zeta_M^n), \nu) > \eta\big).
\end{aligned}$$

By Lemma 7.2, there is $M_\epsilon > \tilde{M}^*$ and $n_\epsilon > 0$ such that

$$\mathbb{P}\Big(\sup_{0\le t\le T} |Z^n(t)| > \frac{M_\epsilon}{|u|}\Big) \le \frac{\epsilon}{2} \qquad \text{for all } n \ge n_\epsilon.$$

Thus, for all $n \ge n_\epsilon$, we have

$$\mathbb{P}\Big(\sup_{0\le t\le T} |\varsigma^n(t)| > K\Big) \le \frac{\epsilon}{2} + \mathbb{P}\Big(Z^n(0) \in G_{M_\epsilon}, \sup_{0\le t\le T} |\varsigma^n(t\wedge \zeta_{M_\epsilon}^n)| > K\Big), \tag{7.9}$$

and

$$\mathbb{P}\big(w_T(\varsigma^n,\nu)>\eta\big)\le\frac{\epsilon}{2}+\mathbb{P}\big(Z^n(0)\in G_{M_\epsilon},w_T\big(\varsigma^n(\cdot\wedge\zeta^n_{M_\epsilon}),\nu\big)>\eta\big). \tag{7.10}$$

Let $\tilde{\varsigma}^n(\cdot)=\varsigma^n(\cdot\wedge\zeta^n_{M_\epsilon})\mathbf{1}_{\{Z^n(0)\in G_{M_\epsilon}\}}$. To estimate the last probabilities in (7.9) and (7.10), we consider $Z^n(\cdot\wedge\zeta^n_{M_\epsilon})\mathbf{1}_{\{Z^n(0)\in G_{M_\epsilon}\}}$ as living in the bounded polyhedron G_M for a suitable $M>M_\epsilon$. Then we can apply a $\mathcal{C}$-tightness criterion developed by Kang and Williams [17] for approximations of reflected processes in domains with piecewise smooth boundary, to prove that $\{\tilde{\varsigma}^n\}$ is $\mathcal{C}$-tight. We can then use this to estimate the last terms in (7.9) and (7.10). The result we will use from [17] is Theorem 4.2. This requires that various assumptions on the domain and directions of reflection be satisfied (see Assumptions (A1)–(A5) in [17]) as well as by the processes $\tilde{\varsigma}^n$ (see Assumption 4.1 in [17]).

We first make the choice of M for G_M and define faces and directions of reflection associated with G_M. We also develop several properties of these quantities. Then we verify that Assumptions (A1)–(A5) in [17] hold.

Truncated State Space G_M, Faces of G_M and Reflection Fields $\{\hat{\gamma}^i:i=1,\dots,d+1\}$. Let $M=M_\epsilon+2\rho$ be fixed, where ρ is given by (4.29). Then

$$G_M=\big\{x\in G:\langle x,u\rangle\le M\big\},$$

and we define

$$\begin{aligned}F_i^M&=\{x\in G_M:x_i=0\},\qquad i=1,\dots,d,\\ F_{d+1}^M&=\big\{x\in G_M:\langle x,u\rangle=M\big\},\\ n_{d+1}&=-u/|u|,\end{aligned}$$

and let G_M^b denote the boundary of G_M. Then $G_M^b=\bigcup_{i=1}^{d+1}F_i^M$.

By Lemmas 4.1 and 4.2, there is $\alpha>0$ such that for $i=1,\dots,d$,

$$\big\langle n_i,\mu(x)\big\rangle\ge\alpha,\qquad\big\langle n_i,\tilde{\mu}^n(x)\big\rangle\ge\alpha\qquad\text{for each }x\in F_i^M,n=1,2,\dots, \tag{7.11}$$

and

$$\begin{aligned}\big\langle n_{d+1},\mu(x)\big\rangle&\ge\frac{\alpha}{|u|},\\ \big\langle n_{d+1},\tilde{\mu}^n(x)\big\rangle&\ge\frac{\alpha}{|u|}\qquad\text{for each }x\in F_{d+1}^M,n=1,2,\dots.\end{aligned} \tag{7.12}$$

In particular, this implies that

$$\big|\mu(x)\big|\ge\frac{\alpha}{|u|},\qquad\big|\tilde{\mu}^n(x)\big|\ge\frac{\alpha}{|u|}\qquad\text{for each }x\in G_M^b,n=1,2,\dots. \tag{7.13}$$

Since the right-hand sides of equations (3.7) and (4.7) are well defined as infinitely differentiable functions on all of $\mathbb{R}^d$, we can use these to extend the definitions of the λ_k, $k=1,\dots,r$ and μ to infinitely differentiable functions defined on all of $\mathbb{R}^d$.

It then follows from the continuity of μ and (7.11)–(7.12) that there is $\varphi > 0$ such that

$$\langle n_i, \mu(x)\rangle \geq \frac{\alpha}{2|u|} \qquad \text{for each } x \in F_i^{M,\varphi}, i = 1, \dots, d+1,$$

where $F_i^{M,\varphi} \triangleq \{x \in \mathbb{R}^d : \mathrm{dist}(x, F_i^M) \leq \varphi\}$, $i = 1, \dots, d+1$, and so

$$|\mu(x)| \geq \frac{\alpha}{2|u|} \qquad \text{for each } x \in G_M^{b,\varphi},$$

where $G_M^{b,\varphi} \triangleq \{x \in \mathbb{R}^d : \mathrm{dist}(x, G_M^b) \leq \varphi\}$.

For $i = 1, \dots, d+1$, let $\psi_i^M \in C_c^\infty(\mathbb{R}^d)$ such that $0 \leq \psi_i^M(x) \leq 1$ for all $x \in \mathbb{R}^d$, $\psi_i^M(x) = 1$ for all $x \in F_i^M$ and $\psi_i^M(x) = 0$ outside of $F_i^{M,\varphi}$. Define, for each $x \in \mathbb{R}^d$, $i = 1, \dots, d+1$,

$$\hat{\mu}^{i,M}(x) = \psi_i^M(x)\mu(x) + \bigl(1 - \psi_i^M(x)\bigr)n_i. \tag{7.14}$$

Then, for each $i = 1, \dots, d+1$, $\hat{\mu}^{i,M}$ is infinitely differentiable on $\mathbb{R}^d$ and we have for all $x \in \mathbb{R}^d$,

$$\langle n_i, \hat{\mu}^{i,M}(x)\rangle = \psi_i^M(x)\langle n_i, \mu(x)\rangle + \bigl(1 - \psi_i^M(x)\bigr) \tag{7.15}$$

$$\geq \psi_i^M(x)\frac{\alpha}{2|u|} + \bigl(1 - \psi_i^M(x)\bigr) \tag{7.16}$$

$$\geq \hat{\alpha}, \tag{7.17}$$

where $\hat{\alpha} \triangleq \min(\frac{\alpha}{2|u|}, 1) > 0$. Thus, for all $x \in \mathbb{R}^d$ and $i = 1, \dots d+1$, $|\hat{\mu}^{i,M}(x)| \geq \hat{\alpha}$. It follows that for each $i = 1, \dots, d+1$,

$$\hat{\gamma}^{i,M}(x) \triangleq \hat{\mu}^{i,M}(x)/|\hat{\mu}^{i,M}(x)| \qquad \text{for } x \in \mathbb{R}^d, \tag{7.18}$$

is well defined as an infinitely differentiable, unit length vector field on $\mathbb{R}^d$, and since it is equal to a constant vector outside of some compact set, it is uniformly Lipschitz continuous on $\mathbb{R}^d$. For each n, $i = 1, \dots, d+1$, we define

$$\hat{\gamma}^{i,M,n}(x) = \begin{cases} \tilde{\gamma}^n(x) & \text{if } x \in F_i^M, \\ \hat{\gamma}^{i,M}(x) & \text{if } x \notin F_i^M, \end{cases} \tag{7.19}$$

where $\tilde{\gamma}^n$ is given by (7.1), (4.23) and (4.15). Then, for each n, i, $\hat{\gamma}^{i,M,n} : \mathbb{R}^d \to \mathcal{S}_1$ is Borel measurable, and

$$|\hat{\gamma}^{i,M}(x) - \hat{\gamma}^{i,M,n}(x)| \leq |\gamma(x) - \tilde{\gamma}^n(x)|\mathbf{1}_{\{x \in F_i^M\}}. \tag{7.20}$$

Combining this with Lemma 7.1, it follows that, for every M, $\hat{\gamma}^{i,M,n}$ converges uniformly to $\hat{\gamma}^{i,M}$ on F_i^M [and hence on $\mathbb{R}^d$ by (7.19)] as $n \to \infty$ for each $i = 1, \dots, d+1$.

From the above, we see that G_M is a bounded, convex polyhedron with nonempty interior, satisfying

$$G_M^\circ = \bigcap_{i=1}^{d} \{x \in \mathbb{R}^d : \langle n_i, x\rangle > 0\} \cap \{x \in \mathbb{R}^d : \langle u, x\rangle < M\}. \tag{7.21}$$

It follows that G_M satisfies Assumptions (A1)–(A3) of Kang and Williams [17]. The vector fields $\{\hat{\gamma}^{i,M}, i = 1, \dots, d+1\}$ defined by (7.18), being uniformly Lipschitz continuous, unit length vector fields on $\mathbb{R}^d$, satisfy Assumption (A4) of [17]. Furthermore, they satisfy Assumption (A5) of [17]. To see this, we note that if $x \in F_i^M$ for some $i \in \{1, \dots, d+1\}$, then by (7.14), (7.15) and (7.18), we have

$$\langle n_i, \hat{\gamma}^{i,M}(x)\rangle = \frac{\langle n_i, \hat{\mu}^{i,M}(x)\rangle}{|\hat{\mu}^{i,M}(x)|} = \frac{\langle n_i, \mu(x)\rangle}{|\mu(x)|} \ge \frac{\hat{\alpha}}{|\mu(x)|} \ge \frac{\hat{\alpha}}{c^M}, \tag{7.22}$$

where $c^M = \sup_{y \in G_M^b} |\mu(y)| < \infty$. Thus, for $x \in G_M^b$, letting $I^M(x) = \{1 \le i \le d+1 : x \in F_i^M\}$ and $b_i(x) = 1/|I^M(x)|$ for $i \in I^M(x)$ and $b_i(x) = 0$ for $i \notin I^M(x)$, we have for $j \in I^M(x)$,

$$\left\langle \sum_{i \in I^M(x)} b_i(x) n_i, \hat{\gamma}^{j,M}(x)\right\rangle = \frac{1}{|I^M(x)|} \sum_{i \in I^M(x)} \left\langle n_i, \frac{\mu(x)}{|\mu(x)|}\right\rangle \ge \frac{\hat{\alpha}}{c^M}, \tag{7.23}$$

and

$$\left\langle \sum_{i \in I^M(x)} b_i(x) \hat{\gamma}^{i,M}(x), n_j\right\rangle = \left\langle \frac{\mu(x)}{|\mu(x)|}, n_j\right\rangle \ge \frac{\hat{\alpha}}{c^M}. \tag{7.24}$$

This establishes Assumption (A5) of [17]. So far, we have established the assumptions on the state space and the directions of reflection in Kang and Williams [17]. We now turn to verifying the assumptions on the stochastic processes required by Assumption 4.1 in [17].

Properties of $\{\tilde{\xi}^n\}$. For each n, the process $\tilde{\xi}^n = (\tilde{Z}^n, \tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n, \tilde{L}^n)$, where

$$\begin{aligned}
\tilde{Z}^n(\cdot) &= Z^n(\cdot \wedge \zeta_{M_\epsilon}^n)\mathbf{1}_{\{Z^n(0) \in G_{M_\epsilon}\}}, \\
\tilde{\mathcal{X}}^n(\cdot) &= \mathcal{X}^n(\cdot \wedge \zeta_{M_\epsilon}^n)\mathbf{1}_{\{Z^n(0) \in G_{M_\epsilon}\}}, \\
\tilde{\mathcal{Y}}^n(\cdot) &= \mathcal{Y}^n(\cdot \wedge \zeta_{M_\epsilon}^n)\mathbf{1}_{\{Z^n(0) \in G_{M_\epsilon}\}}, \\
\tilde{L}^n(\cdot) &= L^n(\cdot \wedge \zeta_{M_\epsilon}^n)\mathbf{1}_{\{Z^n(0) \in G_{M_\epsilon}\}}.
\end{aligned} \tag{7.25}$$

Then the process $\tilde{Z}^n$ lives in $G_{M_\epsilon} \subset G_M$ and, because $M > M_\epsilon + \rho$, $\tilde{Z}^n$ does not reach F_{d+1}^M. Therefore, $\tilde{Z}^n(s) \in G^b$ if and only if $\tilde{Z}^n(s) \in G_M^b$. Furthermore, $\tilde{Z}^n$ agrees with Z^n on the random time interval $(0, \zeta_{M_\epsilon}^n]$. Consequently, by (7.2)–(7.6) and (7.19) we have the following for all $t \geq 0$:

$$\begin{aligned}\tilde{Z}^n(t) &= \tilde{\mathcal{X}}^n(t) + \tilde{\mathcal{Y}}^n(t) + \frac{1}{\sqrt{V}} \int_0^t \tilde{\gamma}^n(\tilde{Z}^n(s))\, d\tilde{L}^n(s) \\ &= \tilde{\mathcal{X}}^n(t) + \tilde{\mathcal{Y}}^n(t) + \sum_{i=1}^{d+1} \int_0^t \hat{\gamma}^{i,M,n}(\tilde{Z}^n(s))\, d\tilde{\mathcal{L}}_i^n(s),\end{aligned}$$

where for $i = 1, \dots, d+1$,

$$\tilde{\mathcal{L}}_i^n(t) = \frac{1}{\delta_n \sqrt{V}} \int_0^{t \wedge \zeta_{M_\epsilon}^n} \mathbf{1}_{\{\tilde{Z}^n(s) \in F_i^M, \tilde{Z}^n(s) \notin F_j^M \text{ for any } j<i\}} |\tilde{\mu}^n(\tilde{Z}^n(s))|\, ds. \tag{7.26}$$

Here, we have split $\frac{1}{\sqrt{V}} \tilde{L}^n(\cdot)$ into a disjoint sum $\sum_{i=1}^{d+1} \tilde{\mathcal{L}}_i^n(\cdot)$ of contributions for each $i = 1, \dots, d+1$ from when $\tilde{Z}^n$ is in F_i^M and not in any of the F_j^M for $j < i$. Note that if $Z^n(0) \notin G_{M_\epsilon}$, then $\zeta_{M_\epsilon}^n \equiv 0$ and $\tilde{\mathcal{L}}_i^n \equiv 0$ for each $i = 1, \dots, d+1$. We let $\tilde{\mathcal{L}}^n = (\tilde{\mathcal{L}}_1^n, \dots, \tilde{\mathcal{L}}_{d+1}^n)$.

From the above, it follows that conditions (i)–(v) of Assumption 4.1 of [17] are satisfied with $W^n = \tilde{W}^n, X^n, \alpha^n, Y^n = \tilde{Y}^n, \beta^n, \delta^n$ in [17] replaced by our $\tilde{Z}^n, \tilde{\mathcal{X}}^n + \tilde{\mathcal{Y}}^n, \mathbf{0}, \tilde{\mathcal{L}}^n, \mathbf{0}, \delta^n$, respectively, and with $\gamma^i(x), \gamma^{i,n}(y, x)$ in [17] replaced by our $\hat{\gamma}^{i,M}(x), \hat{\gamma}^{i,M,n}(x)$, respectively, for all $y, x \in \mathbb{R}^d$. In particular, note that $\tilde{\mathcal{L}}^n$ is continuous, so it has no jumps, and condition (iv) holds because $\hat{\gamma}^{i,M,n}$ converges uniformly to $\hat{\gamma}^{i,M}$ as $n \to \infty$, for $i = 1, \dots, d+1$. The remaining condition of Assumption 4.1 to be verified is that $\{\tilde{\mathcal{X}}^n + \tilde{\mathcal{Y}}^n\}$ is $\mathcal{C}$-tight. For this, we shall prove that $\{(\tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n)\}$ is $\mathcal{C}$-tight. Once this is verified, it follows using Theorem 4.2 of [17], that $\{(\tilde{Z}^n, \tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n, \tilde{\mathcal{L}}^n)\}$ is $\mathcal{C}$-tight, and hence, since $\tilde{L}^n = \sqrt{V} \sum_{i=1}^{d+1} \tilde{\mathcal{L}}_i^n$, that

$$\{\tilde{\varsigma}^n = (\tilde{Z}^n, \tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n, \tilde{L}^n)\} \qquad \text{is } \mathcal{C}\text{-tight.} \tag{7.27}$$

This, together with (7.9), (7.10) and Proposition 7.1, gives that $\{\varsigma^n\}$ is $\mathcal{C}$-tight. Thus, to complete the proof of Theorem 7.1, we only need to show that $\{(\tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n)\}$ is $\mathcal{C}$-tight. We now prove this.

$\{(\tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n)\}$ *is* $\mathcal{C}$*-tight.* To verify the $\mathcal{C}$-tightness of $\{(\tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n)\}$, we use a well-known variant of the $\mathcal{C}$-tightness conditions given in Proposition 7.1, that employs Aldous' criterion. First, since $\tilde{\mathcal{X}}^n$ is continuous and the jump sizes of $\tilde{\mathcal{Y}}^n$ satisfy the following inequality for each $\tilde{T} > 0$:

$$\sup_{0 \leq t \leq \tilde{T}} |\tilde{\mathcal{Y}}^n(t) - \tilde{\mathcal{Y}}^n(t-)| \leq \frac{\delta_n}{\sqrt{V}} \max_{1 \leq k \leq r} |v_k| \to 0 \qquad \text{as } n \to \infty, \tag{7.28}$$

where $\tilde{\mathcal{Y}}(0-) = \tilde{\mathcal{Y}}(0)$, it suffices to prove that $\{(\tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n)\}$ is tight (see, e.g., Proposition 3.26, page 351, of [15]). To prove the tightness, by Theorem 4.5 on page 356 of [15], it suffices to prove compact containment and the following, which is known as Aldous' condition: for each $\tilde{T} > 0$, $\tilde{\eta} > 0$,

$$\lim_{\theta\to 0}\limsup_{n\to\infty}\sup_{\substack{\tau_1,\tau_2\in\mathcal{T}^n_{\tilde{T}}\\ \tau_1\le\tau_2\le\tau_1+\theta}}\mathbb{P}\big(|\tilde{\mathcal{X}}^n(\tau_2)-\tilde{\mathcal{X}}^n(\tau_1)|\vee|\tilde{\mathcal{Y}}^n(\tau_2)-\tilde{\mathcal{Y}}^n(\tau_1)|\ge\tilde{\eta}\big) \tag{7.29}$$
$$=0,$$

where $\mathcal{T}^n_{\tilde{T}}$ denotes the set of stopping times relative to $\{\mathcal{F}^n_{\tilde{T}}\}$ that are bounded by $\tilde{T}$. In fact, if Aldous' criterion holds, the compact containment condition follows immediately using this and the facts that $\tilde{\mathcal{X}}^n(0) = \mathbf{1}_{\{Z^n(0)\in G_{M_\epsilon}\}}Z^n(0)$ is bounded in norm by M_ϵ and $\tilde{\mathcal{Y}}^n(0) = \mathbf{0}$. Thus, it suffices to prove (7.29) for each $\tilde{T} > 0$, $\tilde{\eta} > 0$.

For this, fix $\tilde{T} > 0$, $\tilde{\eta} > 0$. Then, for $\theta > 0$, $\tau_1, \tau_2 \in \mathcal{T}^n_{\tilde{T}}$ such that $\tau_1 \le \tau_2 \le \tau_1 + \theta$, we have by (7.3) and (7.25) that

$$\begin{aligned}|\tilde{\mathcal{X}}^n(\tau_2)-\tilde{\mathcal{X}}^n(\tau_1)|^2 &\le 2\sum_{i=1}^d\left(\int_{\tau_1}^{\tau_2}\mu_i(\tilde{Z}^n(s))\mathbf{1}_{\{\tilde{Z}^n(s)\in G^\circ,s\le\zeta^n_{M_\epsilon}\}}\,ds\right)^2\\ &\quad+\frac{2}{V}\sum_{i=1}^d\left(\sum_{j=1}^d\int_{\tau_1}^{\tau_2}\sigma_{ij}(\tilde{Z}^n(s))\mathbf{1}_{\{\tilde{Z}^n(s)\in G^\circ,s\le\zeta^n_{M_\epsilon}\}}\,dW_j(s)\right)^2\end{aligned}$$

and

$$|\tilde{\mathcal{Y}}^n(\tau_2)-\tilde{\mathcal{Y}}^n(\tau_1)|^2\le\sum_{i=1}^d\big(\mathcal{Y}^n_i(\tau_2\wedge\zeta^n_{M_\epsilon})-\mathcal{Y}^n_i(\tau_1\wedge\zeta^n_{M_\epsilon})\big)^2.$$

Then, on taking expectations in the above, using the fact that μ and σ are bounded on G_{M_ϵ}, together with the L^2-isometry for the stopped stochastic integrals with respect to W and for the stopped compensated terms involving Poisson processes, we have

$$\begin{aligned}&\mathbb{E}\big[|\tilde{\mathcal{X}}^n(\tau_2)-\tilde{\mathcal{X}}^n(\tau_1)|^2\big]\\ &\quad\le 2\theta^2K_1+\frac{2d}{V}\sum_{i=1}^d\sum_{j=1}^d\mathbb{E}\left[\int_{\tau_1}^{\tau_2}\sigma^2_{ij}(\tilde{Z}^n(s))\mathbf{1}_{\{\tilde{Z}^n(s)\in G^\circ,s\le\zeta^n_{M_\epsilon}\}}\,ds\right]\\ &\quad\le 2\left(\theta^2K_1+\frac{\theta\,dK_2}{V}\right)\end{aligned}\tag{7.30}$$

and

$$\begin{aligned}&\mathbb{E}\big[|\tilde{\mathcal{Y}}^n(\tau_2)-\tilde{\mathcal{Y}}^n(\tau_1)|^2\big]\\ &\quad\le\frac{\delta_n^2 r}{V}\sum_{i=1}^d\sum_{k=1}^r v^2_{ik}\mathbb{E}\left[\delta_n^{-2}\int_{\tau_1}^{\tau_2}\tilde{\lambda}^n_k(\tilde{Z}^n(s))\mathbf{1}_{\{\tilde{Z}^n(s)\in G^b,s\le\zeta^n_{M_\epsilon}\}}\,ds\right]\\ &\quad\le\frac{r\theta K_3}{V},\end{aligned}\tag{7.31}$$

where for the last double sum we have used the fact that $0 \le \tilde{\lambda}_k^n \le \lambda_k$ on G, for all n and $k = 1, \dots, r$, and $K_1 \triangleq \sup_{x \in G_{M_\epsilon+\rho}} \sum_{i=1}^d |\mu_i(x)|^2$, $K_2 \triangleq \sup_{x \in G_{M_\epsilon+\rho}} \sum_{i=1}^d \sum_{j=1}^d \sigma_{ij}^2(x)$, and $K_3 \triangleq \sup_{x \in G_{M_\epsilon+\rho}} \sum_{i=1}^d \sum_{k=1}^r v_{ik}^2 |\lambda_k(x)|$ are all finite constants. Hence, using Markov's inequality and taking the supremum over the set of τ_1, τ_2 under consideration, we have for each n and θ that

$$\begin{aligned} &\sup_{\substack{\tau_1,\tau_2 \in \mathcal{T}_T^n \\ \tau_1 \le \tau_2 \le \tau_1+\theta}} \mathbb{P}(|\tilde{\mathcal{X}}^n(\tau_2) - \tilde{\mathcal{X}}^n(\tau_1)| \vee |\tilde{\mathcal{Y}}^n(\tau_2) - \tilde{\mathcal{Y}}^n(\tau_1)| \ge \tilde{\eta}) \\ &\qquad \le \tilde{\eta}^{-2}\left(2\theta^2 K_1 + \frac{2d\theta K_2 + r\theta K_3}{V}\right). \end{aligned}$$

Since the bound above does not depend on n and tends to zero as $\theta \to 0$, it follows that (7.29) holds. This completes the proof of the tightness of $\{(\tilde{\mathcal{X}}^n, \tilde{\mathcal{Y}}^n)\}$, and hence the theorem is proved. □

The following is a corollary of Theorem 7.1 and the proof of Lemma 7.2.

COROLLARY 7.1. *For each $T \ge 0$,*

$$\int_0^T \mathbf{1}_{\{Z^n(s) \in G^b\}}\, ds \to 0 \qquad \textit{in probability as } n \to \infty. \tag{7.32}$$

PROOF. Given $T \ge 0$ and $\epsilon > 0$, by the proof of Lemma 7.2, there is $M_\epsilon > \tilde{M}^*$ and $n_\epsilon > 0$, such that

$$\mathbb{P}(\zeta_{M_\epsilon}^n \le T) < \frac{\epsilon}{2} \qquad \text{for all } n \ge n_\epsilon. \tag{7.33}$$

Then, for $\eta > 0$, using the fact from Lemma 4.2 that $|\tilde{\mu}^n(x)| \ge \alpha > 0$ for all $x \in G^b$ and all n, we have for all $n \ge n_\epsilon$,

$$\begin{aligned} &\mathbb{P}\left(\left|\int_0^T \mathbf{1}_{\{Z^n(s)\in G^b\}}\, ds\right| \ge \eta\right) \\ &\quad \le \mathbb{P}(\zeta_{M_\epsilon}^n \le T) + \mathbb{P}\left(\zeta_{M_\epsilon}^n > T, \left|\int_0^{T\wedge\zeta_{M_\epsilon}^n} \mathbf{1}_{\{Z^n(s)\in G^b\}}\, ds\right| \ge \eta\right) \\ &\quad \le \frac{\epsilon}{2} + \mathbb{P}\left(\zeta_{M_\epsilon}^n > T, \frac{1}{\delta_n}\int_0^{T\wedge\zeta_{M_\epsilon}^n} |\tilde{\mu}^n(Z^n(s))| \mathbf{1}_{\{Z^n(s)\in G^b\}}\, ds \ge \frac{\alpha\eta}{\delta_n}\right) \\ &\quad \le \frac{\epsilon}{2} + \mathbb{P}\left(\zeta_{M_\epsilon}^n > T, L^n(T\wedge\zeta_{M_\epsilon}^n) \ge \frac{\alpha\eta}{\delta_n}\right). \end{aligned}$$

Now, from Theorem 7.1, we know that $\{L^n\}$ is $\mathcal{C}$-tight, and hence satisfies the compact containment condition of Proposition 7.1. Therefore, since $\delta_n \to 0$ as $n \to \infty$, there is $n_{\eta,\epsilon} \ge n_\epsilon$ such that

$$\mathbb{P}\left(L^n(T) \ge \frac{\alpha\eta}{\delta_n}\right) < \frac{\epsilon}{2} \qquad \text{for all } n \ge n_{\eta,\epsilon}.$$

Combining all of the above, we have that for all $n \geq n_{\eta,\epsilon}$,

$$\mathbb{P}\left(\left|\int_0^T \mathbf{1}_{\{Z^n(s)\in G^b\}}\, ds\right| \geq \eta\right) \leq \epsilon.$$

Since $\eta > 0$, $\epsilon > 0$ were arbitrary, the desired result follows. □

7.2. *Characterization of the weak-sense limit.* In this subsection, we show that $\{Z^n\}$ converges weakly to a solution of the constrained Langevin equation with initial condition having the same distribution as $\bar{X}^V(0)$. As we have noted previously, such a solution is unique in law. For convenience, in this subsection, we let ϑ^V denote the distribution of $\bar{X}^V(0)$, which is fixed, since we are fixing V.

Our main technical result is the following.

THEOREM 7.2. *The sequence of processes* $\{(Z^n, W, \mathcal{X}^n, \mathcal{Y}^n, L^n)\}$ *is* $\mathcal{C}$*-tight and any weak limit point*, $(Z^*, W^*, \mathcal{X}^*, \mathcal{Y}^*, L^*)$, *is such that* (Z^*, W^*, L^*) *satisfy the conditions of the Definition 5.1 of a solution of the constrained Langevin equation*, *and*

$$Z^*(t) = \mathcal{X}^*(t) + \frac{1}{\sqrt{V}}\int_0^t \gamma\big(Z^*(s)\big)\, dL^*(s), \qquad t \geq 0, \tag{7.34}$$

$$\begin{aligned}\mathcal{X}^*(t) = \mathcal{X}^*(0) &+ \int_0^t \mu\big(Z^*(s)\big)\, ds \\ &+ \frac{1}{\sqrt{V}}\int_0^t \sigma\big(Z^*(s)\big)\, dW^*(s), \qquad t \geq 0,\end{aligned} \tag{7.35}$$

$$\mathcal{Y}^* \equiv \mathbf{0}, \tag{7.36}$$

where $\mathcal{X}^*(0)$ *has distribution* ϑ^V, *and* W^* *is a martingale with respect to the filtration generated by* (Z^*, W^*, L^*).

PROOF. Since W is a continuous process and it does not depend on n, it follows immediately from Theorem 7.1 that $\{(Z^n, W, \mathcal{X}^n, \mathcal{Y}^n, L^n)\}$ is $\mathcal{C}$-tight. Suppose that $(Z^*, W^*, \mathcal{X}^*, \mathcal{Y}^*, L^*)$ is a weak limit point of the aforementioned sequence. For notational convenience, we shall denote the weakly converging subsequence again by $\{(Z^n, W, \mathcal{X}^n, \mathcal{Y}^n, L^n)\}$. Because of the $\mathcal{C}$-tightness, the weak limit $(Z^*, W^*, \mathcal{X}^*, \mathcal{Y}^*, L^*)$ may be assumed to have continuous paths (surely). Since W does not depend on n, W^* will be a standard d-dimensional Brownian motion. Furthermore, for every n, W is a martingale with respect to the filtration generated by $(Z^n, W, \hat{U}^{n,b}, \hat{Y}^n)$ and so W is a martingale with respect to the smaller filtration generated by (Z^n, W), to which it is still adapted. Since L^n is adapted to the filtration generated by Z^n, it follows that, for every n, W is a martingale with respect to the filtration generated by (Z^n, W, L^n). Thus, for any $0 \leq s \leq t$, any positive integer p, any continuous bounded function $g : \mathbb{R}^{p(2d+1)} \to \mathbb{R}$,

$0 \le s_1 < s_2 < \cdots s_p \le s$, we have by the continuous mapping theorem and dominated convergence that

$$\begin{aligned}
&\mathbb{E}\big[\big(W^*(t) - W^*(s)\big)g\big(Z^*(s_1), W^*(s_1), L^*(s_1), \dots, Z^*(s_p), W^*(s_p), L^*(s_p)\big)\big] \\
&\quad = \lim_{n\to\infty} \mathbb{E}\big[\big(W(t) - W(s)\big) \\
&\qquad \times g\big(Z^n(s_1), W(s_1), L^n(s_1), \dots, Z^n(s_p), W(s_p), L^n(s_p)\big)\big] \\
&\quad = 0,
\end{aligned}$$

where we used the martingale property of W for the last equality. It follows that W^* is a martingale with respect to the filtration generated by (Z^*, W^*, L^*).

For the purposes of verifying the other properties associated with (Z^*, W^*, L^*) being a solution of the CLE and that (7.34)–(7.36) hold, by using the Skorokhod representation theorem, we may suppose that the convergence of $\{(Z^n, W, \mathcal{X}^n, \mathcal{Y}^n, L^n)\}$ to $(Z^*, W^*, \mathcal{X}^*, \mathcal{Y}^*, L^*)$ is almost sure uniform convergence on compact time intervals, rather than just weak convergence.

Let $\mathcal{F}_t = \sigma\{(Z^*(s), W^*(s), L^*(s)) : 0 \le s \le t\}$ for all $t \ge 0$. Then Z^*, W^*, L^* are adapted to this filtration. Since $Z^n(t) \in G$ for all $t \ge 0$, it follows that the limit process Z^* satisfies $Z^*(t) \in G$ for all $t \ge 0$, almost surely. Furthermore, since $Z^n(0)$ has the same distribution ϑ^V for all n, it follows that $Z^*(0)$ will have this same distribution.

The process L^* will be nondecreasing (a.s.) because it is the a.s. limit under uniform convergence on compact time intervals of the nondecreasing processes L^n. To establish the integral property in (iii) of the definition of a solution of the CLE, for each $\epsilon > 0$, let $f_\epsilon : \mathbb{R}^d \to \mathbb{R}$ be a continuous function such that $0 \le f_\epsilon \le 1$, $f_\epsilon = 0$ on G^b, $f_\epsilon(x) = 1$ when $\mathrm{dist}(x, G^b) \ge \epsilon$. Then by Lemma A.4 of Kang and Williams [17], since a.s., $(Z^n, L^n) \to (Z^*, L^*)$ uniformly on compact time intervals, L^n is nondecreasing for each n, and f_ϵ is continuous, we have a.s. for each $\epsilon > 0$ and $t > 0$,

$$\int_0^t f_\epsilon\big(Z^*(s)\big)\, dL^*(s) = \lim_{n\to\infty} \int_0^t f_\epsilon\big(Z^n(s)\big)\, dL^n(s) = 0,$$

where the last equality follows because L^n can only increase when Z^n is on G^b and f_ϵ is zero there. On letting $\epsilon \to 0$, we obtain by the bounded convergence theorem that a.s.,

$$\int_0^t \mathbf{1}_{\{Z^*(s) \notin G^b\}}\, dL^*(s) = 0 \qquad \text{for all } t \ge 0,$$

from which it follows that a.s., L^* can increase only when Z^* is on G^b.

It remains to verify that (7.34)–(7.36) hold, since property (iv) of the definition of a solution of the CLE follows from (7.34)–(7.35). We shall verify this for stopped versions of the processes and then let the stopping times tend to infinity to obtain the desired result. We first define the stopping times that we will use.

For each $M > 0$, recall that

$$\zeta_M^n = \inf\{t \ge 0 : \langle u, Z^n(t)\rangle > M\} \tag{7.37}$$

and define

$$\zeta_M^* = \inf\{t \ge 0 : \langle u, Z^*(t)\rangle > M\}. \tag{7.38}$$

Then there is an increasing sequence of positive constants, $\{M_\ell, \ell = 1, 2, \dots\}$, such that $M_\ell > \tilde{M}^*$ for all ℓ, $M_\ell \to \infty$ as $\ell \to \infty$, and for each ℓ,

$$\mathbb{P}\Big(\lim_{n\to\infty} \zeta_{M_\ell}^n = \zeta_{M_\ell}^*\Big) = 1; \tag{7.39}$$

see [23], pages 13–14. Thus, for each ℓ, the sequence of stopping times $\{\zeta_{M_\ell}^n\}_{n=1}^\infty$ converges a.s. to $\zeta_{M_\ell}^*$.

We first consider the convergence of $\{\mathcal{X}^n\}$. Fix $T > 0$, $\ell \in \{1, 2, \dots\}$. Then for each $t \in [0, T]$,

$$\begin{aligned}\mathcal{X}^n\big(t \wedge \zeta_{M_\ell}^n\big) = \mathcal{X}^n(0) &+ \int_0^{t\wedge\zeta_{M_\ell}^n} \mu\big(Z^n(s)\big)\,ds \\ &+ \frac{1}{\sqrt{V}}\int_0^{t\wedge\zeta_{M_\ell}^n} \sigma\big(Z^n(s)\big)\,dW(s) \\ &- \int_0^{t\wedge\zeta_{M_\ell}^n} \mu\big(Z^n(s)\big)\mathbf{1}_{\{Z^n(s)\in G^b\}}\,ds \\ &- \frac{1}{\sqrt{V}}\int_0^{t\wedge\zeta_{M_\ell}^n} \sigma\big(Z^n(s)\big)\mathbf{1}_{\{Z^n(s)\in G^b\}}\,dW(s).\end{aligned} \tag{7.40}$$

Now, on $\{\zeta_{M_\ell}^* > 0\}$, we have that as $n \to \infty$, a.s., uniformly on $[0, T]$,

$$\mu\big(Z^n(\cdot \wedge \zeta_{M_\ell}^n)\big) \to \mu\big(Z^*(\cdot \wedge \zeta_{M_\ell}^*)\big),$$
$$\sigma\big(Z^n(\cdot \wedge \zeta_{M_\ell}^n)\big) \to \sigma\big(Z^*(\cdot \wedge \zeta_{M_\ell}^*)\big).$$

It follows from this and the fact that $\zeta_{M_\ell}^n \to \zeta_{M_\ell}^*$ a.s. as $n \to \infty$, that on $\{\zeta_{M_\ell}^* > 0\}$, as $n \to \infty$, almost surely, uniformly on $[0, T]$,

$$\int_0^{\cdot\wedge\zeta_{M_\ell}^n} \mu\big(Z^n(s)\big)\,ds \to \int_0^{\cdot\wedge\zeta_{M_\ell}^*} \mu\big(Z^*(s)\big)\,ds. \tag{7.41}$$

Furthermore, by Theorem VI.6.22 and Corollary VI.6.29 of [15], on $\{\zeta_{M_\ell}^* > 0\}$, as $n \to \infty$, in the uniform norm on $[0, T]$,

$$\int_0^{\cdot\wedge\zeta_{M_\ell}^n} \sigma\big(Z^n(s)\big)\,dW(s) \to \int_0^{\cdot\wedge\zeta_{M_\ell}^*} \sigma\big(Z^*(s)\big)\,dW^*(s) \qquad \text{in probability.} \tag{7.42}$$

(We note that in [15], stochastic integrands are left-continuous modifications of the ones used here. However, when integrating against the continuous martingale W, it does not matter whether one uses the left-continuous modification or the original right-continuous process, the value of the stochastic integral is the same.) Finally, using Corollary 7.1 and the L^2-isometry for stochastic integrals, it follows that on $\{\zeta^*_{M_\ell} > 0\}$, as $n \to \infty$, in the uniform norm on $[0, T]$,

$$\int_0^{\cdot\wedge\zeta^n_{M_\ell}} \mu(Z^n(s))\mathbf{1}_{\{Z^n(s)\in G^b\}}\,ds \to 0 \qquad \text{in probability,} \tag{7.43}$$

$$\int_0^{\cdot\wedge\zeta^n_{M_\ell}} \sigma(Z^n(s))\mathbf{1}_{\{Z^n(s)\in G^b\}}\,dW(s) \to 0 \qquad \text{in probability.} \tag{7.44}$$

Substituting (7.41)–(7.44) in (7.40) and letting $n \to \infty$, using the convergence of $\mathcal{X}^n$ to $\mathcal{X}^*$, we see that on $\{\zeta^*_{M_\ell} > 0\}$, we have that almost surely,

$$\begin{aligned}\mathcal{X}^*(t\wedge\zeta^*_{M_\ell}) = \mathcal{X}^*(0) &+ \int_0^{t\wedge\zeta^*_{M_\ell}} \mu(Z^*(s))\,ds \\ &+ \frac{1}{\sqrt{V}}\int_0^{t\wedge\zeta^*_{M_\ell}} \sigma(Z^*(s))\,dW^*(s).\end{aligned} \tag{7.45}$$

We note that on $\{\zeta^*_{M_\ell} = 0\}$, this equation holds trivially. Since Z^* is almost surely continuous, $\zeta^*_{M_\ell} \to \infty$ a.s. when $\ell \to \infty$. So on letting $\ell \to \infty$ in (7.45), we see that (7.35) holds a.s. for all $t \in [0, T]$.

To show the convergence of $\{\mathcal{Y}^n\}$ to the zero process, we first estimate the expectation of the square of $\mathcal{Y}^n(t\wedge\zeta^n_{M_\ell})$. For $T > 0$, ℓ fixed, $i \in \{1, \dots, d\}$, we have for $t \in [0, T]$,

$$\begin{aligned}\mathbb{E}[(\mathcal{Y}^n_i(t\wedge\zeta^n_{M_\ell}))^2] &\le \frac{\delta_n^2}{V}\left(\sum_{k=1}^r v_{ik}^2\right)\sum_{k=1}^r \mathbb{E}[(\hat{U}^b_k(t\wedge\zeta^n_{M_\ell}))^2] \\ &\le \frac{\delta_n^2}{V}\left(\sum_{k=1}^r v_{ik}^2\right)\sum_{k=1}^r \mathbb{E}\left[\delta_n^{-2}\int_0^{t\wedge\zeta^n_{M_\ell}} \tilde{\lambda}^n_k(Z^n(s))\mathbf{1}_{\{Z^n(s)\in G^b\}}\,ds\right] \\ &\le \frac{1}{V}\left(\sum_{k=1}^r v_{ik}^2\right) rK_{M_\ell}\mathbb{E}\left[\int_0^{t\wedge\zeta^n_{M_\ell}} \mathbf{1}_{\{Z^n(s)\in G^b\}}\,ds\right] \\ &\to 0 \qquad \text{as } n\to\infty,\end{aligned}$$

where $K_{M_\ell} = \sup_{k=1}^r \sup_{x\in G_{M_\ell}} \lambda_k(x)$, we have used the fact that $\tilde{\lambda}^n_k(x) \le \lambda_k(x)$ for all $x \in \mathbb{R}^d_+$ and we have used Corollary 7.1 and dominated convergence for the last line. Thus, for each $t \in [0, T]$, $\{\mathcal{Y}^n(t\wedge\zeta^n_{M_\ell})\}$ converges in L^2 to the zero vector as $n \to \infty$. Since we already know that $\{\mathcal{Y}^n(t\wedge\zeta^n_{M_\ell})\}$ converges a.s. to $\mathcal{Y}^*(t\wedge\zeta^*_{M_\ell})$, it follows that a.s., $\mathcal{Y}^*(\cdot\wedge\zeta^*_{M_\ell}) = \mathbf{0}$ on $[0, T]$. On letting $\ell \to \infty$, we see that a.s., $\mathcal{Y}^*(\cdot) = \mathbf{0}$ on $[0, T]$.

For the convergence of the integral with respect to L^n in (7.2), fix $T > 0, \ell$. For each $t \in [0, T]$,

$$\begin{aligned}
&\limsup_{n\to\infty}\left|\int_0^{t\wedge\zeta^*_{M_\ell}} \gamma\big(Z^*(s)\big)\,dL^*(s) - \int_0^{t\wedge\zeta^n_{M_\ell}} \tilde{\gamma}^n\big(Z^n(s)\big)\,dL^n(s)\right| \\
&\quad\le \limsup_{n\to\infty}\left|\int_0^t \gamma\big(Z^*\big(s\wedge\zeta^*_{M_\ell}\big)\big)\,dL^*\big(s\wedge\zeta^*_{M_\ell}\big)\right. \\
&\qquad \left. - \int_0^t \gamma\big(Z^n\big(s\wedge\zeta^n_{M_\ell}\big)\big)\,dL^n\big(s\wedge\zeta^n_{M_\ell}\big)\right| \\
&\qquad + \limsup_{n\to\infty}\left|\int_0^t (\gamma - \tilde{\gamma}^n)\big(Z^n\big(s\wedge\zeta^n_{M_\ell}\big)\big)\,dL^n\big(s\wedge\zeta^n_{M_\ell}\big)\right|.
\end{aligned} \tag{7.46}$$

Now a.s., $L^*(\cdot\wedge\zeta^*_{M_\ell})$ [resp., $L^n(\cdot\wedge\zeta^n_{M_\ell})$] can increase only when $Z^*(\cdot\wedge\zeta^*_{M_\ell}) \in \tilde{G}^b_{M_\ell} = \{x \in G^b_{M_\ell} : x \notin G^\circ\}$ [resp., $Z^n(\cdot\wedge\zeta^n_{M_\ell}) \in \tilde{G}^b_{M_\ell}$], and γ_{M_ℓ} of Lemma 6.1 is a continuous, bounded extension of γ from $\tilde{G}^b_{M_\ell}$ to $\mathbb{R}^d$. Then, since, a.s., $(Z^n(\cdot\wedge\zeta^n_{M_\ell}), L^n(\cdot\wedge\zeta^n_{M_\ell}))$ converges to $(Z^*(\cdot\wedge\zeta^*_{M_\ell}), L^*(\cdot\wedge\zeta^*_{M_\ell}))$ uniformly on $[0, T]$ and L^n is nondecreasing for each n, it follows from Lemma A.4 of [17] that a.s., as $n\to\infty$, uniformly for $t \in [0, T]$,

$$\begin{aligned}
&\int_0^t \gamma\big(Z^n\big(s\wedge\zeta^n_{M_\ell}\big)\big)\,dL^n\big(s\wedge\zeta^n_{M_\ell}\big) \\
&\qquad = \int_0^t \gamma_M\big(Z^n\big(s\wedge\zeta^n_{M_\ell}\big)\big)\,dL^n\big(s\wedge\zeta^n_{M_\ell}\big) \\
&\qquad \to \int_0^t \gamma_M\big(Z^*\big(s\wedge\zeta^*_{M_\ell}\big)\big)\,dL^*\big(s\wedge\zeta^*_{M_\ell}\big) \\
&\qquad = \int_0^t \gamma\big(Z^*\big(s\wedge\zeta^*_{M_\ell}\big)\big)\,dL^*\big(s\wedge\zeta^*_{M_\ell}\big).
\end{aligned}$$

It follows that a.s., for all $t \in [0, T]$, the first lim sup in the right-hand side of the inequality in (7.46) is zero. For the second lim sup in the right-hand side of that inequality, we have by Lemma 7.1 that given $\epsilon > 0$, there is $n_\epsilon > 1$ (depending on M_ℓ) such that for all $n \ge n_\epsilon$ and $x \in \tilde{G}^b_{M_\ell}$,

$$\big|\gamma(x) - \tilde{\gamma}^n(x)\big| < \epsilon.$$

(Here, we used the fact that $|x| \le \langle u, x\rangle \le M_\ell$ for $x \in \tilde{G}^b_{M_\ell}$.) It follows that a.s., for all $t \in [0, T]$, the second lim sup in the right-hand side of the inequality in (7.46) is bounded for each $\epsilon > 0$ by

$$\epsilon \limsup_{n\to\infty} L^n\big(t\wedge\zeta^n_{M_\ell}\big) = \epsilon L^*\big(t\wedge\zeta^*_{M_\ell}\big).$$

On letting ϵ tend to zero, we see that a.s., the right-hand side of (7.46) is zero for all $t \in [0, T]$. Thus, a.s., for all $t \in [0, T]$,

$$\int_0^{t\wedge\zeta^n_{M_\ell}} \tilde{\gamma}^n\big(Z^n(s)\big)\,dL^n(s) \to \int_0^{t\wedge\zeta^*_{M_\ell}} \gamma(Z^*(s)\,dL^*(s). \tag{7.47}$$

On letting $n \to \infty$ in (7.2) stopped at $\zeta^n_{M_\ell}$, since a.s., $(Z^n(\cdot \wedge \zeta^n_{M_\ell}), \mathcal{X}^n(\cdot \wedge \zeta^n_{M_\ell}, \mathcal{Y}^n(\cdot \wedge \zeta^n_{M_\ell}))$ converges to $(Z^*(\cdot \wedge \zeta^*_{M_\ell}), \mathcal{X}^*(\cdot \wedge \zeta^*_{M_\ell}), \mathbf{0})$ uniformly on $[0, T]$ as $n \to \infty$, and we have (7.47), we see that a.s., for all $t \in [0, T]$,

$$Z^*\big(t \wedge \zeta^*_{M_\ell}\big) = \mathcal{X}^*\big(t \wedge \zeta^*_{M_\ell}\big) + 0 + \frac{1}{\sqrt{V}} \int_0^{t\wedge\zeta^*_{M_\ell}} \gamma\big(Z^*(s)\big)\,dL^*(s). \tag{7.48}$$

On letting $\ell \to \infty$, we see that a.s., (7.34) holds for all $t \in [0, T]$. Since $T > 0$ was arbitrary, the desired result follows. □

The following is immediate from the above theorem and weak uniqueness for solutions of the CLE.

COROLLARY 7.2. *The sequence* $\{(Z^n, L^n)\}$ *converges weakly to* (Z^*, L^*) *as* $n \to \infty$, *where* Z^* *is a solution of the constrained Langevin equation with initial distribution* ϑ^V, *and* L^* *is the associated auxiliary local time process.*

PROOF. Since we have $\mathcal{C}$-tightness and uniqueness in law for solutions of the CLE, this follows from a standard real analysis argument. Indeed, by the preceding theorem, any subsequence of $\{(Z^n, W, L^n)\}$ has a further subsequence that converges to some (Z^*, W^*, L^*), which defines a solution of the CLE, together with its two auxiliary processes, where the initial distribution of Z^* is ϑ^V. By the uniqueness in law of Corollary 6.1, given ϑ^V, the distribution of (Z^*, L^*) is unique, no matter what converging subsequence was used. It follows, by a standard argument by contradiction, that the sequence $\{(Z^n, L^n)\}$ converges weakly to (Z^*, L^*), where Z^* is a solution of the CLE with initial distribution ϑ^V and L^* is the associated local time process. □

8. Numerical examples. In this section, we present results of some numerical experiments, to illustrate the practical use of the constrained Langevin approximation proposed here. We consider three different examples of reaction networks: in the first example, a simple chemical reaction network whose deterministic model RRE has only one stable steady-state (near the boundary) is considered; for the second example, we consider the Brusselator chemical reaction set that has a limit cycle for the deterministic RRE; and, for the third example, we consider a (bistable) chemical reaction network for which the RRE has two stable fixed points.

For the first and second examples, we performed simulations of the system evolving over a long period of time to observe the system near steady state. These simulations were repeated for 10 different runs in order to calculate confidence intervals. For each example, we performed simulations for the scaled *Markov Chain Model (MCM)*, the *Linear Noise Approximation (LNA)* and the *Constrained Langevin Approximation (CLA)* proposed here. For the first example, in order to compare the behavior near the boundary, we also simulated two modifications of the Langevin approximation that guarantee nonnegative concentrations. First, we considered a Langevin approximation in the interior of the orthant, extended by normal reflection at the boundary. Notice that this contrasts with the CLA proposed here since, in our approximation, the directions of reflection at the boundary are often oblique. We also considered a modification of the Langevin approximation which chops off negative values. That is, we simulated a discrete approximation to the Langevin equation but, after each step of the simulation, any negative concentrations of chemical species were set to zero. We call the first extension the "*Langevin Equation with Normal Reflection (LE-NR)*" and the second the "*Langevin Equation with Chopping (LE-Chop)*."

In order to simulate LNA and LE-Chop, the Euler–Maruyama method for SDEs was used [26, 29]. For the constrained diffusions, that is, for the CLA and LE-NR, we used the method proposed by Bossy et al. [3], which is also based on the Euler–Maruyama method for the interior behavior and approximates the reflection at the boundary by jumping in the direction of γ. It is important to mention here that the method proposed by Bossy et al. [3] assumes that the SDER has reflection on a *smooth* boundary. Although this assumption does not hold here, this was not a problem for our simulations since they did not cross more than one face of $\mathbb{R}^d_+$ in one simulation step, and hence the boundary was treated as though it were smooth. Bossy et al. also assumed a bounded domain, but since our simulations were for a finite time, with finite replication, we used an approximation by a sufficiently large bounded domain. Indeed, we believe that the results in [17] could be used to prove that the processes used in our simulation scheme approximate the SDER solution in distribution. This is an interesting avenue for future research; furthermore, it would be interesting to develop convergence rate guarantees for such numerical approximations. In order to simulate the deterministic model, which is used during the LNA simulation, we used a fourth-order Runge–Kutta method to increase accuracy. For simulating the Markov chain model, the discrete-event Doob–Gillespie algorithm was implemented. The codes were written in the R programming language [30] and ran on a 2.6 GHz Intel Core i5 processor.

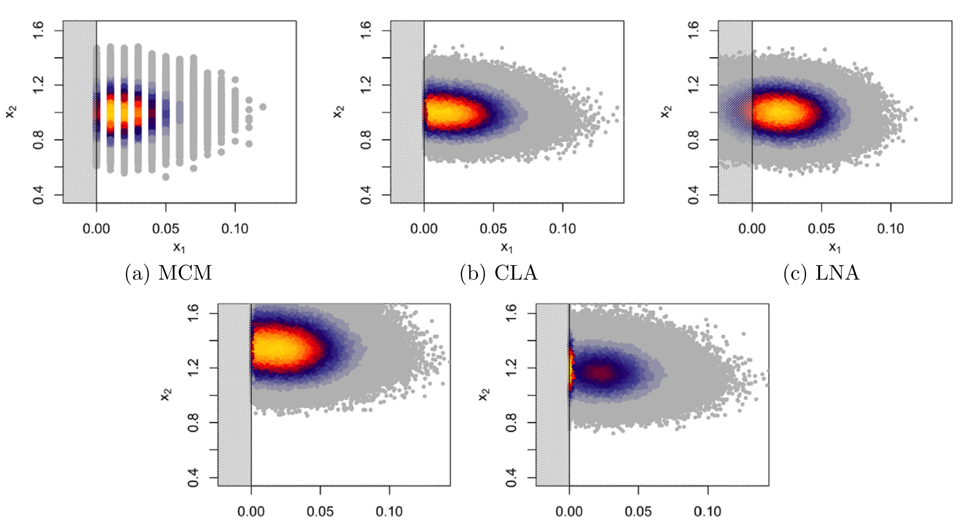
8.1. *Example* 1. In this subsection, we consider the following simple example of a chemical reaction network comprised of two species S_1 and S_2 which are involved in six reactions, as follows:

$$S_1 \underset{c_2}{\overset{c_1}{\rightleftharpoons}} \varnothing, \qquad S_2 \underset{c_4}{\overset{c_3}{\rightleftharpoons}} \varnothing, \qquad S_1 \underset{c_6}{\overset{c_5}{\rightleftharpoons}} S_2.$$

Notice that we labeled each reaction arrow with their respective normalized reaction rate constants. In order to observe a system with a low concentration of

S_1 near steady state, we used the following values: $c_1 = 10^{-4}$, $c_2 = 1$, $c_3 = 1$, $c_4 = 10^{-4}$, $c_5 = 100$, $c_6 = 1$. The constant V, which denotes the volume of the vessel times Avogadro's number, was set to 100. As a starting point for the simulations, we used the steady-state for the deterministic model, which is approximately given by $\bar{x}^o = (0.02, 1.00)'$. For the Markov chain simulation, we used the rounded starting point $\bar{x}^o$ in order to have an initial condition in the lattice $G^V \triangleq \{y/V : y \in \mathbb{Z}_+^2\}$. For the simulation of the diffusion approximations using the Euler–Maruyama method (or the method proposed by Bossy et al. [3]), we used the following time step $h = 0.01$. The simulations were performed up to time $T = 10^4$, and sampled every $\Delta = 0.1$. This sampling was performed in order to compare the Markov chain and the approximations over the same discrete time steps.

Figure 1 gives a scatter plot of the points sampled during the 10 runs of the simulation. In this plot, the points corresponding to the Markov chain simulation lay on the lattice G^V. Notice LE-NR and LE-Chop are shifted upwards. This can be explained by the the influence of the reflection direction at the boundary, since the "correct" reflection direction, as in the CLA, should be oblique, and, near the origin, is approximately 45° from the inward normal, pointing toward the origin. Also, the LE-Chop spends more time at the boundary than the Markov chain, suggesting that this truncated approximation is inappropriate to represent the behavior of the Markov chain near the boundary. In addition, the linear noise approximation predicts negative concentration values. The mean time consumed to generate the



(a) MCM (b) CLA (c) LNA

(d) LE-NR (e) LE-Chop

FIG. 1. (*Example* 1) *"Heat" scatter plot of the points generated by* 10 *runs of the simulations up to time* $T = 10^4$ (*the figures were generated with the LSD R package* [33]). *The axis labels* x_1 *and* x_2 *denote concentrations of species* S_1 *and* S_2, *respectively* (*or their appropriate approximations*).

TABLE 1
(*Example* 1) *Mean running time in seconds for the simulations. The mean running time is calculated by averaging the running times over the* 10 *independent runs. The* 95%*-confidence interval for this average is also calculated*

Model	**Mean running time**	**95%-C.I.**
MCM	310.338 s	(309.624, 311.052)
LNA	320.178 s	(319.645, 320.711)
CLA	308.922 s	(307.686, 310.159)
LE-NR	282.607 s	(281.706, 283.509)
LE-Chop	251.664 s	(251.584, 251.743)

simulations is reported in Table 1. Notice that the time is approximately the same for every algorithm, since the number of molecules of each species is low. This example was chosen to show the influence of the oblique reflection direction at the boundary.

In order to have a more precise measure of the quality of the approximations, we estimated a discrete density by counting the number of points that lay within a regular grid of square bins with side lengths given by $1/V$, which are centered around each point of the Markov chain state space (i.e., the two-dimensional lattice G^V). The density is estimated by counting the number of points within each bin divided by the total number of points and the area of the square bin. This process is performed for every approximation algorithm at each run. In addition, the absolute error is calculated by integrating the modulus of the difference between the discrete density of an approximation and that of the Markov chain. The results are displayed in Table 2. Since the densities integrate to 1, the maximum possible absolute error is 2. Notice that CLA gives the lowest error with respect to the Markov chain.

TABLE 2
(*Example* 1) *Mean absolute error calculated by integrating the absolute difference between the discrete density for the approximation and that of the Markov chain. The mean error is calculated by averaging this result over* 10 *independent runs. The* 95%*-confidence interval for this average is also calculated*

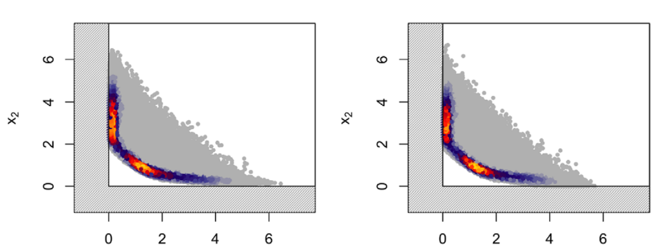
Model	**Mean absolute error**	**95%-C.I.**
LNA	0.395791	(0.393330, 0.398252)
CLA	0.206943	(0.203844, 0.210042)
LE-NR	1.775523	(1.770973, 1.780073)
LE-Chop	1.234770	(1.225156, 1.244384)

8.2. *Example* 2. In this subsection, we consider the Brusselator reaction set, where both species S_1 and S_2 have inflow and outflow reactions. The reactions are as follows:

$$S_1 \underset{c_2}{\overset{c_1}{\rightleftharpoons}} \varnothing, \qquad S_2 \underset{c_4}{\overset{c_3}{\rightleftharpoons}} \varnothing, \qquad S_1 \overset{c_5}{\to} S_2, \qquad 2S_1 + S_2 \overset{c_6}{\to} 3S_1.$$

The following normalized rate constants were chosen for this example: $c_1 = 1$, $c_2 = 1$, $c_3 = 10^{-4}$, $c_4 = 10^{-4}$, $c_5 = 11$ and $c_6 = 10$. With these parameters, the deterministic RRE model for this chemical reaction system exhibits a stable limit cycle. The constant V, which denotes the volume of the vessel times Avogadro's number, is set to 100. The starting point for the simulations was set to $\bar{x}^o = (2, 1)'$. For the simulations of diffusion approximations using the Euler–Maruyama method (or the method proposed by Bossy et al. [3]), we used the following time step $h = 0.01$. The simulations were performed up to time $T = 10^4$, and sampled every $\Delta = 0.1$.

Figure 2 gives a scatter plot of the points generated by the simulations. Here, we can observe that the Markov chain model displays fluctuation around a limit cycle, which the CLA captures well. However, as has been noted in previous works (e.g., [34, 39]), LNA does not capture this behavior properly. In fact, oscillations build up as time progresses. In Figure 3, plots of x_1 and x_2 as functions of time are shown, which shows that LNA oscillates and diverges. The mean times taken for the simulations are given in Table 3. The mean running time for LNA is omitted since the simulations diverge when executed for a long period of time. Notice now that there is a gain in simulating the CLA with respect to the Markov chain since the number of molecules in the system is larger than in Example 1.



(a) MCM (b) CLA

FIG. 2. (*Example* 2) *"Heat" scatter plot of the points generated by* 10 *runs of the simulations up to time* $T = 10^4$ (*the figures were generated with the LSD R package* [33]). *The axis labels* x_1 *and* x_2 *denote concentrations of species* S_1 *and* S_2, *respectively* (*or their appropriate approximations*). *The scatter plot for LNA is not shown since LNA oscillates and diverges during a long simulation.*

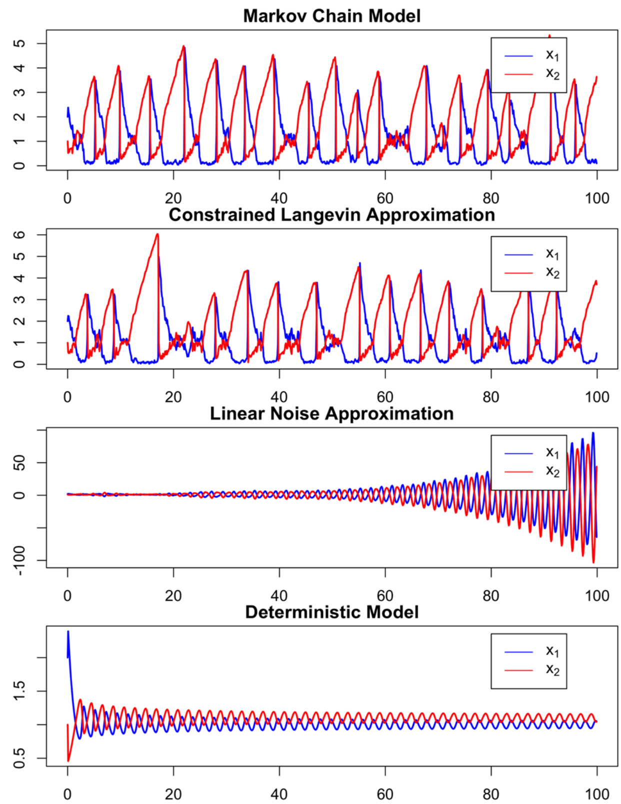


FIG. 3. (*Example* 2) *This figure displays plots of* x_1 *and* x_2 *versus time over a time period of* $T = 100$, *where* x_1 *and* x_2 *denote the concentration of molecular species* S_1 *and* S_2, *respectively* (*or their approximations*). *This illustrates that LNA increases in oscillation and diverges during a long simulation*.

8.3. *Example* 3. For this example, we consider a reaction set that was introduced in [41]. We modify this system slightly by allowing inflow and outflow reactions for every species and also by keeping track of the substrate concentration in order to satisfy the mass-dissipating assumption (given by Assumption 3.1).

TABLE 3
(*Example* 2) *Mean running time in seconds for the simulation of MCM and CLA. The mean running time is calculated by averaging the running times over* 10 *independent runs. The* 95%-*confidence interval from this average is also calculated*

Model	**Mean running time**	**95%-C.I.**
MCM	915.332 s	(912.949, 917.715)
CLA	255.127 s	(254.843, 255.410)

The reactions are as follows:

$$S_1 \underset{c_2}{\overset{c_1}{\rightleftharpoons}} \varnothing, \qquad S_2 \underset{c_4}{\overset{c_3}{\rightleftharpoons}} \varnothing, \qquad S_3 \underset{c_6}{\overset{c_5}{\rightleftharpoons}} \varnothing,$$

$$S_3 + S_2 \xrightarrow{c_7} 2S_1, \qquad 2S_1 \xrightarrow{c_8} S_1 + S_2, \qquad S_1 + S_2 \xrightarrow{c_9} S_2.$$

The following values for the normalized reaction rate constants were used: $c_1 = 1/\sqrt{10}$, $c_2 = 0.01$, $c_3 = 1$, $c_4 = 0.01$, $c_5 = 1$, $c_6 = 10$, $c_7 = 8/10$, $c_8 = 1$, $c_9 = 1.5/\sqrt{10}$. With these parameters, the deterministic model of this system exhibits bistability with two real stable points near $\bar{x}^* \triangleq (1.2679 \cdot 10^{-1}, 2.90328 \cdot 10^{-3}, 9.97683)$ and $\bar{x}^\dagger \triangleq (2.96686, 2.31681, 3.50454)$. For the simulations of diffusion approximations using the Euler–Maruyama method (or the method proposed by Bossy et al. [3]), we used the time step $h = 0.005$. The constant V was again set to 100 and the starting point for the simulations was fixed at $\bar{x}^o = (0.1, 0.1, 10)$.

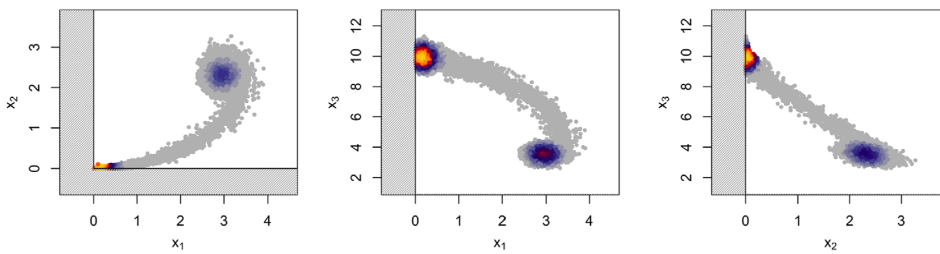
With this starting point, the deterministic model converges to the first stable point $\bar{x}^*$. However, the Markov chain model, starting at the same point $\bar{x}^o$, spends some time near the first stable point $\bar{x}^*$, but eventually moves away to the vicinity of the second stable point $\bar{x}^\dagger$, where it stays for a longer period of time.

In order to capture this change of stable points and illustrate this behavior, the system was simulated during a shorter time period of $T = 100$. Simulation samples were collected for every time period of size $\Delta = 0.1$. A total of 100 runs of duration T were constructed for each algorithm and used to generate the scatter plot of Figure 4. From this figure, we see that the Markov chain spends some time near $\bar{x}^*$, but eventually reaches $\bar{x}^\dagger$. This behavior is reproduced by the CLA. However, since LNA is a *linear* noise approximation, it only captures the behavior near the first stable point.

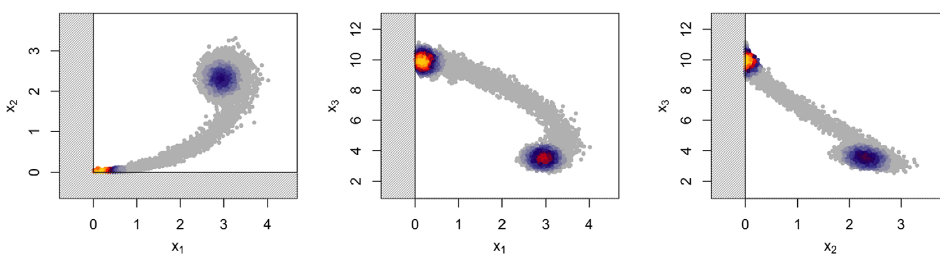
Table 4 shows the average running time taken for each simulation. Again, we can observe that the mean simulation time for CLA is shorter than that for the Markov chain. For the LNA, the mean running time is longer than the time for the CLA due to the fact that we used the fourth-order Runge–Kutta method for its deterministic part. (It took on average 3.301 seconds to run LNA using Euler's method for its deterministic part.)

APPENDIX: PROOFS OF SOME LEMMAS

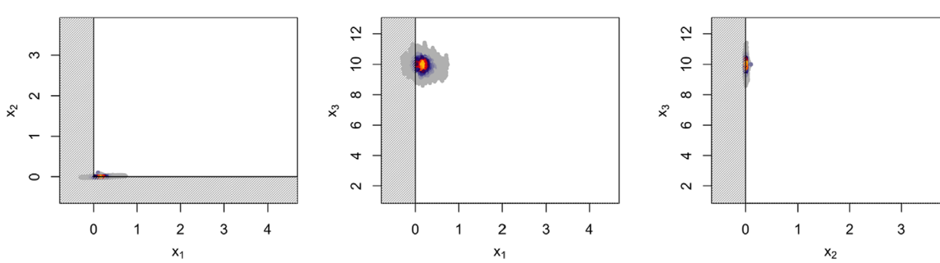
PROOF OF LEMMA 3.1. Let ζ denote the explosion time for X: $\zeta = \lim_{n\to\infty} \tau_n$, where $\tau_n = \inf\{t \ge 0 : |X(t)| \ge n\}$. We shall prove that $\zeta = \infty$ a.s. For a proof by contradiction, suppose that $\mathbb{P}(\zeta < \infty) > 0$ and let $t > 0$ be such that $\mathbb{P}(\Omega_t) > 0$ for $\Omega_t = \{\zeta < t\}$. Let $\mathcal{I}$ denote the index set for the external input reactions. Let $u \in \mathbb{R}^d$, $u \ge \mathbb{1}$, be a vector such that $\langle u, v_k\rangle \le 0$ for all k such that



(a) MCM



(b) CLA



(c) LNA

FIG. 4. (*Example* 3) *"Heat" scatter plot of the points generated by* 100 *runs of the simulations up to* $T = 100$ (*the figures were generated with the LSD R package* [33]). *The axis labels* x_1, x_2 *and* x_3 *denote concentrations of species* S_1, S_2 *and* S_3, *respectively* (*or their appropriate approximations*).

TABLE 4
(*Example* 3) *Mean running time in seconds for each simulation. The mean running time is calculated by averaging the running times over* 100 *independent runs. The* 95%-*confidence interval from this average is also calculated*

Model	**Mean running time**	**95%-C.I.**
MCM	9.1989 s	(8.782, 9.615)
LNA	4.4227 s	(4.398, 4.448)
CLA	2.5721 s	(2.518, 2.626)

$R_k \in \mathscr{R}_1$. Then on Ω_t, for $0 \le s < \zeta < t$, we have

$$\begin{aligned}\langle u, X(s)\rangle &= \langle u, X(0)\rangle + \sum_{k=1}^{r} \langle u, v_k\rangle N_k\left(\int_0^s \Lambda_k(X(w))\,dw\right)\\ &\le \langle u, X(0)\rangle + \sum_{k\in\mathcal{I}} \langle u, v_k\rangle N_k\left(\int_0^s \Lambda_k(X(w))\,dw\right) \qquad \text{(A.1)}\\ &\le \langle u, X(0)\rangle + \max_{k\in\mathcal{I}}\{\langle u, v_k\rangle\} \sum_{k\in\mathcal{I}} N_k(\kappa_k t),\end{aligned}$$

where we used the facts that $\langle u, v_k\rangle \le 0$ for all $k \notin \mathcal{I}$ in the first inequality, that $\Lambda_k(x) = \kappa_k$ for all $k \in \mathcal{I}$, and that N_k is nondecreasing and $s < t$ in the last line.

Then, on Ω_t,

$$\begin{aligned}\sup_{s<\zeta} \sum_{i=1}^{d} |X_i(s)| &\le \sup_{s<\zeta} \langle u, X(s)\rangle\\ &\le \langle u, X(0)\rangle + \max_{k\in\mathcal{I}}\{\langle u, v_k\rangle\} \sum_{k\in\mathcal{I}} N_k(\kappa_k t) < \infty \qquad \text{a.s.},\end{aligned}$$

since $N_k(\kappa_k t) < \infty$ a.s. for all $k \in \mathcal{I}$. Since all Euclidean norms are equivalent, we have that $\sup_{s<\zeta} |X(s)| < \infty$ a.s. on Ω_t. However, by the definition of ζ, $\sup_{s<\zeta} |X(s)| = \infty$ on $\{\zeta < \infty\}$, which contains Ω_t. This contradicts the assumption that $\mathbb{P}(\Omega_t) > 0$. $\square$

PROOF OF LEMMA 4.1. Let $x \in G^b$ and $j \in I(x)$. Using the fact that $x_j = 0$ and the definition of λ_k, we have

$$\begin{aligned}\langle \mu(x), n_j\rangle &= \sum_{k=1}^{r} c_k\left(\prod_{l=1}^{d} x_l^{v_{lk}^-}\right)\langle v_k, n_j\rangle = \sum_{k=1}^{r} c_k\left(\prod_{l=1(l\neq j)}^{d} x_l^{v_{lk}^-}\right) x_j^{v_{jk}^-} v_{jk}\\ &= \sum_{k:v_{jk}^-=0} c_k\left(\prod_{l=1}^{d} x_l^{v_{lk}^-}\right)(v_{jk}^+ - 0), \qquad \text{(A.2)}\end{aligned}$$

since the kth summand in the second equality will be zero if $v_{jk}^- > 0$ because $x_j = 0$. Now the summands in the last equality in (A.2) are nonnegative and so by Assumption 3.2(a), which implies that all species have inflows, we have

$$\langle \mu(x), n_j\rangle \ge c_{k_j^\uparrow} \triangleq \alpha_j > 0, \qquad \text{(A.3)}$$

where the last line defines the constant α_j.

Now, let $x \in G$ and let $\mathcal{I}$ be the set of indexes associated with external input reactions. Then we have

$$\langle \mu(x), -u\rangle = \sum_{k=1}^{r} \lambda_k(x)\langle v_k, -u\rangle = \sum_{k\notin\mathcal{I}} \lambda_k(x)\left|\langle v_k, u\rangle\right| - \sum_{k\in\mathcal{I}} c_k\left|\langle v_k, u\rangle\right|, \qquad \text{(A.4)}$$

where we used the fact that $\langle v_k, u\rangle \le 0$ for $k \notin \mathcal{I}$, by Assumption 3.1, and that $\langle v_k, u\rangle > 0$ for $k \in \mathcal{I}$, since $u \ge \mathbb{1}$. Let $k_i^{\downarrow}$, $i = 1, \dots, d$, be as in Assumption 3.2, then

$$\begin{aligned}\langle \mu(x), -u\rangle &\ge \Bigl(\min_{i=1}^{d} c_{k_i^{\downarrow}}\Bigr) \sum_{i=1}^{d} x_i u_i - \sum_{k\in\mathcal{I}} c_k \bigl|\langle v_k, u\rangle\bigr| \\ &= \Bigl(\min_{i=1}^{d} c_{k_i^{\downarrow}}\Bigr)\langle x, u\rangle - \sum_{k\in\mathcal{I}} c_k \bigl|\langle v_k, u\rangle\bigr|.\end{aligned} \tag{A.5}$$

Then, there exists an $\alpha_{d+1} > 0$ such that $\langle \mu(x), -u\rangle \ge \alpha_{d+1}$ for all $x \in G$ such that $\langle x, u\rangle > M^*$. Therefore the proof is complete with $\alpha = \min_{j=1}^{d+1} \alpha_j$. $\square$

PROOF OF LEMMA 4.2. Let $\delta \in (0, \frac{1}{\sqrt{V}}]$, $x \in G^b$ and $j \in I(x)$. Then, using analogous arguments to those used to derive (A.2), we have that

$$\langle \tilde{\mu}^{\delta}(x), n_j\rangle = \sum_{k: v_{jk}^- = 0} c_k \Biggl(\prod_{l=1}^{d} x_l^{v_{lk}^-} \mathbf{1}_{\{x_l \ge \delta v_{lk}^-/\sqrt{V}\}}\Biggr)(v_{jk}^+ - 0).$$

Since the summands on the right-hand side of the equation are all nonnegative and at least one of them corresponds to an input reaction by Assumption 3.2, we have that

$$\langle \tilde{\mu}^{\delta}(x), n_j\rangle \ge c_{k_j^{\uparrow}} \triangleq \alpha_j > 0,$$

where the constant α_j is defined by the last line.

Now let $x \in G$ be such that $\langle x, u\rangle > \tilde{M}^*$. Define the vectors $x^{\delta}, \bar{x}^{\delta} \in \mathbb{R}^d$ as follows:

$$x_i^{\delta} = x_i \mathbf{1}_{\{x_i \ge \delta/\sqrt{V}\}} \quad \text{and} \quad \bar{x}_i^{\delta} = \frac{\delta}{\sqrt{V}} \mathbf{1}_{\{x_i < \delta/\sqrt{V}\}},$$

for $1 \le i \le d$. Then we have that $\langle x^{\delta} + \bar{x}^{\delta}, u\rangle \ge \langle x, u\rangle > \tilde{M}^* = M^* + \sum_{i=1}^{d} u_i/V$, which implies that

$$\langle x^{\delta}, u\rangle > M^* + \frac{1}{V}\sum_{i=1}^{d} u_i - \langle \bar{x}^{\delta}, u\rangle > M^* + \frac{1}{V}\sum_{i=1}^{d} u_i - \frac{\delta}{\sqrt{V}}\sum_{i=1}^{d} u_i \ge M^*, \tag{A.6}$$

since $\delta \le 1/\sqrt{V}$. Then, by analogous arguments to those used to derive (A.5), we have that

$$\begin{aligned}\langle \tilde{\mu}^{\delta}(x), -u\rangle &\ge \min_{i=1}^{d} c_{k_i^{\downarrow}} \sum_{i=1}^{d} u_i x_i \mathbf{1}_{\{x_i \ge \delta v_{ik_i^{\downarrow}}^-/\sqrt{V}\}} - \sum_{k\in\mathcal{I}} c_k \bigl|\langle v_k, u\rangle\bigr| \\ &= \Bigl(\min_{i=1}^{d} c_{k_i^{\downarrow}}\Bigr)\langle u, x^{\delta}\rangle - \sum_{k\in\mathcal{I}} c_k \bigl|\langle v_k, u\rangle\bigr|,\end{aligned}$$

since $v^-_{k_i^\downarrow} = e_i$. Then (A.6) implies that there is an $\alpha_{d+1} > 0$ such that $\langle \tilde{\mu}^\delta(x), -u\rangle \geq \alpha_{d+1}$. The result follows by setting $\alpha = \min_{j=1}^{d+1} \alpha_j$. $\square$

PROOF OF LEMMA 4.3. Fix $m > 0$, $M > m$, and suppose that Z^δ starts from $z \in G$ satisfying $\langle u, z\rangle \leq m$. Then by (4.12)–(4.17), and the facts that $\mu(x) = \sum_{k=1}^r v_k \lambda_k(x)$ for all $x \in G$, $\langle u, v_k\rangle \leq 0$ for all $k \notin \mathcal{I}$, and $\lambda_k(x) = c_k$ for $k \in \mathcal{I}$, we have that for $t \geq 0$,

$$\begin{aligned}\langle u, Z^\delta(t \wedge \zeta_M^\delta)\rangle \leq \langle u, z\rangle &+ \sum_{k \in \mathcal{I}} \langle u, v_k\rangle \int_0^{t\wedge\zeta_M^\delta} c_k \mathbf{1}_{\{Z^\delta(s) \in G^\circ\}}\,ds \\ &+ \frac{1}{\sqrt{V}} \int_0^{t\wedge\zeta_M^\delta} \langle u, \sigma(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\,dW(s) \\ &+ \frac{\delta}{\sqrt{V}} \sum_{k\in\mathcal{I}} \langle u, v_k\rangle N_k^b\left(\delta^{-2}\int_0^{t\wedge\zeta_M^\delta} c_k \mathbf{1}_{\{Z^\delta(s)\in G^b\}}\,ds\right).\end{aligned} \tag{A.7}$$

Let

$$C^\delta(t) = \left(1 + \frac{\delta^{-1}}{\sqrt{V}}\right) \sum_{k\in\mathcal{I}} |\langle u, v_k\rangle| c_k t. \tag{A.8}$$

Then, on taking expectations in (A.7), we have that

$$\mathbb{E}[\langle u, Z^\delta(t \wedge \zeta_M^\delta)\rangle] \leq m + C^\delta(t), \tag{A.9}$$

where we have used the martingale properties of the stopped stochastic integral with respect to W and the fact that

$$N_k^b\left(\delta^{-2}\int_0^{t\wedge\zeta_M^\delta} c_k \mathbf{1}_{\{Z^\delta(s)\in G^b\}}\,ds\right) - \delta^{-2}\int_0^{t\wedge\zeta_M^\delta} c_k \mathbf{1}_{\{Z^\delta(s)\in G^b\}}\,ds$$

is a martingale with respect to a right-continuous filtration to which it is adapted. From (A.9), we conclude that

$$M\mathbb{P}(\zeta_M^\delta \leq t) \leq m + C^\delta(t) \tag{A.10}$$

and so

$$\sup_{z\in G:\langle u,z\rangle\leq m} P_z^\delta(\zeta_M \leq t) \leq \frac{m + C^\delta(t)}{M}, \tag{A.11}$$

from which the desired result follows. $\square$

PROOF OF LEMMA 4.4. Recall the definitions of ρ and Θ_m from (4.29) and (4.31). For $0 < \delta \leq \frac{1}{\sqrt{V}}$, $t \geq 0$, $m \geq \tilde{M}^* + \rho$, $M > m + 5\rho$, let

$$C_{m,M}^\delta(t) = \sup_{z\in\Theta_m} P_z^\delta(\zeta_M \leq t). \tag{A.12}$$

For any $z \in G$ such that $m \leq \langle u, z\rangle \leq M$, for Z^δ starting from z, we have by (4.12)–(4.17), (4.23) and (4.30) that for each $t \geq 0$,

$$\begin{aligned}\langle u, Z^\delta(t \wedge \tau_m^\delta \wedge \zeta_M^\delta)\rangle &= \langle u, z\rangle + \int_0^{t\wedge\tau_m^\delta\wedge\zeta_M^\delta} \langle u, \mu(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\, ds \\ &\quad + \frac{1}{\sqrt{V}}\int_0^{t\wedge\tau_m^\delta\wedge\zeta_M^\delta} \langle u, \sigma(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\, dW(s) \\ &\quad + \frac{1}{\sqrt{V}}\langle u, \hat{Y}^\delta(t\wedge\tau_m^\delta\wedge\zeta_M^\delta)\rangle \\ &\quad + \frac{\delta^{-1}}{\sqrt{V}}\int_0^{t\wedge\tau_m^\delta\wedge\zeta_M^\delta} \langle u, \tilde{\mu}^\delta(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^b\}}\, ds.\end{aligned} \tag{A.13}$$

Since $M^* \leq \tilde{M}^* < m \leq \langle u, Z^\delta(s)\rangle \leq M$ for all $s \in (0, t\wedge\tau_m^\delta\wedge\zeta_M^\delta)$, it follows from Lemmas 4.1 and 4.2 that there is a constant $\alpha > 0$ such that for all $s \in (0, t\wedge\tau_m^\delta\wedge\zeta_M^\delta)$,

$$\langle u, \mu(Z^\delta(s))\rangle \leq -\alpha \quad \text{and} \quad \langle u, \tilde{\mu}^\delta(Z^\delta(s))\rangle \leq -\alpha. \tag{A.14}$$

Thus, we have that

$$\begin{aligned}&\langle u, Z^\delta(t\wedge\tau_m^\delta\wedge\zeta_M^\delta)\rangle \\ &\quad\leq \langle u, z\rangle + \frac{1}{\sqrt{V}}\int_0^{t\wedge\tau_m^\delta\wedge\zeta_M^\delta} \langle u, \sigma(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\, dW(s) \\ &\qquad + \frac{1}{\sqrt{V}}\langle u, \hat{Y}^\delta(t\wedge\tau_m^\delta\wedge\zeta_M^\delta)\rangle.\end{aligned} \tag{A.15}$$

Upon taking expectations in the above expression, using the martingale property of the stopped stochastic integrals with respect to W and of the stopped compensated Poisson process $\hat{Y}^\delta(\cdot\wedge\tau_m^\delta\wedge\zeta_M^\delta)$, we obtain

$$\mathbb{E}[\langle u, Z^\delta(t\wedge\tau_m^\delta\wedge\zeta_M^\delta)\rangle] \leq \langle u, z\rangle. \tag{A.16}$$

Writing this in terms of the law of Z^δ on the canonical space $(\mathcal{D}^\partial, \mathcal{M}^\partial)$, we have

$$E^{P_z^\delta}[\langle u, \omega(t\wedge\tau_m\wedge\zeta_M)\rangle] \leq \langle u, z\rangle, \tag{A.17}$$

where $\tau_m = \inf\{s \geq 0 : \langle u, \omega(s)\rangle < m\}$, $\tau_m = +\infty$ if $\omega(s) = \partial$ for all $s \geq 0$, and $\langle u, \partial\rangle = 0$. We will shortly prove that for any $m \geq \tilde{M}^* + \rho$, $M > m + 5\rho$, $m \leq \langle u, z\rangle \leq M$,

$$P_z^\delta(\tau_m \wedge \zeta_M < \infty) = 1. \tag{A.18}$$

Assuming this holds for the moment, on letting $t \to \infty$ in (A.17), we obtain

$$E^{P_z^\delta}[\langle u, \omega(\tau_m\wedge\zeta_M)\rangle] \leq \langle u, z\rangle. \tag{A.19}$$

Taking account of the fact that $\omega(\cdot)$ may jump as it exits $\{x \in G : m \leq \langle u, x\rangle \leq M\}$, we have that P_z^δ-a.s., $\langle u, \omega(\tau_m)\rangle \geq m - \rho$ on $\{\tau_m < \zeta_M\}$ and $M \leq \langle u, \omega(\zeta_M)\rangle$ on $\{\zeta_M < \tau_m\}$. Hence, (A.19) yields

$$(m-\rho)P_z^\delta(\tau_m < \zeta_M) + MP_z^\delta(\zeta_M < \tau_m) \leq \langle u, z\rangle. \tag{A.20}$$

This implies that

$$P_z^\delta(\zeta_M < \tau_m) \leq \frac{\langle u, z\rangle - m + \rho}{M - m + \rho}, \tag{A.21}$$

and so for any $m \geq \tilde{M}^* + \rho$, $M > m + 5\rho$, $m \leq K < M$, we have

$$\sup_{z \in G : m \leq \langle u, z\rangle \leq K} P_z^\delta(\zeta_M < \tau_m) \leq \frac{K - m + \rho}{M - m + \rho}. \tag{A.22}$$

Now, for $m \geq \tilde{M}^* + \rho$, $\tilde{m} > m + 4\rho$ and $M > \tilde{m} + \rho$, for $z \in \Theta_{\tilde{m}}$, we have by the strong Markov property that

$$\begin{aligned} P_z^\delta(\tau_{m+\rho} < \zeta_M \leq t) &\leq E^{P_z^\delta}\big[\mathbf{1}_{\{\tau_{m+\rho} < \zeta_M\}} P^\delta_{\omega(\tau_{m+\rho})}(\zeta_M \leq t)\big] \\ &\leq P_z^\delta(\tau_{m+\rho} < \zeta_M) C^\delta_{m,M}(t) \\ &\leq C^\delta_{m,M}(t). \end{aligned} \tag{A.23}$$

Then for $z \in \Theta_m$, using the strong Markov property again, we have

$$\begin{aligned} &P_z^\delta(\zeta_M \leq t) \\ &\quad = P_z^\delta(\zeta_{\tilde{m}} \leq t, \zeta_M \leq t) \\ &\quad \leq E^{P_z^\delta}\big[\mathbf{1}_{\{\zeta_{\tilde{m}} \leq t\}} P^\delta_{\omega(\zeta_{\tilde{m}})}(\zeta_M \leq t)\big] \\ &\quad = E^{P_z^\delta}\big[\mathbf{1}_{\{\zeta_{\tilde{m}} \leq t\}}\big(P^\delta_{\omega(\zeta_{\tilde{m}})}(\tau_{m+\rho} < \zeta_M \leq t) \\ &\qquad + P^\delta_{\omega(\zeta_{\tilde{m}})}(\zeta_M < \tau_{m+\rho}, \zeta_M \leq t)\big)\big] \\ &\quad \leq E^{P_z^\delta}\Big[\mathbf{1}_{\{\zeta_{\tilde{m}} \leq t\}}\Big(C^\delta_{m,M}(t) + \sup_{x \in \Theta_{\tilde{m}}} P_x^\delta(\zeta_M < \tau_{m+\rho})\Big)\Big] \\ &\quad \leq E^{P_z^\delta}\left[\mathbf{1}_{\{\zeta_{\tilde{m}} \leq t\}}\left(C^\delta_{m,M}(t) + \frac{\tilde{m} - m + \rho}{M - m}\right)\right], \end{aligned} \tag{A.24}$$

for $M > m + 5\rho$, where we used (A.23) in the second to last line and (A.22) in the last line, with m replaced by $m + \rho$ and $K = \tilde{m} + \rho$.

Letting $\varepsilon^\delta_{m,\tilde{m}}(t) = \sup_{z \in \Theta_m} P_z^\delta(\zeta_{\tilde{m}} \leq t)$ and taking the supremum over $z \in \Theta_m$ in (A.24), we have

$$C^\delta_{m,M}(t) \leq \varepsilon^\delta_{m,\tilde{m}}(t)\left(C^\delta_{m,M}(t) + \frac{\tilde{m} - m + \rho}{M - m}\right). \tag{A.25}$$

If we can prove that there is δ_0 (depending on $m, \tilde{m}, t$) such that $0 < \delta_0 \leq \frac{1}{\sqrt{V}}$ and

$$\hat{\varepsilon}_{m,\tilde{m}}(t) \triangleq \sup_{0<\delta\leq\delta_0} \varepsilon^{\delta}_{m,\tilde{m}}(t) < 1, \tag{A.26}$$

then we have that

$$\sup_{0<\delta\leq\delta_0} C^{\delta}_{m,M}(t) \leq \frac{\hat{\varepsilon}_{m,\tilde{m}}(t)}{1-\hat{\varepsilon}_{m,\tilde{m}}(t)} \frac{\tilde{m}-m+\rho}{M-m}, \tag{A.27}$$

where the right member tends to zero as $M \to \infty$, which proves the desired result.

Thus, it remains to prove (A.26) and (A.18). We first prove (A.18). For m, M as at the beginning of this proof, Z^δ starting from $z \in G$ satisfying $m \leq \langle u, z\rangle \leq M$, on $\{\tau^\delta_m \wedge \zeta^\delta_M = +\infty\}$, we have from (A.15) that for all $t \geq 0$,

$$\langle u, Z^\delta(t)\rangle \leq \langle u, z\rangle + \frac{1}{\sqrt{V}} \mathcal{M}^\delta(t), \tag{A.28}$$

where

$$\mathcal{M}^\delta(t) \triangleq \int_0^t \langle u, \sigma(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\, dW(s) + \langle u, \hat{Y}^\delta(t)\rangle, \qquad t \geq 0, \tag{A.29}$$

defines a local martingale with respect to the filtration $\{\mathcal{F}_t\}$ and $\hat{Y}^\delta$ is given by (4.17). The predictable quadratic variation of $\mathcal{M}^\delta$ is given by

$$\langle \mathcal{M}^\delta\rangle(t) = \int_0^t u'\Gamma(Z^\delta(s))u\mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\, ds + \int_0^t u'\tilde{\Gamma}^\delta(Z^\delta(s))u\mathbf{1}_{\{Z^\delta(s)\in G^b\}}\, ds,$$

where

$$\tilde{\Gamma}^\delta(x) \triangleq \sum_{k=1}^r v_k v_k' \tilde{\lambda}^\delta_k(x), \qquad x \in G. \tag{A.30}$$

In a similar manner to that for (4.9), we have

$$\Bigl(\min_{i=1}^d c_{k_i^\uparrow}\Bigr)|\theta|^2 \leq \langle \theta, \tilde{\Gamma}^\delta(x)\theta\rangle \leq \tilde{K}(x)|\theta|^2 \qquad \text{for all } \theta \in \mathbb{R}^d, x \in G^b, \tag{A.31}$$

where $\tilde{K}(x) = \sum_{k=1}^r |v_k|^2 \lambda_k(x) \geq \sum_{k=1}^r |v_k|^2 \tilde{\lambda}^\delta_k(x)$, since $\tilde{\lambda}^\delta_k$ is a truncation of λ_k. It then follows that

$$\langle \mathcal{M}^\delta\rangle(t) \geq \Bigl(\min_{i=1}^d c_{k_i^\uparrow}\Bigr)|u|^2 t \to \infty \qquad \text{as } t \to \infty, \tag{A.32}$$

where we have used (4.9) and (A.31). By a law of the iterated logarithm for local martingales [24] (see also [40]), it follows that a.s., on $\{\tau^\delta_m \wedge \zeta^\delta_M = +\infty\}$,

$$\liminf_{t\to\infty} \frac{\mathcal{M}^\delta(t)}{\sqrt{\langle \mathcal{M}^\delta\rangle(t)}} = -\infty. \tag{A.33}$$

Combining (A.28), (A.32) and (A.33), we see that on $\{\tau_m^\delta \wedge \zeta_M^\delta = +\infty\}$, a.s.,

$$\liminf_{t\to\infty} \langle u, Z^\delta(t)\rangle = -\infty. \tag{A.34}$$

But on $\{\tau_m^\delta \wedge \zeta_M^\delta = +\infty\}$,

$$\langle u, Z^\delta(t)\rangle \geq m, \tag{A.35}$$

for all $t \geq 0$, and so we must have that

$$\mathbb{P}(\tau_m^\delta \wedge \zeta_M^\delta = +\infty) = 0. \tag{A.36}$$

Hence, (A.18) holds.

We now fix $m \geq \tilde{M}^* + \rho, \tilde{m} > m + 4\rho, t \geq 0$, and prove that (A.26) holds for some $\delta_0 \in (0, \frac{1}{\sqrt{V}}]$. For Z^δ starting from $z \in \Theta_m$, in a similar manner to that in which (A.28) was derived, we have that if $0 \leq s_1 < s_2 \leq t$ are such that $m \leq \langle u, Z^\delta(s)\rangle \leq \tilde{m}$ for $s_1 < s < s_2$, then

$$\langle u, Z^\delta(s_2)\rangle - \langle u, Z^\delta(s_1)\rangle \leq \frac{1}{\sqrt{V}}\big(\mathcal{M}^\delta(s_2) - \mathcal{M}^\delta(s_1)\big), \tag{A.37}$$

where $\mathcal{M}^\delta$ is defined by (A.29). Now, the stochastic integral with respect to W in $\mathcal{M}^\delta$ can be time-changed to a Brownian motion, $B^{\delta,1}$, so that

$$\mathcal{M}^{\delta,1}(v) \triangleq \int_0^v \langle u, \sigma\big(Z^\delta(s)\big)\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\, dW(s) = B^{\delta,1}\big(\langle \mathcal{M}^{\delta,1}\rangle(v)\big), \tag{A.38}$$

where the quadratic variation of $\mathcal{M}^{\delta,1}$ is given by

$$\langle \mathcal{M}^{\delta,1}\rangle(v) = \int_0^v \langle u, \Gamma\big(Z^\delta(s)\big)u\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\, ds. \tag{A.39}$$

Let $C(V, m, \tilde{m}) = \frac{\sqrt{V}(\tilde{m}-(m+4\rho))}{4}$ and $K_{\tilde{m}}$ be a constant such that for all $x \in G$ satisfying $\langle u, x\rangle \leq \tilde{m}$ we have

$$u'\tilde{\Gamma}^\delta(x)u \leq u'\Gamma(x)u \leq K_{\tilde{m}}. \tag{A.40}$$

Then

$$\begin{aligned} &\mathbb{P}\Big(\max_{0\leq s\leq t} \big|\mathcal{M}^{\delta,1}\big(s \wedge \zeta_{\tilde{m}}^\delta\big)\big| \geq C(V, m, \tilde{m})\Big) \\ &\quad\leq \mathbb{P}\big(\langle \mathcal{M}^{\delta,1}\rangle\big(t \wedge \zeta_{\tilde{m}}^\delta\big) \geq \Upsilon^\delta(m, \tilde{m})\big) \\ &\quad\leq \mathbb{P}\big(K_{\tilde{m}} t \geq \Upsilon^\delta(m, \tilde{m})\big) = \beta, \end{aligned} \tag{A.41}$$

where $\Upsilon^\delta(m, \tilde{m}) = \inf\{s \geq 0 : |B^{\delta,1}(s)| \geq C(V, m, \tilde{m})\}$ is the first time that the absolute value of a one-dimensional Brownian motion is of size $C(V, m, \tilde{m})$, and $\beta \in (0, 1)$ depends only on $t, m, \tilde{m}$, but not on δ or the starting point $z \in \Theta_m$ for Z^δ.

Considering upcrossings from where $\langle u, Z^\delta\rangle \le m + 2\rho$ to where $\langle u, Z^\delta\rangle > \tilde{m}$ that are needed to ensure that $\zeta^\delta_{\tilde{m}} \le t$, for any $0 < \delta \le \delta_0$, $z \in \Theta_m$ with $Z^\delta(0) = z$, we have

$$\begin{aligned}
P^\delta_z(\zeta_{\tilde{m}} \le t) &\le \mathbb{P}\big(\langle u, Z^\delta(s_2)\rangle - \langle u, Z^\delta(s_1)\rangle \ge \tilde{m} - (m + 4\rho) \\
&\qquad \text{and } m + 2\rho < \langle u, Z^\delta(s)\rangle < \tilde{m} \text{ for all } s \in (s_1, s_2), \\
&\qquad \text{for some } 0 \le s_1 < s_2 \le \zeta^\delta_{\tilde{m}} \le t\big) \\
&\le \mathbb{P}\Big(\max_{0\le s\le t} \big|\mathcal{M}^{\delta,1}(s \wedge \zeta^\delta_{\tilde{m}})\big| \ge C(V, m, \tilde{m})\Big) \\
&\qquad + \mathbb{P}\big(\langle u, \hat{Y}^\delta(s_2)\rangle - \langle u, \hat{Y}^\delta(s_1)\rangle \ge \sqrt{V}\big(\tilde{m} - (m + 4\rho)\big)/2 \\
&\qquad \text{and } m + 2\rho < \langle u, Z^\delta(s)\rangle < \tilde{m} \text{ for all } s \in (s_1, s_2), \\
&\qquad \text{for some } 0 \le s_1 < s_2 \le \zeta^\delta_{\tilde{m}} \le t\big).
\end{aligned} \tag{A.42}$$

By the preceding paragraph, the second last probability in the above is bounded by $\beta \in (0, 1)$ that is independent of δ and $z \in \Theta_m$. We now focus on showing that the last probability can be made arbitrarily small [in particular, less than $(1 - \beta)/2$, uniformly for all $z \in \Theta_m$], provided that δ is sufficiently small.

First, notice that for $z \in G$ satisfying $m + \rho \le \langle u, z\rangle \le \tilde{m}$, (A.13) holds with the stopping time $\tau^\delta_m \wedge \zeta^\delta_M$ replaced by $\tau^\delta_{m+\rho} \wedge \zeta^\delta_{\tilde{m}}$. Therefore, using (A.14) and in a similar manner to how (A.16) was derived, we have

$$\begin{aligned}
m &\le \mathbb{E}\big[\big\langle u, Z^\delta\big(t \wedge \tau^\delta_{m+\rho} \wedge \zeta^\delta_{\tilde{m}}\big)\big\rangle\big] \\
&\le \langle u, z\rangle - \frac{\delta^{-1}\alpha}{\sqrt{V}} \mathbb{E}\bigg[\int_0^{t\wedge\tau^\delta_{m+\rho}\wedge\zeta^\delta_{\tilde{m}}} \mathbf{1}_{\{Z^\delta(s)\in G^b\}}\, ds\bigg].
\end{aligned}$$

Hence, for any $z \in G$ satisfying $m + \rho \le \langle u, z\rangle \le m + 3\rho$, we have

$$E^{P^\delta_z}\bigg[\int_0^{t\wedge\tau_{m+\rho}\wedge\zeta_{\tilde{m}}} \mathbf{1}_{\{\omega(s)\in G^b\}}\, ds\bigg] \le \frac{\sqrt{V}3\rho\delta}{\alpha}. \tag{A.43}$$

We shall use this inequality further below.

Now, let $\tau^{\delta,0}_{m+\rho} = 0$ and for $k = 1, 2, \ldots,$ inductively define

$$\begin{aligned}
\zeta^{\delta,k-1}_{m+2\rho} &= \inf\big\{s \ge \tau^{\delta,k-1}_{m+\rho} : \big\langle u, Z^\delta(s)\big\rangle > m + 2\rho\big\}, \\
\tau^{\delta,k}_{m+\rho} &= \inf\big\{s \ge \zeta^{\delta,k-1}_{m+2\rho} : \big\langle u, Z^\delta(s)\big\rangle < m + \rho\big\}.
\end{aligned}$$

Also, in order to simplify the notation in the expression below, let us define $D_k \triangleq \max_{\zeta^{\delta,k-1}_{m+2\rho}\le s\le\tau^{\delta,k}_{m+\rho}\wedge t\wedge\zeta^\delta_{\tilde{m}}} |\langle u, \hat{Y}^\delta(s)\rangle - \langle u, \hat{Y}^\delta(\zeta^{\delta,k-1}_{m+2\rho})\rangle|$ on $\{\zeta^{\delta,k-1}_{m+2\rho} \le t \wedge \zeta^\delta_{\tilde{m}}\}$. Then, for $Z^\delta(0) = z \in \Theta_m$, we have

$$
\begin{aligned}
&\mathbb{P}\big(\langle u, \hat{Y}^\delta(s_2)\rangle - \langle u, \hat{Y}^\delta(s_1)\rangle \geq \sqrt{V}\big(\tilde{m} - (m+4\rho)\big)/2 \text{ and}\\
&\quad m + 2\rho < \langle u, Z^\delta(s)\rangle < \tilde{m} \text{ for all } s \in (s_1, s_2),\\
&\quad \text{for some } 0 \leq s_1 < s_2 \leq \zeta^\delta_{\tilde{m}} \leq t\big)\\
&\qquad \leq \sum_{k=1}^\infty \mathbb{P}\big(D_k \geq C(V, m, \tilde{m}) \text{ and } \zeta^{\delta,k-1}_{m+2\rho} \leq t \wedge \zeta^\delta_{\tilde{m}}\big)\\
&\qquad \leq \sum_{k=1}^\infty \mathbb{E}\big[\mathbb{P}\big(D_k \geq C(V, m, \tilde{m}) | \mathcal{F}_{\zeta^{\delta,k-1}_{m+2\rho}}\big)\mathbf{1}_{\{\zeta^{\delta,k-1}_{m+2\rho} \leq t \wedge \zeta^\delta_{\tilde{m}}\}}\big].
\end{aligned} \tag{A.44}
$$

On $\{\zeta^{\delta,k-1}_{m+2\rho} \leq t \wedge \zeta^\delta_{\tilde{m}}\}$, we have, using a conditional version of Doob's L^2-maximal inequality and the strong Markov property of Z^δ, that

$$
\begin{aligned}
&\mathbb{P}\Big(\max_{\zeta^{\delta,k-1}_{m+2\rho} \leq s \leq \tau^{\delta,k}_{m+\rho} \wedge t \wedge \zeta^\delta_{\tilde{m}}} \big|\langle u, \hat{Y}^\delta(s)\rangle - \langle u, \hat{Y}^\delta(\zeta^{\delta,k-1}_{m+2\rho})\rangle\big| \geq C(V, m, \tilde{m}) \Big| \mathcal{F}_{\zeta^{\delta,k-1}_{m+2\rho}}\Big)\\
&\qquad \leq \frac{4}{(C(V, m, \tilde{m}))^2}\mathbb{E}\Big[\int_{\zeta^{\delta,k-1}_{m+2\rho}}^{\tau^{\delta,k}_{m+\rho} \wedge t \wedge \zeta^\delta_{\tilde{m}}} u'\tilde{\Gamma}^\delta\big(Z^\delta(s)\big)u\mathbf{1}_{\{Z^\delta(s) \in G^b\}}\, ds \Big| \mathcal{F}_{\zeta^{\delta,k-1}_{m+2\rho}}\Big]\\
&\qquad \leq \frac{4K_{\tilde{m}}}{(C(V, m, \tilde{m}))^2} E^{P^\delta_{\zeta^{\delta,k-1}_{m+2\rho}}}\Big[\int_0^{\tau_{m+\rho} \wedge t \wedge \zeta_{\tilde{m}}} \mathbf{1}_{\{\omega(s) \in G^b\}}\, ds\Big]\\
&\qquad \leq \frac{12K_{\tilde{m}}}{(C(V, m, \tilde{m}))^2}\frac{\sqrt{V}\rho\delta}{\alpha},
\end{aligned} \tag{A.45}
$$

where we used (A.43) for the last inequality.

Substituting the above in (A.44) and then using that in (A.42), we obtain for all δ and $z \in \Theta_m$,

$$
\begin{aligned}
P^\delta_z(\zeta_{\tilde{m}} \leq t) &\leq \beta + \delta\tilde{C}(V, m, \tilde{m}, \rho)\sum_{k=1}^\infty \mathbb{P}\big(\zeta^{\delta,k-1}_{m+2\rho} \leq t \wedge \zeta^\delta_{\tilde{m}}\big)\\
&\leq \beta + \delta\tilde{C}(V, m, \tilde{m}, \rho)\sum_{k=1}^\infty \mathbb{P}\big(\tau^{\delta,k-1}_{m+\rho} \leq t \wedge \zeta^\delta_{\tilde{m}}\big),
\end{aligned} \tag{A.46}
$$

where $\tilde{C}(V, m, \tilde{m}, \rho) = \frac{12K_{\tilde{m}}\sqrt{V}\rho}{\alpha(C(V,m,\tilde{m}))^2}$ and we have used the fact that $\zeta^{\delta,k-1}_{m+2\rho} \geq \tau^{\delta,k-1}_{m+\rho}$.

We now proceed to estimate the sum in (A.46). We first note that

$$
\sum_{k=1}^\infty \mathbb{E}\big[\mathbb{E}\big[\big(\zeta^{\delta,k-1}_{m+2\rho} - \tau^{\delta,k-1}_{m+\rho}\big) \wedge t | \mathcal{F}_{\tau^{\delta,k-1}_{m+\rho}}\big]\mathbf{1}_{\{\tau^{\delta,k-1}_{m+\rho} \leq t \wedge \zeta^\delta_{\tilde{m}}\}}\big] \leq 2t. \tag{A.47}
$$

On $\{\tau^{\delta,k-1}_{m+\rho} \le t \wedge \zeta^{\delta}_{\tilde{m}}\}$, we have $Z^\delta(\tau^{\delta,k-1}_{m+\rho}) \in \Theta_m$ and by the strong Markov property of Z^δ,

$$\mathbb{E}\big[(\zeta^{\delta,k-1}_{m+2\rho} - \tau^{\delta,k-1}_{m+\rho}) \wedge t | \mathcal{F}_{\tau^{\delta,k-1}_{m+\rho}}\big] \ge \inf_{x\in\Theta_m} E^{P^\delta_x}[\zeta_{m+2\rho} \wedge t]. \tag{A.48}$$

Let f be a twice continuously differentiable, nondecreasing function on the real line that is an approximation to the function $y \to (y-m)^+$. In particular, we choose f such that $f(y) = 0$ for $y \le m$ and $f(y) = (y-m)^+$ for $y \ge m+\rho$, f' is nondecreasing, $0 \le f''(y) \le 2$ for $m \le y \le m+\rho$. [Note that $f''(y) = 0$ for $y \le m$ and $y \ge m+\rho$.] Then by applying Itô's formula (see [15], page 57) to f and $\langle u, Z^\delta\rangle$ [cf. (A.13)], with $Z^\delta(0) \in \Theta_m$, for any stopping time $\eta \le t \wedge \zeta^\delta_{\tilde{m}}$, we have

$$\begin{aligned}
\mathcal{U}^\delta(\eta) &\triangleq f(\langle u, Z^\delta(\eta)\rangle) \\
&= f(\langle u, Z^\delta(0)\rangle) + \int_0^\eta f'(\langle u, Z^\delta(s)\rangle)\langle u, \mu(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\,ds \\
&\quad + \frac{1}{\delta\sqrt{V}} \int_0^\eta f'(\langle u, Z^\delta(s)\rangle)\langle u, \tilde{\mu}^\delta(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^b\}}\,ds \\
&\quad + \frac{1}{\sqrt{V}} \int_0^\eta f'(\langle u, Z^\delta(s)\rangle)\langle u, \sigma(Z^\delta(s))\rangle \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\,dW(s) \\
&\quad + \frac{1}{\sqrt{V}} \int_0^\eta f'(\langle u, Z^\delta(s-)\rangle)\,d\langle u, \hat{Y}^\delta(s)\rangle \\
&\quad + \frac{1}{2V} \int_0^\eta f''(\langle u, Z^\delta(s)\rangle) u'\Gamma(Z^\delta(s)) u \mathbf{1}_{\{Z^\delta(s)\in G^\circ\}}\,ds \\
&\quad + \sum_{0<s\le\eta} \big(f(\langle u, Z^\delta(s)\rangle) - f(\langle u, Z^\delta(s-)\rangle) \\
&\quad - f'(\langle u, Z^\delta(s-)\rangle)\langle u, \Delta Z^\delta(s)\rangle\big),
\end{aligned} \tag{A.49}$$

where $\Delta Z^\delta(s) = Z^\delta(s) - Z^\delta(s-)$ denotes the jump of Z^δ at s (if there is one). By using Taylor's theorem with remainder the last term above equals

$$\begin{aligned}
&\frac{1}{2}\sum_{0<s\le\eta} f''(\langle u, \tilde{Z}^\delta(s-)\rangle)(\langle u, \Delta Z^\delta(s)\rangle)^2 \\
&\qquad = \frac{\delta^2}{2V} \sum_{0<s\le\eta} f''(\langle u, \tilde{Z}^\delta(s-)\rangle) \sum_{k=1}^r \langle u, v_k\rangle^2 \Delta\tilde{U}^{\delta,b}_k(s),
\end{aligned} \tag{A.50}$$

where $\tilde{Z}^\delta(s)$ is a point on the line segment between $Z^\delta(s-)$ and $Z^\delta(s)$. Substituting this in (A.49), and using the facts that $0 \le f' \le 1$ everywhere, $f' = 0$ and $f'' = 0$ on $(-\infty, m]$, $f'' = 0$ on $[m+\rho, \infty)$, $0 \le f'' \le 2$, and $\langle u, \mu(x)\rangle \le -\alpha$,

$\langle u, \tilde{\mu}^\delta(x)\rangle \le -\alpha$ whenever $\langle u, x\rangle \ge m$, we obtain

$$
\begin{aligned}
\mathcal{U}^\delta(\eta) \le{}& \mathcal{U}^\delta(0) - \alpha \int_0^\eta f'(\langle u, Z^\delta(s)\rangle) \mathbf{1}_{\{\langle u, Z^\delta(s)\rangle > m, Z^\delta(s) \in G^\circ\}}\, ds \\
& - \frac{\alpha}{\delta\sqrt{V}} \int_0^\eta f'(\langle u, Z^\delta(s)\rangle) \mathbf{1}_{\{\langle u, Z^\delta(s)\rangle > m, Z^\delta(s) \in G^b\}}\, ds \\
& + \frac{1}{\sqrt{V}} \int_0^\eta f'(\langle u, Z^\delta(s)\rangle)\langle u, \sigma(Z^\delta(s))\rangle \\
& \times \mathbf{1}_{\{\langle u, Z^\delta(s)\rangle > m, Z^\delta(s) \in G^\circ\}}\, dW(s) \\
& + \frac{1}{\sqrt{V}} \int_0^\eta f'(\langle u, Z^\delta(s-)\rangle)\, d\langle u, \hat{Y}^\delta(s)\rangle \\
& + \frac{1}{V} \int_0^\eta u'\Gamma(Z^\delta(s)) u \mathbf{1}_{\{m < \langle u, Z^\delta(s)\rangle < m+\rho, Z^\delta(s) \in G^\circ\}}\, ds \\
& + \frac{\delta^2}{V} \sum_{k=1}^r \langle u, v_k\rangle^2 \tilde{U}_k^{\delta,b}(\eta).
\end{aligned} \tag{A.51}
$$

Now, with $Z^\delta(0) \in \Theta_m$, for $a \in (0,t)$, since f is nondecreasing and $f(y) = (y-m)^+$ for $y \ge m + \rho$, we have

$$
\begin{aligned}
\mathbb{P}(\zeta^\delta_{m+2\rho} \ge a) &\ge 1 - \mathbb{P}\Big(\sup_{0\le s\le a} \big(\mathcal{U}^\delta(s \wedge \zeta^\delta_{\tilde{m}}) - \mathcal{U}^\delta(0)\big) \ge \rho\Big) \\
&\ge 1 - \mathbb{P}\Big(\max_{0\le\eta\le a\wedge\zeta^\delta_{\tilde{m}}} \mathcal{N}^{\delta,i}(\eta) \ge \rho/4 \text{ for some } i \in \{1,2,3,4\}\Big),
\end{aligned}
$$

where we have used (A.51) for the second inequality and

$$
\begin{aligned}
\mathcal{N}^{\delta,1}(\eta) &= \frac{1}{\sqrt{V}} \int_0^\eta f'(\langle u, Z^\delta(s)\rangle)\langle u, \sigma(Z^\delta(s))\rangle \mathbf{1}_{\{\langle u, Z^\delta(s)\rangle > m, Z^\delta(s) \in G^\circ\}}\, dW(s), \\
\mathcal{N}^{\delta,2}(\eta) &= \frac{1}{\sqrt{V}} \int_0^\eta f'(\langle u, Z^\delta(s-)\rangle)\, d\langle u, \hat{Y}^\delta(s)\rangle, \\
\mathcal{N}^{\delta,3}(\eta) &= \frac{1}{V} \int_0^\eta u'\Gamma(Z^\delta(s)) u \mathbf{1}_{\{m < \langle u, Z^\delta(s)\rangle < m+\rho, Z^\delta(s) \in G^\circ\}}\, ds, \\
\mathcal{N}^{\delta,4}(\eta) &= \frac{\delta^2}{V} \sum_{k=1}^r \langle u, v_k\rangle^2 \tilde{U}_k^{\delta,b}(\eta).
\end{aligned}
$$

The quadratic variation for $\sqrt{V}\mathcal{N}^{\delta,1}$ satisfies

$$
\begin{aligned}
\langle \sqrt{V}\mathcal{N}^{\delta,1}\rangle(\eta) &= \int_0^\eta \big(f'(\langle u, Z^\delta(s)\rangle)\big)^2 u'\Gamma(Z^\delta(s)) u \mathbf{1}_{\{\langle u, Z^\delta(s)\rangle > m, Z^\delta(s) \in G^\circ\}}\, ds \\
&\le K_{\tilde{m}}\eta,
\end{aligned}
$$

whenever $\eta \le \zeta^\delta_{\tilde m}$. It follows in a similar manner to (A.41) that

$$\mathbb{P}\Big(\max_{0\le s\le a\wedge\zeta^\delta_{\tilde m}} \mathcal{N}^{\delta,1}(s) \ge \rho/4\Big) \le \mathbb{P}\big(K_{\tilde m} a \ge \Upsilon(\sqrt{V}\rho/4)\big),$$

where $\Upsilon(\sqrt{V}\rho/4)$ is the first time a one-dimensional Brownian motion reaches the level $\sqrt{V}\rho/4$. The right-hand side above tends to zero as $a \to 0$ and so there is $a_1 \in (0,t)$ (not depending on δ and valid for all initial values of Z^δ in Θ_m) such that the probability on the right above is less than $1/8$ for all $0 < a \le a_1$. By Doob's L^2 maximal inequality, we have in a similar manner to that in (A.45)

$$\begin{aligned}\mathbb{P}\Big(\max_{0\le s\le a\wedge\zeta^\delta_{\tilde m}} |\mathcal{N}^{\delta,2}(s)| \ge \rho/4\Big) &\le \frac{64}{V\rho^2}\mathbb{E}\Big[\int_0^{a\wedge\zeta^\delta_{\tilde m}} u'\tilde\Gamma^\delta\big(Z^\delta(s)\big)u\mathbf{1}_{\{Z^\delta(s)\in G^b\}}\,ds\Big]\\ &\le \frac{64K_{\tilde m}a}{V\rho^2}.\end{aligned}$$

Let $a_2 \in (0,a_1)$ such that the last expression is less than $1/8$ for all $a \le a_2$. For $\mathcal{N}^{\delta,3}$, we will have $\max_{0\le s\le a\wedge\zeta^\delta_{\tilde m}} \mathcal{N}^{\delta,3}(s) < \rho/4$ whenever $a \le a_3 = \rho V/(8K_{\tilde m})$. Since $\mathcal{N}^{\delta,4}$ is an increasing process, we have by Markov's inequality

$$\begin{aligned}&\mathbb{P}\Big(\max_{0\le s\le a\wedge\zeta^\delta_{\tilde m}} \mathcal{N}^{\delta,4}(s) \ge \rho/4\Big)\\ &\quad\le \frac{4\delta^2}{\rho V}\sum_{k=1}^r \langle u, v_k\rangle^2\mathbb{E}\big[\tilde U_k^{\delta,b}\big(a\wedge\zeta^\delta_{\tilde m}\big)\big]\\ &\quad\le \frac{4\delta^2}{\rho V}\sum_{k=1}^r \langle u, v_k\rangle^2\mathbb{E}\Big[\delta^{-2}\int_0^{a\wedge\zeta^\delta_{\tilde m}} \tilde\lambda_k^\delta\big(Z^\delta(s)\big)\mathbf{1}_{\{Z^\delta(s)\in G^b\}}\,ds\Big]\\ &\quad\le \frac{4C_{\tilde m}a}{\rho V},\end{aligned}$$

where $C_{\tilde m} = \sum_{k=1}^r \langle u, v_k\rangle^2 \sup\{\lambda_k(x) : x \in G, \langle u, x\rangle \le \tilde m\}$ and we have used the fact that $\tilde\lambda_k^\delta(x) \le \lambda_k(x)$ for $x \in G$. Let $a_4 \in (0, a_2 \wedge a_3)$ such that the last term above is less than $1/8$ for all $a \le a_4$.

Combining all of the above, it follows that for all $a \in (0, a_4]$,

$$\mathbb{P}\big(\zeta^\delta_{m+2\rho} \ge a\big) \ge 1 - \frac{3}{8} \ge 1/2.$$

This is true for all starting points of Z^δ in Θ_m, and so it follows that since $a_4 < t$,

$$\inf_{x\in\Theta_m} E^{P_x^\delta}[\zeta_{m+2\rho}\wedge t] \ge \frac{a_4}{2}. \tag{A.52}$$

Combining this with (A.48) and (A.47), we obtain

$$\sum_{k=1}^\infty \mathbb{E}[\mathbf{1}_{\{\tau^{\delta,k-1}_{m+\rho}\le t\wedge\zeta^\delta_{\tilde m}\}}] \le \frac{4t}{a_4} \tag{A.53}$$

and substituting this in (A.46), we obtain

$$P_z^\delta(\zeta_{\tilde{m}} \le t) \le \beta + \frac{4\delta t \tilde{C}(V, m, \tilde{m}, \rho)}{a_4}. \tag{A.54}$$

It follows that there is $\delta_0 \in (0, 1/\sqrt{V})$ such that the above expression on the right is strictly less than $(1-\beta)/2$, for all $z \in \Theta_m$ and all $0 < \delta \le \delta_0$. This completes the proof of (A.26). □

PROOF OF LEMMA 6.1. Consider λ_k, $k = 1, \dots, r$, μ, and Γ to be defined on $\mathbb{R}^d$ by (3.7), (4.7) and (4.8), respectively. We begin with the definition of Γ_M. For $\epsilon \in (0, 1)$, define $G_M^\epsilon \triangleq \{x \in \mathbb{R}^d : \operatorname{dist}(x, G_M) \le \epsilon\}$. Notice that for $x \in G_M^\epsilon$, we have $-\epsilon \le x_i \le M/u_i + \epsilon$ for each $1 \le i \le d$. Since $u \ge \mathbb{1}$ and $\epsilon < 1$, we have that $M/u_i + \epsilon \le M + 1$ for $1 \le i \le d$. Then, for $x \in G_M^\epsilon$,

$$\begin{aligned}\lambda_k(x) &= c_k \prod_{i=1}^d x_i^{v_{ik}^-} = c_k \Biggl(\prod_{i:x_i<0} x_i^{v_{ik}^-}\Biggr)\Biggl(\prod_{i:x_i\ge 0} x_i^{v_{ik}^-}\Biggr) \\ &\ge c_k(-\epsilon)\Biggl(\prod_{i:x_i\ge 0} (M+1)^{v_{ik}^-}\Biggr) \ge -c_k\epsilon(M+1)^{[v_k^-]},\end{aligned} \tag{A.55}$$

where $[v_k^-] \triangleq \sum_{i=1}^d v_{ik}^-$. Let $\theta \in \mathbb{R}^d$, $x \in G_M^\epsilon$, and denote by $\mathcal{I}$ the index set associated with the external input reactions. Then

$$\langle \theta, \Gamma(x)\theta\rangle = \sum_{k=1}^r \lambda_k(x)\langle\theta, v_k\rangle^2 = \sum_{k\in\mathcal{I}} c_k\langle\theta, v_k\rangle^2 + \sum_{k\notin\mathcal{I}} \lambda_k(x)\langle\theta, v_k\rangle^2. \tag{A.56}$$

Let $\{k_i^\uparrow, 1 \le i \le d\}$ denote the indices for the input reactions of Assumption 3.2(a). Then, by (A.55) and (A.56),

$$\langle \theta, \Gamma(x)\theta\rangle \ge \sum_{i=1}^d c_{k_i^\uparrow}\theta_i^2 - \sum_{k\notin\mathcal{I}} c_k\epsilon(M+1)^{[v_k^-]}|\theta|^2|v_k|^2 \ge |\theta|^2(b_0 - \epsilon b_1),$$

where $b_0 \triangleq \min_{i=1}^d c_{k_i^\uparrow}$ and $b_1 \triangleq \sum_{k\notin\mathcal{I}} c_k(M+1)^{[v_k^-]}|v_k|^2$. Therefore, Γ is uniformly elliptic on G_M^ϵ for $\epsilon < b_0/b_1$. Fix an $\epsilon \in (0, b_0/b_1)$ and let $\varphi_\epsilon \in C_c^\infty(\mathbb{R}^d)$ be such that $0 \le \varphi_\epsilon(x) \le 1$ for all $x \in \mathbb{R}^d$, $\varphi_\epsilon = 1$ on G_M and $\varphi_\epsilon = 0$ outside of G_M^ϵ. Then the function $\Gamma_M : \mathbb{R}^d \to \mathbb{S}^{d\times d}$ defined as

$$\Gamma_M(x) \triangleq \varphi_\epsilon(x)\Gamma(x) + \bigl(1 - \varphi_\epsilon(x)\bigr)I \tag{A.57}$$

is uniformly elliptic on $\mathbb{R}^d$. In fact, for $x, \theta \in \mathbb{R}^d$,

$$\langle\theta, \Gamma_M(x)\theta\rangle \ge |\theta|^2\bigl((b_0 - \epsilon b_1) \wedge 1\bigr).$$

Now we show that the symmetric positive definite square root σ_M of Γ_M is uniformly Lipschitz continuous on $\mathbb{R}^d$. First, note that Γ is continuously differentiable on $\mathbb{R}^d$, and it, together with its first partial derivatives, is uniformly bounded on the compact set G_M^ϵ. It follows from (A.57) that Γ_M is continuously differentiable on $\mathbb{R}^d$ and its first partial derivatives are bounded on $\mathbb{R}^d$. Hence Γ_M is uniformly Lipschitz continuous on $\mathbb{R}^d$. Then, by Theorem 5.2.2 of [35], we have that σ_M is uniformly Lipschitz continuous on $\mathbb{R}^d$.

Similarly, define $\mu_M : \mathbb{R}^d \to \mathbb{R}^d$ by $\mu_M(x) \triangleq \varphi_\epsilon(x)\mu(x)$. Since μ is continuously differentiable on $\mathbb{R}^d$ and μ_M has compact support, it follows that μ_M is uniformly Lipschitz continuous on $\mathbb{R}^d$.

We now define γ_M. Lemma 4.1, together with the fact that $n_{d+1} = -u/|u|$, implies that there is a constant $\alpha_M > 0$ such that

$$\langle \mu(x), n_j \rangle \ge \alpha_M \qquad \text{for all } j \in I_M(x), x \in G_M^b. \tag{A.58}$$

For $\eta \in (0, 1)$, let

$$G_M^{b,\eta} \triangleq \{x \in \mathbb{R}^d : \operatorname{dist}(G_M^b, x) \le \eta\}. \tag{A.59}$$

We have that $|\mu(x)| \ge \alpha_M$ for all $x \in G_M^b$ by (A.58). Then by the continuity of μ, there is an $\eta \in (0, 1)$ such that

$$|\mu(x)| \ge \frac{\alpha_M}{2} \qquad \text{for all } x \in G_M^{b,\eta}. \tag{A.60}$$

Let $\psi_\eta \in C_c^\infty(\mathbb{R}^d)$ be such that $0 \le \psi_\eta(x) \le 1$, $\psi_\eta(x) = 1$ for $x \in G_M^{b,\frac{\eta}{2}}$ and $\psi_\eta(x) = 0$ outside of $G_M^{b,\eta}$. Define $\gamma_M : \mathbb{R}^d \to \mathbb{R}^d$ as follows:

$$\gamma_M(x) \triangleq \psi_\eta(x)\frac{\mu(x)}{|\mu(x)|} + (1 - \psi_\eta(x))e_1. \tag{A.61}$$

Then since μ is twice continuously differentiable on $\mathbb{R}^d$ and we have (A.60), γ_M is twice continuously differentiable on $\mathbb{R}^d$, and γ_M agrees with $\gamma = \frac{\mu}{|\mu|}$ on $\tilde{G}_M^b$. Moreover, $\gamma_M = \frac{\mu}{|\mu|}$ on $G_M^{b,\frac{\eta}{2}}$ and so is of unit length in a neighborhood of G_M^b. Also, being equal to a constant vector outside of $G_M^{b,\eta}$, γ_M is a constant vector outside of a compact set.

It remains to show that there is a $\beta \in (0, 1)$ such that

$$\bigcup_{0 \le t \le \beta} \mathbb{B}_{t\beta}(x - t\gamma_M(x)) \subset (G_M^\circ)^c \qquad \text{for } x \in G_M^b.$$

Let

$$\beta \in \left(0, \frac{\alpha_M}{\sup_{z \in G_M^b} |\mu(z)|} \wedge 1\right).$$

Suppose $x \in G_M^b$, $t \in [0, \beta]$ and $y \in \mathbb{B}_{t\beta}(x - t\gamma_M(x))$. To prove that $y \in (G_M^\circ)^c$, since $G_M^\circ = \{z \in \mathbb{R}^d : \langle z, n_j \rangle > 0,\ \text{for } j = 1, \dots, d, \text{ and } \langle z, n_{d+1} \rangle > -M/|u|\}$, it suffices to show that

$$\langle n_j, y - x \rangle \le 0 \qquad \text{for some } j \in I_M(x).$$

Indeed, since $y \in \mathbb{B}_{t\beta}(x - t\gamma_M(x))$, there is $d_y \in \mathbb{R}^d$ such that $|d_y| \leq t\beta$ and

$$y = x - t\gamma_M(x) + d_y,$$

and then for $j \in I_M(x)$,

$$\begin{aligned}\langle n_j, y - x\rangle &= \langle n_j, -t\gamma_M(x) + d_y\rangle \\ &= -t\langle n_j, \gamma_M(x)\rangle + \langle n_j, d_y\rangle \\ &\leq -t\langle n_j, \gamma_M(x)\rangle + |n_j||d_y| \\ &\leq \frac{-t\alpha_M}{\sup_{z \in G_M^b} |\mu(z)|} + t\beta \leq 0,\end{aligned}$$

by (A.58), the definition of γ_M and choice of β. Thus, $y \in (G_M^\circ)^c$. $\square$

PROOF OF LEMMA 7.1. Let $x \in G^b$, we have the following by the triangle inequality and Lemma 4.1(i):

$$|\tilde{\gamma}^n(x) - \gamma(x)| \leq \frac{2}{\alpha}|\tilde{\mu}^n(x) - \mu(x)| \leq \frac{2}{\alpha}\sum_{k=1}^r |v_k||\tilde{\lambda}_k^n(x) - \lambda_k(x)|.$$

Therefore, it is enough to show that the statement holds with $\tilde{\lambda}_k^n$ and λ_k replacing $\tilde{\gamma}^n$ and γ, respectively, for $1 \leq k \leq r$.

Let $M > 1$ and $\varepsilon > 0$. Since δ_n decreases monotonically to zero as n increases, there is $n_\varepsilon > 1$ such that

$$\delta_n < \frac{\varepsilon\sqrt{V}}{\max_{k,j} c_k v_{jk}^- M^{[v_k^-]} + 1}, \tag{A.62}$$

for all $n \geq n_\varepsilon$, where $[v_k^-] \triangleq \sum_{i=1}^d v_{ik}^-$. Then, for $x \in G^b$ such that $|x| \leq M$, and $n \geq n_\varepsilon$, let us define

$$\mathcal{N}_n^k(x) = \left\{1 \leq j \leq d : 0 \leq x_j < \frac{v_{jk}^- \delta_n}{\sqrt{V}}\right\},$$

for $1 \leq k \leq r$. Notice that if $\mathcal{N}_n^k(x) = \varnothing$, then $\lambda_k(x) = \tilde{\lambda}_k^n(x)$, which implies that

$$|\tilde{\lambda}_k^n(x) - \lambda_k(x)| < \varepsilon. \tag{A.63}$$

Suppose now that there is a $j \in \mathcal{N}_n^k(x)$. Then we have

$$|\tilde{\lambda}_k^n(x) - \lambda_k(x)| = \left|0 - c_k \prod_{i=1}^d x_i^{v_{ik}^-}\right| = c_k x_j^{v_{jk}^-} \prod_{\substack{i=1 \\ i \neq j}}^d x_i^{v_{ik}^-} \leq c_k \frac{v_{jk}^- \delta_n}{\sqrt{V}} M^{[v_k^-]},$$

where we used the facts that $x_j < v_{jk}^- \delta_n/\sqrt{V}$ and that $|x_i| \leq M$, for $1 \leq i \leq d$. Since $n \geq n_\varepsilon$, δ_n satisfies (A.62), and therefore (A.63) holds. Since the choice of $x \in G^b$ such that $|x| \leq M$, $n \geq n_\varepsilon$, and $1 \leq k \leq r$, was arbitrary, the result follows. $\square$

Acknowledgments. We are grateful to Paul Dupuis for permitting us to include elements of a private communication with him, relating to a modification of the uniqueness argument given in [5]. We would also like to thank Bence Melykuti for helpful feedback, as well as an Associate Editor and two anonymous referees for constructive comments and suggestions, on the first version of this paper.

REFERENCES

[1] ANDERSON, D. F., HIGHAM, D. J., LEITE, S. C. and WILLIAMS, R. J. (2019). On constrained Langevin equations and (bio)chemical reaction networks. *Multiscale Model. Simul.* **17** 1–30. MR3895328

[2] ANDERSON, D. F. and KURTZ, T. G. (2011). Continuous time Markov chain models for chemical reaction networks. In *Design and Analysis of Biomolecular Circuits* (H. Koeppl, G. Setti, M. d. Bernardo and D. Densmore, eds.) 3–42. Springer, New York.

[3] BOSSY, M., GOBET, E. and TALAY, D. (2004). A symmetrized Euler scheme for an efficient approximation of reflected diffusions. *J. Appl. Probab.* **41** 877–889. MR2074829

[4] DUPUIS, P. and ISHII, H. (1990). On oblique derivative problems for fully nonlinear second-order elliptic partial differential equations on nonsmooth domains. *Nonlinear Anal.* **15** 1123–1138. MR1082287

[5] DUPUIS, P. and ISHII, H. (1993). SDEs with oblique reflection on nonsmooth domains. *Ann. Probab.* **21** 554–580. Correction **36**(5) (2008) 1992–1997. MR1207237

[6] ETHIER, S. N. and KURTZ, T. G. (2005). *Markov Processes*: *Characterization and Convergence*, 2nd ed. Wiley, New York. MR0838085

[7] FEINBERG, M. (1979). Lectures on chemical reaction networks. Delivered at the Mathematics Research Center, Univ. Wisconsin-Madison. Available at https://crnt.osu.edu/lectures-chemical-reaction-networks.

[8] GILLESPIE, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22** 403–434. MR0503370

[9] GILLESPIE, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81** 2340–2361.

[10] GILLESPIE, D. T. (1992). A rigorous derivation of the chemical master equation. *Phys. A* **188** 404–425.

[11] GILLESPIE, D. T. (2000). The chemical Langevin equation. *J. Chem. Phys.* **113** 297–306.

[12] GILLESPIE, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115** 1716–1733.

[13] HIGHAM, D. J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* **43** 525–546. MR1872387

[14] HIGHAM, D. J. (2008). Modeling and simulating chemical reactions. *SIAM Rev.* **50** 347–368. MR2403055

[15] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877

[16] KANG, W. and RAMANAN, K. (2017). On the submartingale problem for reflected diffusions in domains with piecewise smooth boundaries. *Ann. Probab.* **45** 404–468. MR3601653

[17] KANG, W. and WILLIAMS, R. J. (2007). An invariance principle for semimartingale reflecting Brownian motions in domains with piecewise smooth boundaries. *Ann. Appl. Probab.* **17** 741–779. MR2308342

[18] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940

[19] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131. MR0375412

[20] KURTZ, T. G. (1972). The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.* **57** 2976–2978.

[21] KURTZ, T. G. (1976). Limit theorems and diffusion approximations for density dependent Markov chains. In *Stochastic Systems: Modeling, Identification and Optimization, I* (R. J.-B. Wets, ed.). *Mathematical Programming Studies* **5** 67–78. Springer, Berlin.

[22] KURTZ, T. G. (1977/78). Strong approximation theorems for density dependent Markov chains. *Stochastic Process. Appl.* **6** 223–240. MR0464414

[23] KURTZ, T. G. (1981). *Approximation of Population Processes. CBMS-NSF Regional Conference Series in Applied Mathematics* **36**. SIAM, Philadelphia, PA. MR0610982

[24] LÉPINGLE, D. (1978). Sur le comportement asymptotique des martingales locales. In *Séminaire de Probabilités, XII* (*Univ. Strasbourg, Strasbourg*, 1976/1977). *Lecture Notes in Math.* **649** 148–161. Springer, Berlin. MR0520004

[25] MANNINEN, T., LINNE, M. and RUOHONEN, K. (2006). Developing Itô stochastic differential equation models for neuronal signal transduction pathways. *Comput. Biol. Chem.* **30** 280–291. DOI:10.1016/j.compbiolchem.2006.04.002.

[26] MARUYAMA, G. (1955). Continuous Markov processes and stochastic equations. *Rend. Circ. Mat. Palermo* (2) **4** 48–90. MR0071666

[27] MCADAMS, H. H. and ARKIN, A. (1997). Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA* **94** 814–819.

[28] MÉLYKÚTI, B. (2010). Theoretical advances in the modelling and interrogation of biochemical reaction systems: Alternative formulations of the chemical Langevin equation and optimal experiment design for model discrimination. Ph.D. thesis, Dept. Statistics, Univ. Oxford. Available at: http://ora.ouls.ox.ac.uk/objects/uuid:d368c04c-b611-41b2-8866-cde16b283b0d.

[29] PLATEN, E. (1999). An introduction to numerical methods for stochastic differential equations. *Acta Numer.* **8** 197–246. MR1819646

[30] R CORE TEAM (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[31] ROGERS, L. C. G. and WILLIAMS, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol.* 2. *Itô Calculus*. Cambridge Univ. Press, Cambridge. Reprint of the second (1994) edition. MR1780932

[32] SCHNOERR, D., SANGUINETTI, G. and GRIMA, R. (2014). The complex chemical Langevin equation. *J. Chem. Phys.* **141** 024103. DOI:10.1063/1.4885345.

[33] SCHWALB, B., TRESCH, A., TORKLER, P., DUEMCKE, S. and DEMEL, C. (2015). LSD: Lots of superior depictions. R package version 3.0.

[34] SCOTT, M. and INGALLS, B. P. (2005). 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, California*.

[35] STROOCK, D. W. and VARADHAN, S. R. S. (2006). *Multidimensional Diffusion Processes*. Springer, Berlin. Reprint of the 1997 edition. MR2190038

[36] SZPRUCH, L. and HIGHAM, D. J. (2009/10). Comparing hitting time behavior of Markov jump processes and their diffusion approximations. *Multiscale Model. Simul.* **8** 605–621. MR2600298

[37] TURNER, T. E., SCHNELL, S. and BURRAGE, K. (2004). Stochastic approaches for modelling in vivo reactions. *Comput. Biol. Chem.* **28** 165–178.

[38] VAN KAMPEN, N. G. (1961). A power series expansion of the master equation. *Can. J. Phys.* **39** 551–567. MR0128921

[39] WALLACE, E. W. J., PETZOLD, L. R., GILLESPIE, D. T. and SANFT, K. R. (2012). Linear noise approximation is valid over limited times for any chemical system that is sufficiently large. *IET Syst. Biol.* **6** 102–115.

[40] WANG, J. G. (1995). The asymptotic behavior of locally square integrable martingales. *Ann. Probab.* **23** 552–585. MR1334161

[41] WILHELM, T. (2009). The smallest chemical reaction system with bistability. *BMC Syst. Biol.* **3** 90.

[42] WILKIE, J. and WONG, Y. M. (2008). Positivity preserving chemical Langevin equations. *Chem. Phys.* **353** 132–138.

[43] WILKINSON, D. J. (2006). *Stochastic Modelling for Systems Biology*. Chapman & Hall/CRC, Boca Raton, FL. MR2222876

CENTRO DE MATEMÁTICA, COMPUTAÇÃO
E COGNIÇÃO
UNIVERSIDADE FEDERAL DO ABC
AV. DOS ESTADOS 5001
SANTO ANDRÉ, SP, 09210-580
BRAZIL
E-MAIL: saul.leite@ufabc.edu.br

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, SAN DIEGO
9500 GILMAN DRIVE
LA JOLLA, CALIFORNIA 92093-0112
USA
E-MAIL: williams@math.ucsd.edu