On the stochastic modelling of surface reactions through reflected chemical Langevin equations

M Pineda[1] and M Stamatakis[2]

*Department of Chemical Engineering, University College London, Roberts Building,
Torrington Place, London WC1E 7JE, United Kingdom*

Abstract

Modelling of small-scale heterogeneous catalytic systems with master equations captures the impact of molecular noise, but can be computationally expensive. On the other hand, the chemical Fokker-Planck approximation offers an excellent alternative from an efficiency perspective. The Langevin equation can generate stochastic realisations of the Fokker-Planck equation; yet, these realisations may violate the conditions $0 \leq \theta \leq 1$ (where θ is surface coverage). In this work, we adopt Skorokhod's formulations to impose reflective boundaries that remedy this issue. We demonstrate the approach on a simple system involving a single species and describing adsorption, desorption, reaction and diffusion processes on a lattice. We compare different numerical schemes for the solution of the resulting reflected Langevin equation and calculate rates of convergence. Our benchmarks should guide the choice of appropriate numerical methods for the accurate and efficient simulation of chemical systems in the catalysis field.

Keywords: `elsarticle.cls`, LATEX, Elsevier, template

2010 MSC: 00-01, 99-00

Email addresses: [1]`m.pineda@ucl.ac.uk` (M Pineda[1] and M Stamatakis[2]),
[2]`m.stamatakis@ucl.ac.uk` (M Pineda[1] and M Stamatakis[2])

1. Introduction

Traditionally, the linear and non-linear kinetics of chemical reactions on solid surfaces has been neatly analysed using deterministic ordinary or partial differential equations for the average concentrations of surface adsorbates or the coverages [1]. However, it is also well recognised that in many cases the predictions of these deterministic approaches can depart from real situations due to the stochasticity in the occurrence of the elementary processes of reaction and diffusion [2]. This stochasticity is commonly referred to as intrinsic noise and manifests itself by the occurrence of fluctuations in the observables (e.g. coverage) when the sizes of the catalytic surfaces are small. Normally, the relative amplitude of these fluctuations varies as $N^{-1/2}$ with N being the number of adsorption sites on the surface [2].

Small scale chemical reactions on surfaces, where these fluctuations may be large enough to be relevant, are of great practical importance as testified by experiments on supported catalysts [3]. These catalysts consist of many small metal particles of few nanometers diameter composed by reactive facets, typically containing a few hundred to a few thousand surface atoms. In several experimental studies performed with CO oxidation on oxide-supported Pd particles it was shown that the kinetic bistability (the existence of two kinetic regimes for the same reaction conditions) vanishes with decreasing particle size [4]. Field emitter tips (FETs) have been also implemented as model systems to study certain aspects of catalytic reactions on supported catalysts. The surface of a tip consists of small reactive facets of similar sizes as the nanoscale particles of a supported catalyst [5]. Interestingly, using a Pt field emitter tip as catalyst for CO oxidation, it was shown that local coverage fluctuations induce transitions between the two kinetic stationary states that coexist in the so-called bistable range [6, 7]. The role of fluctuations on bistability and oscillations occurring in other relevant chemical reactions on FETs and nanoparticles has also been studied [8, 9, 9, 10, 11, 12]. It is also interesting to mention that in the field of electrochemical reactions it has been theoretically predicted that, in

2

the case of nano-electrodes, the stochastic nature of electrochemical reactions can induce random fluctuations of the electrode potential that result in significantly enhanced reaction kinetics in comparison with macroscopic electrodes [13, 14, 15].

The most popular and successful mathematical formalism used to analyse chemical reactions on nano- and mesoscale surfaces is the so-called stochastic kinetic approach [16, 17]. Within this framework, reversible adsorption, reaction, and diffusion are typically simulated using the on-lattice kinetic Monte Carlo (KMC) method [18, 19, 20, 21, 22]. However, this approach can become extremely computationally intensive when the surface of interest is large or when the diffusion of some of the adsorbates is much faster than other processes on the surfaces [23]. Nevertheless, in the limit of ultrafast diffusion the well-mixed approximation holds and an interpretation in terms of the so-called chemical master equation (CME) for the probability to find a number of adsorbates on the surface is possible [24, 25, 12]. This equivalence between the on-lattice KMC approach and the well-mixed CME has been already established [26]. Under some circumstances the well-mixed CME can be solved directly, or realisations can be obtained by the so-called stochastic simulation algorithm (SSA) by Gillespie [27]. The SSA is a method of generating individual realisations of the stochastic process whose probability distribution reproduces that of the CME. But, it also becomes computationally expensive as the number of adsorbates increases or/and when the model exhibits time scale separation (between fast and slow reactions) [28]. Furthermore, since the system is stochastic, many simulations are needed to make a proper inference concerning global kinetics.

An interesting alternative to incorporate stochasticity in the modelling of catalytic surface reactions at both the nano- and mesoscales is the so-called chemical Langevin equation (CLE) [29, 30] and its associated chemical Fokker-Planck equation (CFPE) [9, 31]. Several attempts to implement the CLE to study chemical reactions on surfaces have been already reported. Phenomena like for example intrinsic noise stochastic resonance and optimal particle size for

3

reaction rate oscillation on nanometer-sized particles have been predicted based on this approach [32, 33].

The CLE is a stochastic differential equation (SDE) that describes the time evolution of the coverage of adsorbates on a well-mixed surface, and constitutes a link between the CME and the macroscopic description of surface reactions. Moreover it has been recognised that this approach leads to substantial savings in computational time when the number of adsorbates is large. However, issues concerning the validity of the CLE have been raised. Given that the solution of the CLE describes the temporal evolution of coverage, it must be positive and less than (or equal to) one to have a physical meaning. Yet, the mathematical formulation of a Langevin equation can lead to negative solutions under certain circumstances. For instance, for a catalytic system this can happen when the noise term is finite, due to e.g an adsorption process, at coverage close to zero. Furthermore, since the stochastic term of the CLE equation contains the square root of some combination of coverages, the numerical solutions can then become imaginary [17, 34, 35].

Several alternatives have been proposed to force the solution to remain non negative. It is common to set the numerical solution to zero when it becomes negative or to impose in a simplistic manner reflected boundary conditions on traditional numerical schemes. In this work we adopted the later approach when numerically solving CLEs using the traditional Euler-Maruyama (EM) scheme, and call it the mirror-boundary EM scheme. However, we will show that these reflections do not necessarily lead to results in accordance with the solutions of the corresponding CFPEs. Another method is to modify the drift and noise terms of the CLE in order to ensure positivity [36]. However, this approach has been found to be inaccurate and even physically inconsistent [37]. A more sophisticated approach is to extend the domain of the CLE to the complex space [34]. Although it has been found that the so-called complex CLE accurately predicts real values for the mean concentrations/coverages and autocorrelation functions, the molecular numbers are normally complex and therefore their phys-

4

ical meaning is questionable. In addition, it is not clear how the complex CLE would impose the restriction a surface coverage cannot exceed 100%. A different approach is to incorporate, in a mathematically rigorous way, boundaries into the CLE [38, 39]. All the aforementioned approaches have attracted a lot of focus in the biomathematics/computational biology communities. However, they have not yet been adopted in the catalysis community. We believe that formulating catalytic kinetics models in terms of CLEs which naturally preserve the positivity of the solution, and additionally enable us to impose a maximum of 100% for surface coverages, would be of interest for the efficient simulation of catalysts at the nano- and mesoscales.

The purpose of this paper is to address the problem of ensuring realistic solutions to the CLEs of catalytic systems occurring on metal surfaces. We do it by using a reflected SDE formalism similar to the one implemented in other research fields [40, 41, 42, 43, 39]. As a illustrative example we consider the dynamics of adsorption, desorption, reactions, and diffusion on a lattice, involving a single species [26]. After discussing the shortcomings of the traditional stochastic approaches, we proceed to present the reflected CLE (RCLE) of the system. We propose it as the SDE that correctly produces the individual realisations of the stochastic process described by the chemical Fokker-Planck equation (CFPE), which is obtained after the truncation of Kramers-Moyal expansion of the CME, and solved after imposing reflected boundary conditions. We numerically integrate this RCLE with an Euler-type numerical scheme introduced by Lépingle [44], and analyse the order of convergence of the scheme as a function of the strength/magnitude of the intrinsic noise [45, 46, 47]. We continue by comparing the numerical solutions of the RCLE with the solutions of the CFPE and CME. Finally, we explore the computational savings offered by the RCLE formalism.

The paper is organised as follows. In Sec. 2.1 we introduce the illustrative model and its corresponding CME. In the same section we derive the CFPE and present the associated CLE. Then, we discuss the shortcomings of implementing
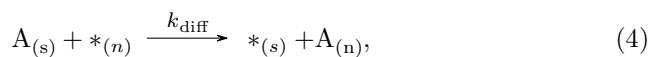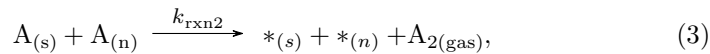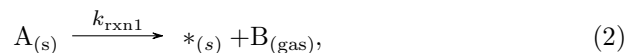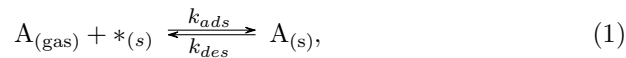
5

this CLE. In Sec 2.2 we introduce the RCLE and the numerical scheme used to integrate it. Subsequently, in Sec 3, we compute the order of convergence of the scheme, compare the solutions of the RCLE with the solution of the CFPE and CME, and analyse the computational efficiency of the RCLE. In Sec. 4, we summarise and mention a number of possible applications and extensions of the methodology presented in this work.

2. Theory

This section starts with the introduction of the catalytic system that serves as a benchmark in our studies, followed by the pertinent models. Under the assumption of ultrafast diffusion we derive a well-mixed CME, and under the additional assumption of mesoscopic system size, we formulate a CLE. The boundary conditions imposed express the physical requirements that the (fractional) surface coverage is bounded between 0 and 1. This naturally gives rise to the reflected CLE, which can be solved numerically with the schemes we discuss. The results of these schemes will then be presented in the next section.

2.1. Model formulation and traditional stochastic analysis

To illustrate the results of this work, we consider the following dynamics of adsorption, desorption, reactions, and diffusion on a lattice, involving a single species [26]:

$$A_{(gas)} + *_{(s)} \xrightleftharpoons[k_{des}]{k_{ads}} A_{(s)}, \tag{1}$$

$$A_{(s)} \xrightarrow{k_{rxn1}} *_{(s)} + B_{(gas)}, \tag{2}$$

$$A_{(s)} + A_{(n)} \xrightarrow{k_{rxn2}} *_{(s)} + *_{(n)} + A_{2(gas)}, \tag{3}$$

$$A_{(s)} + *_{(n)} \xrightarrow{k_{diff}} *_{(s)} + A_{(n)}, \tag{4}$$

6

where $*_{(s)}$ denotes a free site and $*_{(n)}$ its neighbour. The lattice consists of N free sites and coordination number ζ. In the first elementary step, $A_{(gas)}$ represents a chemical species in the gas phase (or in general in a bulk phase) that can occupy a free site to give rise to the adsorbed state $A_{(s)}$. The adsorbed state $A_{(s)}$ can subsequently desorb as $A_{(gas)}$. The second elementary step represents a species in the adsorbed state $A_{(s)}$ that can isomerise into $B_{(gas)}$ and rapidly desorb (1-site, 1st-order reaction). The third elementary step describes a situation in which an adsorbed species $A_{(s)}$ reacts with another species $A_{(n)}$ bound to a neighbouring site, and rapidly desorb as $A_{2(gas)}$ dimer (2-site, 2nd-order reaction or dimerisation reaction). Finally, the last step represents the typical case of diffusion in which a chemical specie in the adsorbate state $A_{(s)}$ can jump to a neighbour free site $*_{(n)}$ to form the adsorbed state $A_{(n)}$. Parameters k_{ads}, k_{des}, k_{rxn1}, and k_{rxn2} represent the rate constants for adsorption, desorption, 1-site reaction, and 2-site reaction, respectively. The parameter k_{diff} is the rate constant for diffusion. In this work no lateral interactions between chemical species are being taken into account. This kinetic scheme (a single species on a lattice) was already implemented in [26] to illustrate the equivalence of on-lattice stochastic chemical kinetics with the well-mixed chemical master equations in the limit of fast diffusion. This reaction network is generic, not necessarily capturing a known mechanism; however, its elementary steps appear in several interesting catalytic reactions, e.g. steps 1 and 3 would be relevant in electrochemical H evolution [48].

In the following subsections, we present the traditional mean-field stochastic description of this model and discuss its shortcomings.

2.1.1. The chemical master equation approach

Let us assume from now on that the rate constant for diffusion is much larger than any other rate constants of the model (the assumption of fast diffusion is usually valid for catalytic reactions on surfaces [23]). In this limit of very fast diffusion the well-mixed assumption holds and an interpretation in terms of the

7

Table 1: Processes, population changes, and transition rates for our well-mixed CME treatment of the dynamics of N_A for a lattices with N available sites. Parameters k_{ads} and k_{des} represent the rate constants for adsorption and desorption, respectively. k_{diff} is the rate constant for diffusion. Parameters k_{rxn1} and k_{rxn2} are the rate constants for the 1-site and 2-site reactions, respectively. ζ is the coordination number of the lattice or the number of nearest neighbours of a site.

Process	Population change	Transition rate
Adsorption	$N_A \to N_A + 1$	$W_{ads} = k_{ads}(N - N_A)$
Desorption	$N_A \to N_A - 1$	$W_{des} = k_{des}N_A$
1st-order reaction	$N_A \to N_A - 1$	$W_{rxn1} = k_{rxn1}N_A$
2nd-order reaction	$N_A \to N_A - 2$	$W_{rxn2} = \frac{\zeta k_{rxn2}}{2(N-1)}N_A(N_A - 1)$

CME already derived by Stamatakis and Vlachos is possible [26]. This CME was obtained after applying singular perturbation analysis to a generic on-lattice master equation for the four elementary steps described above (see Eqs. 1-4). This equation describes the temporal evolution of the probability $P(N_A, t)$ that the lattice or surface has N_A adsorbates at time t. Thus, this equation is a well-mixed stochastic mean-field approach in which one only tracks the total number of N_A adsorbates on the surface and their decrement and increment. This number changes stochastically due to the adsorption, desorption, and reactions in the manner according to the corresponding transition rates indicated in Table 1. These transition rates are evident for adsorption, desorption, and 1st-order reaction. Only the transition rate for 2nd-order reaction involves a nontrivial expression that contains the coordination number ζ and the total number of sites on the surface (reference [26] provides a formal derivation of these transition rates from the on-lattice master equation description). As expected, on a totally covered periodic lattice, one obtains $W_{rxn2} = k_{rxn2}\frac{\zeta N}{2}$, where $\frac{\zeta N}{2}$ is the number of occupied pairs. Note also that in a 2D regular lattice ζ can take one of the following values: 3 for a honeycomb-type lattice, 4 for a square lattice, and 6

for a hexagonal lattice [26]. It is now possible to write down the CME as

$$\frac{\partial P(N_A;t)}{\partial t} = W_{ads}(N_A - 1)P(N_A - 1;t) - W_{ads}(N_A)P(N_A;t)$$
$$+ W_{des}(N_A + 1)P(N_A + 1;t) - W_{des}(N_A)P(N_A;t)$$
$$+ W_{rxn1}(N_A + 1)P(N_A + 1;t) - W_{rxn1}(N_A)P(N_A;t) \tag{5}$$
$$+ W_{rxn2}(N_A + 2)P(N_A + 2;t) - W_{rxn2}(N_A)P(N_A;t),$$

where $0 \leq N_A \leq N$. Because N_A cannot be negative or larger that N, $P(N_A, t)$ must satisfy the following equations at the boundaries (reflective boundaries):

$$\frac{\partial P(0;t)}{\partial t} = -k_{ads}NP(0;t) + (k_{des} + k_{rxn1})P(1;t) + \frac{\zeta k_{rxn2}}{(N-1)}P(2;t), \tag{6}$$

$$\frac{\partial P(N;t)}{\partial t} = k_{ads}P(N - 1;t) - (k_{des} + k_{rxn1})NP(N;t) - \frac{\zeta N k_{rxn2}}{2}P(N;t), \tag{7}$$

where we have evaluated the corresponding transition rates and considered that $P(-1;t) = P(N + 1;t) = P(N + 2;t) = 0$. One can obtain the stationary probability distribution $P_{st}(N_A;t)$ by direct solution of a system of $N + 1$ algebraic equations by writing Eq. 5 at steady state for $N_A = 1, \cdots, N - 1$, along with $\frac{\partial P(0;t)}{\partial t} = \frac{\partial P(N;t)}{\partial t} = 0$. One could also opt to numerically solve Eq. 5 or generate stochastic realisations of this equation by implementing the SSA [26, 28]. However, as the number of molecules increases, solving or simulating the CME becomes computationally expensive. This restricts the implementation of the CME when modelling more complex multi-scale catalytic systems like for example the set of coupled reactive facets of a field emitter tip [8].

The so-called stochastic simulation algorithm or SSA is a method that generates individual realisations of the stochastic process whose probability distribution reproduces that of the CME [28]. One simply generates random numbers to determine the next chemical process to occur as well as the time interval. For our system, this algorithm is described below [16]:

Stochastic simulation algorithm. **1)** Set the time $t = 0$ and the initial number of adsorbates, N_A.

210

9

2) Calculate the total transition rate as

$$W_{tot} = \sum_{r=1}^{4} W_r,\tag{8}$$

where the values of W_r are given in table 1.

3) Get two random numbers p_1 and p_2 from the interval $[0,1]$.

4) Determine the time interval for the next process as

$$\delta t = -\frac{ln(p_1)}{W_{tot}}.\tag{9}$$

5) Find the process q to happen. To do this, take q as the smallest integer satisfying

$$\sum_{r=1}^{q} W_r > p_2 W_{tot} \geq \sum_{r=1}^{q-1} W_r.\tag{10}$$

6) Change the N_A to reflect the occurrence of this process according to table 1. Finally, increase the time by δt and repeat steps 2-6 until the final time is reached.

In the following section we discuss advantages and disadvantages of two popular approximations developed to overcome the computationally expensive problems involving the two aforementioned stochastic approaches.

2.1.2. The chemical Fokker-Planck equation approach

The chemical Fokker-Planck equation or CFPE offers a stochastic description at a more coarse grained level than the CME [49, 50, 51]. It involves a continuous stochastic variable; the fraction of adsorbates or coverage, whereas the CME considers the total number of adsorbates on the surface. To obtain the CFPE from the CME, let us consider a large system and define the coverage $y = N_A/N$, with $0 \leq y \leq 1$. Then, after replacing N_A in Eq. 5, one gets

$$\begin{aligned}
\frac{\partial P(y;t)}{\partial t} &= \frac{k_{ads}}{\epsilon}\left[1-(y-\epsilon)\right]P(y-\epsilon;t) + \frac{k_{des}+k_{rxn1}}{\epsilon}(y+\epsilon)P(y+\epsilon;t) \\
&+ \frac{\zeta k_{rxn2}}{2\epsilon(1-\epsilon)}(y+2\epsilon)(y+\epsilon)P(y+2\epsilon;t) \\
&- \left[\frac{k_{ads}}{\epsilon}(1-y) + \frac{k_{des}+k_{rxn1}}{\epsilon}y + \frac{\zeta k_{rxn2}}{2\epsilon(1-\epsilon)}y(y-\epsilon)\right]P(y;t),
\end{aligned}\tag{11}$$

10

where $\epsilon = \frac{1}{N}$. Note that we are introducing a new probability for y, which is like $P(N_A, t)$ but renormalised according to $y = N_A/N$. Now, we proceed to derive the so-called Kramers-Moyal expansion of Eq. 11 [52]. This is obtained by applying the Taylor expansion,

$$f(x-z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} z^n \frac{d^n}{dx^n} f(x), \tag{12}$$

to the first three terms of Eq. 11 and collecting terms of the same order. If we assume that N is large enough, the terms of order greater that two can be neglected. Then, after considering

$$
\begin{aligned}
\left[1 - (y - \epsilon)\right] P(y - \epsilon; t) \approx{}& (1-y)P(y;t) - \epsilon \partial_y \left[(1-y)P(y;t)\right] \\
&+ \frac{\epsilon^2}{2} \partial_y^2 \left[(1-y)P(y;t)\right],
\end{aligned}
\tag{13}
$$

$$(y+\epsilon)P(y+\epsilon;t) \approx yP(y;t) + \epsilon \partial_y \left[yP(y;t)\right] + \frac{\epsilon^2}{2} \partial_y^2 \left[yP(y;t)\right], \tag{14}$$

and

$$
\begin{aligned}
(y+2\epsilon)(y+\epsilon)P(y+2\epsilon;t) \approx{}& y(y-\epsilon)P(y;t) + 2\epsilon \partial_y \left[y(y-\epsilon)P(y;t)\right] \\
&+ \frac{(2\epsilon)^2}{2} \partial_y^2 \left[y(y-\epsilon)P(y;t)\right],
\end{aligned}
\tag{15}
$$

one gets the following CFPE

$$\frac{\partial P(y;t)}{\partial t} = -\frac{\partial}{\partial y}[a(y)P(y;t)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[b(y)P(y;t)], \tag{16}$$

where the drift and diffusion coefficients are given by

$$a(y) = k_{ads}(1-y) - (k_{des} + k_{rxn1})y - \frac{\zeta k_{rxn2}}{(1-\epsilon)}y(y-\epsilon), \tag{17}$$

and

$$b(y) = \epsilon \left[k_{ads}(1-y) + (k_{des} + k_{rxn1})y + \frac{2\zeta k_{rxn2}}{(1-\epsilon)}y(y-\epsilon)\right], \tag{18}$$

respectively. This CFPE describes the evolution of the probability density function (PDF), $P(y;t)$, and can be also written as

$$\frac{\partial P(y;t)}{\partial t} + \frac{\partial J(y,t)}{\partial y} = 0, \tag{19}$$

11

where

$$J(y,t) = a(y)P(y;t) - \frac{1}{2}\frac{\partial}{\partial y}[b(y)P(y;t)], \tag{20}$$

is the so-called probability current. Then, it follows that at the steady state $P(y;t) = P_{st}(y)$ and $J(y) = constant$. However, note that the CFPE approximation (Eq. 16) does not satisfy the boundary conditions of the CME (Eq. 5) at $N_A = 0$ ($y = 0$) and $N_A = N$ ($y = 1$), since the noise can drive the system beyond those boundaries [52, 53] (we will come back to this point when analysing the SDE or chemical Langevin equation corresponding to the CFPE). Therefore, if we wish to confine the probability to the interval $[0, 1]$, appropriate boundary conditions must be imposed [49, 54]. In this work (because y can not leave the interval $[0, 1]$), the appropriate boundary conditions are total reflections; thus the probability current must be zero at the two boundaries. In other words, $J(0) = J(1) = 0$, and

$$a(y)P_{st}(y) - \frac{1}{2}\frac{\partial}{\partial y}[b(y)P_{st}(y)] = 0, \tag{21}$$

at $y = 0$ and $y = 1$ [49].

Note that, if we assume stationarity, the reflective boundary conditions also ensure a zero probability current everywhere ($J(y) = 0$). Thus, Eq. 21 can be integrated to give the well-known stationary solution

$$P_{st}(y) = \frac{C}{b(y)}exp\left\{2\int_0^y dx\frac{a(x)}{b(x)}\right\}, \tag{22}$$

where C is a constant which has to be chosen so that $P_{st}(y)$ is normalised. This probability density function is indeed 0 outside the interval $[0, 1]$ [49].

Before continuing with our analysis it is interesting to note that, when approaching the low boundary at $y = 0$, it may always occur that $y < \epsilon$. Thus, the terms of the drift and diffusion coefficients (Eqs. 17 and 18) corresponding to the 2nd-order reaction can acquire negative values. This is in contrast with the always positive values that the corresponding transition rate of the CME assumes (see W_{rxn2} in Table 1). Therefore, to ensure positivity of this term, in this work we replace the 2nd-order reaction terms of the drift and diffusion
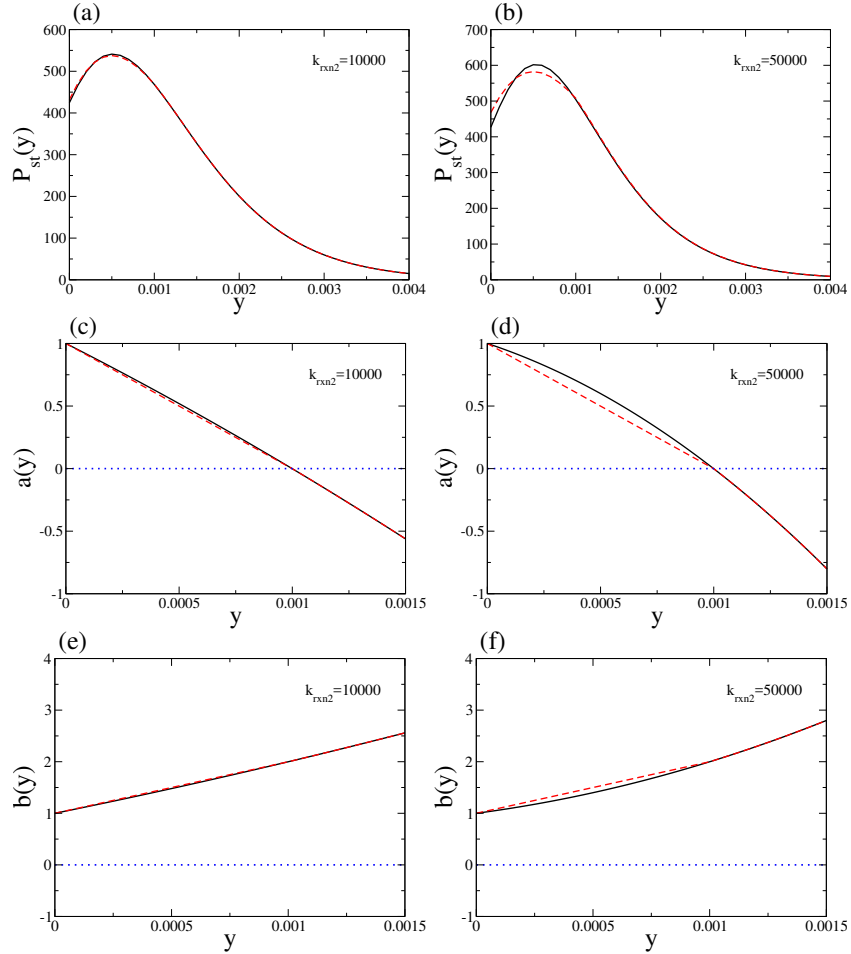
Figure 1: Stationary probability density function obtained with Eq. 22. (a) $k_{rxn2} = 10000$ and (b) $k_{rxn2} = 50000$. In (c) and (d) we plot the corresponding drift coefficient $a(y)$ as a function of y. Black solid lines correspond to plotting Eq. 22 with the Kramers-Moyal expansion term for the 2nd-order reaction. Red dashed lines – or gray on black/white prints correspond to Eq. 22 but with that term replaced by $\left(\frac{\zeta k_{rxn2}}{(1-\epsilon)} y(y - \epsilon) \vee 0 \right)$, where $\vee$ stands for maximum. Similarly, in (e) and (f) we plot the corresponding diffusion coefficient $b(y)$ as a function of y. For simplicity we assume that $k_{rxn1} = 0$. Other parameters are $k_{ads} = 1$, $k_{des} = 1000$, $\zeta = 4$ (square lattice), and $N = 1000$. Because the negative values appear for $y \leq \epsilon = 0.001$ and the 2nd-order reaction term is proportional to $y(y - \epsilon)$, we decided to consider large enough values of k_{rxn2} in order to have a clear appreciation of the impact of the replacement.

13

coefficients obtained by the truncation of the Kramers-Moyal expansion with the new term, $\left(\frac{\zeta k_{rxn2}}{(1-\epsilon)} y(y - \epsilon) \vee 0 \right)$, where $\vee$ stands for maximum. Physically, this means that the 2nd order reaction can only happen if 2 or more molecules of A exist on the surface. In Figs. 1(a) and 1(b) we plot the stationary PDF calculated using the the 2nd-order reaction term obtained from the Kramers-Moyal expressions (black lines) and the one calculated using the new positivity preserving term (red-dashed lines) (see Eq. 22). For the simulations we use $N = 1000$ ($\epsilon = 0.001$), and to have solutions very close to the low boundary ($y = 0$), we chose a very small value of k_{ads}. Moroever, to have a clear appreciation of the impact of the new term, we considered large enough values of k_{rxn2} (note that the 2nd-order reaction term is proportional to $y(y - \epsilon)$). The figures show that, although for $k_{rxn2} = 50000$ the impact of this replacement on the PDF is important, for $k_{rxn2} = 10000$ the two PDFs are practically indistinguishable. As an example, Figs. 1(c) and 1(d) also show the corresponding drift coefficients, $a(y)$. From these figures, it is clear that for $y < \epsilon = 1/N = 0.001$ the drift coefficients differ in both cases. However, when $k_{rxn2} = 10000$, such a difference is very small. A similar effect is observed for the diffusion coefficients plotted in Figs. 1(e) and 1(f). Note that for the following discussions concerning the low boundary at $y = 0$, we will consider $k_{rxn2} = 10000$.

Although it is often simpler to use the approximation based on the CFPE than the exact CME, in many cases one is interested in single realisations of the stochastic process under consideration and not only on probability densities or averages values. Therefore, in the following section we discuss the so-called chemical Langevin equation approach or CLE which in principle serves as a generator of single trajectories of the process described by the CFPE.

2.1.3. The Chemical Langevin equation approach

It is well known that the trajectories of the diffusion process y whose probability density is governed by Eq. 19 can be described by the so-called chemical Langevin equation or CLE. In particular, there exists a standard prescription

14

to go from a CLE for a stochastic variable y to the CFPE for the probability density of this variable $P(y;t)$ [49, 50, 51]. In our one dimensional problem, the CLE is an Itô SDE that takes the form

$$dy = a(y)dt + \sqrt{b(y)}dW, \tag{23}$$

where the drift ($a(y)$) and diffusion ($b(y)$) coefficients are given by Eqs. 17 and 18, respectively. The paramater $\epsilon = 1/N$ accompanying the diffusion coefficient (Eq. 18) is known as the intensity of the intrinsic noise or random fluctuations of the coverage. To avoid the possibility of having negative values in the 2nd-order reaction term of the drift and diffusion coefficients, when $y < \epsilon$, in the subsequent analysis of this SDE we implement the positivity preserving term introduced in Sec. 2.1.2. The term $dW = W(t+dt) - W(t)$ is a Wiener increment of the Wiener process W. A Wiener process is a stochastic process whose increments are independent and normally distributed with mean 0 and variance dt. The advantage of adopting the CLE description is that single trajectories of this stochastic equation can be obtained with comparable efficiency to that of ordinary differential equations (ODE). One can also note that as N increases, the fluctuating part of the CLE (or diffusion coefficient) will decrease relative to the drift coefficient. And in the thermodynamic limit of $N \to \infty$ ($\epsilon \to 0$), the last term of the CLE becomes negligible and the stochastic equation collapses to the corresponding deterministic rate equation given by

$$\frac{dy}{dt} = k_{ads}(1 - y) - (k_{des} + k_{rxn1})y - (\zeta k_{rxn2})y^2. \tag{24}$$

However, it is interesting to mention that, similar to the CFPE (Eq. 16), one can encounter issues with Eq. 23. Assume for example that we start the system very close to the boundary of zero concentrations ($y \to 0$ at $t = 0$). Provided that adsorption is dominant the drift terms will drive the system to positive coverages; however, the symmetric noise term due to adsorption $\epsilon k_{ads}(1 - y)$ can drive the system to negative coverages with a finite probability in a finite time interval. Something similar occurs when y is driven towards $y = 1$. In this case, the variable y becomes larger than one with a finite probability.

15

One simple method for numerically solving Eq. 23 is the Euler-Maruyama or EM scheme. This method has a strong rate of convergence equal to $1/2$ and a weak rate of convergence equal to 1 [45]. Note that weak convergence refers to the error of the mean, whereas strong convergence is related to the mean of the error in each individual path (we will come back to this point later) [45]. The scheme is based on time discretisation with points

$$0 = \tau_0 < \tau_1 < \cdots < \tau_i < \cdots < \tau_M = T, \tag{25}$$

in the interval $[0, T]$. Let $\Delta t = T/M$ and define $\tau_i = i\Delta t$. The numerical approximation to $y(\tau_i)$ will be y_i. For Eq. 23 the EM approximation gives the following recursive equation

$$y_{i+1} = y_i + a(y_i)\Delta t + \sqrt{b(y_i)}\Delta W_i, \tag{26}$$

with $y_0 = x_0$. Here $\Delta W_i = W(\tau_{i+1}) - W(\tau_i)$ are the increments of the Wiener process in the interval $[\tau_i, \tau_{i+1}]$ and are represented by independent $\mathcal{N}(0, \Delta t) = \sqrt{\Delta t}\,\mathcal{N}(0, 1)$ Gaussian random variables with mean zero and variance Δt [45].

The issues with Eq. 23 when trying to ensure y being inside interval $[0, 1]$ are presented in Eq. 26. For instance, it is easy to verify that Eq. 26 with the drift and diffusion coefficients given by Eqs. 17 and 18 will result in negative values for y. Suppose that at some time $y_i \approx 0$, then according to Eq. 26 the noise term has a non-zero variance ϵk_{ads}, and since the Gaussian distribution is symmetric, a negative Gaussian random variable can be generated resulting in a negative y_{i+1} (a similar argument applies when $y_i \approx 1$). In conclusion, the paths obtained by the original EM scheme can leave the interval $[0, 1]$ at some point in the simulation, which would lead to unphysical situations (coverage becoming negative or higher than 100%). Thus, we need to reformulate the CLE in such a way that it respects the boundary conditions of the CFPE and CME. One could reflect in a simplistic manner the occurrences that result in negative or positive ($y > 1$) values (mirror-boundary EM scheme); however, this will not necessarily result in correct solutions, as we will show later.

16

The issue of unphysical solutions of the CLE is clearly not just of a numerical nature. Although, to satisfy the boundaries of the CME, we imposed reflected boundary conditions on the CFPE, in switching to the CLE these boundary conditions were lost, and therefore, there is not guarantee that the solutions of Eq. 23 lie in $[0, 1]$. Consequently, a probability density function obtained from trajectories of the CLE is not necessarily comparable to the one obtained by solving the CFPE under reflected boundary conditions (or Eq. 5 (CME) in the limit of large N). In the following section we discuss a manner to reformulate the CLE in such a way that it respects the desired boundary conditions.

2.2. Reflected chemical Langevin equation with two-sided barriers

In the analysis of the CFPE the reflected boundary conditions were imposed by assuming no net flow of probability across $y = 0$ and $y = 1$. This condition guaranties reflection at the boundaries of the continuous process y because the probability to leave the domain $[0, 1]$ should be zero. However, as mentioned above the solution of the CLE could leave this domain with a non-zero probability.

To impose the boundaries condition given by Eq. 21 on the CLE (Eq. 23), let us decompose the continuous process y into the sum of two stochastic continuous processes $y(t) = x(t) + k(t)$ [40, 41]. The temporal dynamics of $x(t)$ is governed by Eq. 23, and therefore, it determines the behaviour of $y(t)$ inside $(0, 1)$. The new process $k(t)$ is the minimal process which forces $y(t)$ to remain in the unit interval. This reflecting process determines the behaviour at the boundaries and at $t = 0$ its initial value is usually assumed to be zero [40, 41]. Thus one has that $y(0) = x(0)$ and that $y(t) = x(t)$ on $(0, 1)$. Now we can introduce the SDE that takes into account the necessary reflective boundary conditions. This equation is called reflected stochastic differential equation (RSDE) or reflected chemical Langevin equation (RCLE), and for our representative model in differential form it is given by

$$dy = a(y)dt + \sqrt{b(y)}dW + dk, \qquad (27)$$

17

where $y(t) \geq 0$, $k(0) = 0$, and $k(t) = k^0(t) - k^1(t)$. The processes $k^0(t)$ and $k^1(t)$ are non-decreasing functions such that

$$k^0(t) = \int_0^t \mathbf{1}_{\{y(s)=0\}} dk^0(s), \qquad (28)$$

and

$$k^1(t) = \int_0^t \mathbf{1}_{\{y(s)=1\}} dk^1(s), \qquad (29)$$

with $k^0(0) = k^1(0) = 0$ This means that $k(t)$ changes only when $y(t) = 0$ or 1 (note that the indicator variables $\mathbf{1}_{\{y(s)=0\}}$ and $\mathbf{1}_{\{y(s)=1\}}$ evaluate to 1 if $y(s) = 0$ and 1, respectively). In other words, $dk(t)$ disappears when $y(t)$ does not lie on the boundaries, and Eq. 27 reduces to Eq. 23 [44, 55]. When $y(t) = 0$, $k(t)$ must push the process in the positive direction, while at $y(t) = 1$, $k(t)$ must push the process in the opposite direction

The first proof of the existence and uniqueness of solutions $(y(t), k(t))$ to reflected stochastic differential equations of the type, Eq. 27, was provided by Skorokhod [40, 41]. Normally, these types of equations are approximated using numerical methods, and a popular one is the projection method for RSDE [44]. This method is a simple extension of the EM scheme, given in Sec. 2.1.3. At time t the non-reflected process is evaluated at the next step $t + \Delta t$ using the traditional EM scheme (Eq. 26). If this value lies within the defined domain, then the process at the next time step is set to this value. Otherwise it is equal to the orthogonal projection of this point onto the boundary of the domain. In contrast to the EM scheme, it has been shown that the projection method has a strong order of convergence of $1/2 - \sigma$, for $\sigma > 0$ [56, 57]. The lower rate of convergence is normally attributed to the fact that, between t and $t + \Delta t$, the process may leave the domain and lie again in the domain without its excursions having any effect on the value of the numerical approximation at $t + \Delta t$. Better approximations are normally obtained by watching the path between $t + \Delta t$ and t and not allowing it to leave the domain. Therefore, in this work, we choose to implement a method due to Lépingle which solves this problem by sampling

18

from the exact distribution of the reflecting process at each time step [44]. This method has been shown to have a similar strong rate of convergence as that of the EM scheme. This numerical scheme is summarised in the following section.

2.2.1. The Euler-Lépingle scheme for two-sided barriers

In references [44, 58], Lépingle introduced a feasible numerical scheme to solve one-dimensional RSDEs for the simple situation in which a reflective plane exists at $y = 0$ (see supplementary information for details of this method). The method is based on the following explicit expression

$$k^0(t) = \sup_{0 \leq s \leq t} \left[-x(s) \right]^+ , \tag{30}$$

where $a^+ = a \vee 0$, $\vee$ stands for maximum, sup stands for supremum, and $x(s)$ is the unreflected process governed by Eq. 23. The role of the process $k^0(t)$ is to ensure that the modified process $y(t) = x(t) + k^0(t)$ remains greater than or equal to zero, where $x(t)$ is the unreflected process described by Eq. 23 [40, 41].

However, although for the one-dimensional case with two-side barriers there is also an expression for the corresponding reflecting process [56, 59], it is rather cumbersome to be implemented in a numerical method. Nevertheless, in direct analogy to Eq. 30, it is easy to see that $k^1(t)$ defined by

$$k^1(t) = \sup_{0 \leq s \leq t} \left[x(s) - 1 \right]^+ , \tag{31}$$

insures that the modified diffusion process $y(t) = x(t) - k^1(t)$ remains less than or equal to one, where $x(t)$ is also governed Eq. 23 [59].

The process $k^0(t)$ represents the pushing up from zero that is needed to keep $y(t) \geq 0$ for all time, and $k^1(t)$ represents the pushing down from 1 that is needed to keep $y(t) \leq 1$ for all time. Thus, after assuming that a path can not be simultaneously close to the low and upper boundaries of the domain and that in a small interval of time there is very small chance that lower and upper reflecting process both have to work, Lépingle also proposed a numerical scheme for the two-side barriers problem which is based on the Euler scheme together

19

with Eqs. 30 and 31 [44]. For our case study, it is given by

$$y_{i+1} = 0 \vee \left(y_i + a(y_i)\Delta t + \sqrt{b(y_i)}\Delta W_i + C_{i+1} \right) \wedge 1, \tag{32}$$

where

$$C_{i+1} = \mathbf{1}_{\{y_i < \bar{\alpha}\}} \left((\Gamma_i^0 - y_i) \vee 0 \right) - \mathbf{1}_{\{y_i > \bar{\beta}\}} \left((\Gamma_i^1 + (y_i - 1)) \vee 0 \right), \tag{33}$$

with

$$\Gamma_i^0 = \sup_{\tau_i \leq s \leq \tau_{i+1}} \{ -a(y_i)(s - \tau_i) - \sqrt{b(y_i)}(W(s) - W(\tau_i)) \}, \tag{34}$$

and

$$\Gamma_i^1 = \sup_{\tau_i \leq s \leq \tau_{i+1}} \{ a(y_i)(s - \tau_i) + \sqrt{b(y_i)}(W(s) - W(\tau_i)) \}, \tag{35}$$

where $y_0 = x_0$, the time discretisation is given by Eq. 25 ($\Delta t = \tau_{i+1} - \tau_i$), and $\Delta W_i = W(\tau_{i+1}) - W(\tau_i) = \sqrt{\Delta t}\, \mathcal{N}(0,1)$. Note that $0 < \bar{\alpha} < \bar{\beta} < 1$ and $\wedge$ stands for minimum. The indicator variables $\mathbf{1}_{\{y_i < \bar{\alpha}\}}$ and $\mathbf{1}_{\{y_i > \bar{\beta}\}}$ evaluate to 1 if $y_i < \bar{\alpha}$ and $y_i > \bar{\beta}$, respectively. The simulation of this scheme requires the simulation of the pair $(\Delta W_i, \Gamma_i^0)$ or $(\Delta W_i, \Gamma_i^1)$ at each time step. This scheme is feasible because we already know how to simulate ΔW_i (increments of a Wiener process W), and proceeding as for the case of a single barrier at $y = 0$ (see supplementary material and references [44, 58]), one has that

$$\Gamma_i^0 = \frac{1}{2} \left(\Lambda_i + \sqrt{b(y_i)\vartheta_i + \Lambda_i^2} \right), \tag{36}$$

and

$$\Gamma_i^1 = \frac{1}{2} \left(-\Lambda_i + \sqrt{b(y_i)\vartheta_i + \Lambda_i^2} \right), \tag{37}$$

with

$$\Lambda_i = -a(y_i)\Delta t - \sqrt{b(y_i)}\Delta W_i, \tag{38}$$

where the term ϑ_i is an exponential random variable with rate parameter $(2\Delta t)^{-1}$. If one does not want to compute too many correction terms, one can take $\bar{\alpha}$ close to 0 and $\bar{\beta}$ close to 1. In summary, the Lépingle or Euler-Lépingle scheme for two barriers amounts to performing exact reflection on the lower boundary on $[\tau_i, \tau_{i+1}]$ when $y_i < \bar{\alpha}$ and exact reflection on the upper boundary

20

on $[\tau_i, \tau_{i+1}]$ when $y_i > \bar{\beta}$. Then, when at the end of the time (iteration) step the computed value is smaller that 0 or greater than 1, y_{i+1} is respectively given the value 0 or 1. It was shown that the aforementioned numerical method has a strong rate of convergence equal to 1/2 (similar to the EM scheme) [44]. In the supplementary material we present a detailed description of the Lépingle's numerical scheme to solve one-dimensional RSDEs for the simple situation in which a reflective plane exists at $y = 0$. This description contains a proposition that leads to Eqs. 36 and 37.

The rest of this article deals with the comparison of the EL numerical solutions of the RCLE for our single species model, with the solutions obtained using the mirror-boundary EM scheme, the CFPE with reflected boundary conditions, and the CME and SSA. We also analyse the rates of convergence of the EL scheme.

3. Simulation results

For the purpose of testing the EL scheme, the choice of parameter values (reaction constants) is motivated by our aim of testing the performance of this computational scheme close to the low or upper boundaries. Thus, all the simulation results presented below are obtained with reaction constants giving stochastic paths very close to the boundaries $y = 0$ and $y = 1$. In particular, for the low boundary cases we consider, without loss of generality, parameter values in line with Fig. 1(a).

Figure 2 shows a case for parameter values giving paths very close to the low barrier $y = 0$. Panel (a) of Fig. 2 presents a stochastic path from a simulation of the CLE using a mirror-boundary EM scheme where, if the numerical solution becomes negative, we consider as input to the next time step the absolute value of it (one simply reflects the instances that result in negative values for y). On the other hand, panel (c) shows a stochastic path obtained solving Eq. 27 with the more rigorous Euler-Lépingle (EL) scheme for two barriers. The plots of
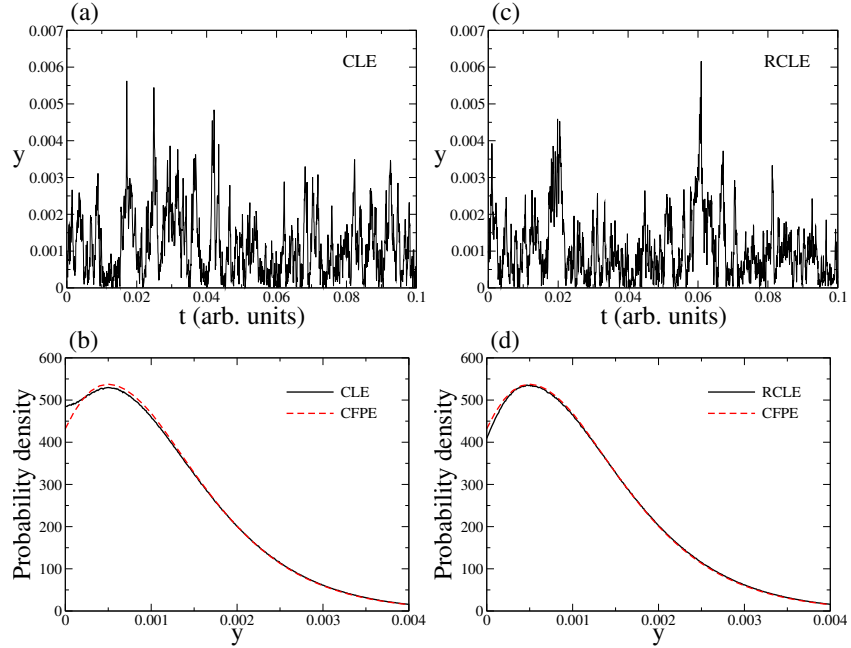
21

Figure 2: Panel (a) shows transient simulations obtained by solving the CLE using the mirror-boundary EM scheme. This is a scheme that reflects back to the domain the instances where y tries to cross the boundary 0. Panel (c) shows transient simulations obtained by solving the RCLE with the EL scheme. Panels (b) and (d) show stationary PDFs obtained with Eq. 22 and by solving the CLE with the mirror-boundary EM scheme (black solid lines) and the RCLE with the EL scheme (red dashed lines – or gray on black/white prints). For the PDFs we consider $t_{ini} = 0$ and $t_{fin} = T = 5000$. For all simulations $k_{ads} = 1$, $k_{des} = 1000$, $k_{rxn1} = 0$, $k_{rxn2} = 10000$, $\zeta = 4$ (square lattice), and $N = 1000$. For the EL scheme we use $\bar{\alpha} = 0.1$ and $\bar{\beta} = 0.9$. For time seres and PDFs we use $\Delta t = 2 \times 10^{-14}$.

panel (a) and (c) are just representative transients. The error introduced by the mirror-boundary EM scheme becomes apparent if one calculates the respective

470 stationary probability density functions and compares them with Eq. 22 (panels (b) and (d)). The mirror-boundary EM scheme introduces a significant error close to $y = 0$, with the probability profile appearing to artificially flatten-out towards the boundary (see Fig. 2(b)). On the other hand, the shape of the solution of Eq. 22 is better reproduced by the EL scheme (see Fig. 2(d)).

475 Figure 3 shows a case of paths very close to the upper barrier. In this case, we
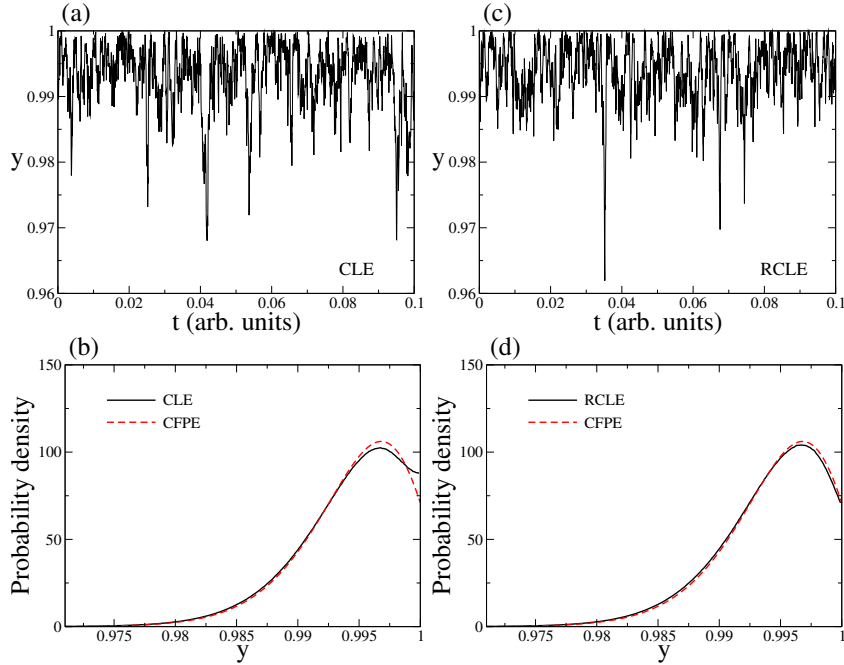
Figure 3: Panel (a) shows transient simulations obtained by solving the CLE using the mirror-boundary EM scheme. This is a scheme that reflects back to the domain the instances where y tries to cross the boundary 1. Panel (c) shows transient simulations obtained by solving the RCLE with the EL scheme. Panels (b) and (d) show stationary PDFs obtained with Eq. 22 and by solving the CLE with the mirror-boundary EM scheme (black solid lines) and the RCLE with the EL scheme (red dashed lines – or gray on black/white prints). For the PDFs we consider $t_{ini} = 0$ and $t_{fin} = T = 5000$. For all simulations $k_{ads} = 2000$, $k_{des} = 10$, $k_{rxn1} = 0$, $k_{rxn2} = 0.1$, $\zeta = 4$ (square lattice), and $N = 250$. For the EL scheme we use $\bar{\alpha} = 0.1$ and $\bar{\beta} = 0.9$. For time series and PDFs we use $\Delta t = 2 \times 10^{-14}$.

reduced the rate constants of desorption (k_{des}) and 2nd-order reaction (k_{rxn2}) but increased the one for adsorption (k_{ads}). For simplicity we also assume that $k_{rxn1} = 0$. Figures 3(b) and (d) shows that, close to the upper barrier $y = 1$, the EL scheme also reproduces very well the solution of the CFPE, while the mirror-boundary EM scheme fails to reproduce it. In this case the mirror-boundary reflects the instances that result in values for y larger than one, while solving Eq. 23 using the EM scheme (if the numerical solution becomes larger than one, we consider as input to the next time step one minus the amount by which

23

the solution overshoots the $y = 1$ boundary). As before, the path presented in

485 Figs. 3(a) and (c) are just representative transients.

To continue elucidating the advantage of using the EL scheme to numerically integrate Eq. 27, in the following section we numerically estimate the strong and weak rates of convergence of it and compare them with the corresponding rates of convergence of the traditional EM scheme.
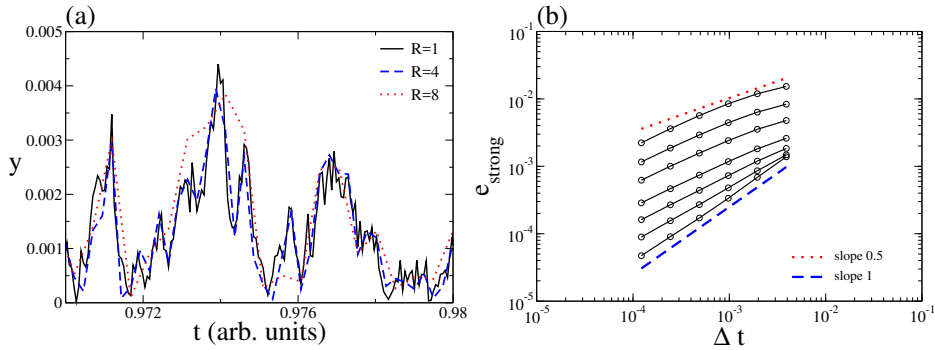

Figure 4: (a) Solutions of the RCLE obtained with the EL scheme for time steps $\Delta t = R\delta t$, where $R = 2^{s-1}$ for $s = 1$, 3, and 4. In this case $N = 1000$ and $\delta t = 2^{-14}$ (exact solution). (b) Strong error plotted in circles for $N = 25$, 50, 100, 250, 500, 1000, 2000 from top to bottom (solid lines are guides to the eye). The assumed exact solution obtained with $\delta t = 2^{-14}$ is compared with other numerical solutions with $\Delta t = R\delta t$, where $R = 2^{s-1}$ for $s = 2, \cdots, 7$. For all simulations $k_{ads} = 1$, $k_{des} = 1000$, $k_{rxn1} = 0$, $k_{rxn2} = 10000$, and $\zeta = 4$ (square lattice). We use $\bar{\alpha} = 0.1$ and $\bar{\beta} = 0.9$. In (b) dashed blue line and dotted red line are the appropriate reference slope in each case. Note that $\epsilon = 1/N$. The error was calculated from Eq. 41 at time $T = 1$.

490 *3.0.1. Convergence rates of Euler-Lépingle scheme*

In this section we use our illustrative model to calculate the rates of convergence of the EL scheme. If y_i is the approximate numerical solution of the RCLE and $y(t)$ is the exact one, we say that the EL scheme has strong order of convergence equal to γ if there exists a constant C such that

$$\mathbb{E}\left(\left|y_i - y(t)\right|\right) \leq C\Delta t^{\gamma}, \tag{39}$$

24

for any fixed $t = i\Delta t \in [0, T]$ and Δt sufficiently small. Note that $\mathbb{E}$ denotes the expected value. In our numerical analysis, we focus on the error at the end point $t = T$, so we have

$$e_{strong} = \mathbb{E}\left(\left|y_M - y(T)\right|\right),\tag{40}$$

where $M\Delta t = T$. If Eq. 39 holds at any point inside $[0, T]$, it also holds at the end point. Thus

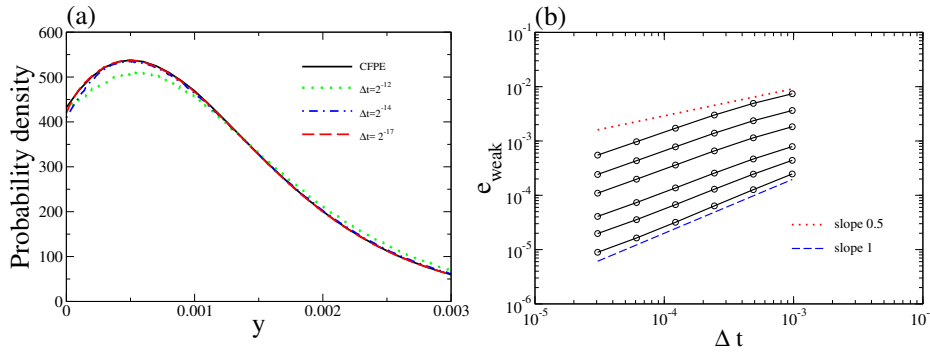$$e_{strong} \leq C\Delta t^\gamma.\tag{41}$$

Figure 5: (a) Stationary probability density function of the CFPE obtained with Eq. 22 (back solid line) and three numerical stationary PDFs with $\Delta t = 2^{-12}$ (green dotted line), $\Delta t = 2^{-14}$ (blue dash-dotted line), and $\Delta t = 2^{-17}$ (red dashed line) approaching it from below. (b) Weak error plotted in circles for $N = 25, 50, 100, 250, 500, 1000$ from top to bottom (solid lines are guides to the eye). The stationary mean obtained from Eq. 22 is compared with the solutions of the RCLE obtained by using the EL scheme with $\Delta t = 2^{s-16}$, for $s = 1, \cdots, 6$. For all simulations $k_{ads} = 1$, $k_{des} = 1000$, $k_{rxn1} = 0$, $k_{rxn2} = 10000$, and $\zeta = 4$ (square lattice). We use $\bar{\alpha} = 0.1$ and $\bar{\beta} = 0.9$. In (b) dashed blue line and dotted red line are the appropriate reference slope in each case. Note that $\epsilon = 1/N$. The error was calculated from Eq. 42 at time $T = 5000$.

For our calculations we first compute a discretised Brownian path over $[0, T]$ with $\delta t = 2^{-14}$ and use it to iterate Eq. 32. This iteration is considered as the exact solution. Then, we iterate again Eq. 32 using the same Brownian path

25

but for time steps $\Delta t = R \delta t$, where $R = 2^{s-1}$ and s an integer number larger than one. This ensures that the set of points in which the discretised Brownian path is based contains the points τ_i at which the EL solution is computed. Figure 4(a) shows several EL solutions for parameter values that generate paths close to the low barrier of $y = 0$. It is clear that the path with $s = 1$ ($R = 1$) corresponds to the assumed exact solution, whereas the other two time series show that effectively the EL solutions approximate the exact solution better as Δt decreases. In Fig. 4(b) we plot the end point error as a function of Δt for several system sizes N (or noise intensities ϵ). The figure shows that the slope of the curves drops from 1 to $1/2$ as the system size N decreases (or the noise intensity ϵ increases). These results are consistent with a strong order of convergence of the EL scheme changing between $\gamma = 1$ and $\gamma = 1/2$ as the intrinsic noise intensity increases. This is in agreement with Lépingle's mathematical proof of a strong order of convergence of one-half [44] (note that the commonly used projection method has a smaller strong order of convergence [56, 57]). The order of convergence of $\gamma = 1$ in the small noise limit coincides with the order of convergence of the deterministic Euler scheme which is usually implemented to solve Eq. 24 [45]. This exchange is due to the fact that the noise terms of the EL scheme decrease relative to the deterministic one (or drift coefficient), as N increases. Eventually, the noise terms are negligible and the EL recursive equation approach the well-known Euler recursive equation. Note that if the numerical scheme is convergent with order γ, and we make the step, Δt, l times smaller then the approximation error will decrease by a factor of l^γ. Therefore, the order of convergence equal one means that if we want to decrease the error 10 times, we have to make the step 10 time smaller. The order equal to one-half means that if we want to decrease the error 10 times, we have to make the step $10^2 = 100$ smaller. And the computational time grows by the same factor.

It is also interesting to explore the weak order of convergence of the EL scheme. Figure 5(a) shows that stationary solutions of the EL scheme in fact

converge in probability to the stationary solution of the CFPE when Δt decreases. In this work we calculate the weak error as the absolute value of difference $\mathbb{E}_{st}(y_i) - \mathbb{E}_{st}(y)$, where $\mathbb{E}_{st}(y) = \int_0^1 y P_{st}(y)$ with $P_{st}(y)$ given by Eq. 22, $\mathbb{E}_{st}(y_i)$ is obtained with the EL scheme after averaging over several independent iterations, and i is the iteration step of the scheme. Here, $\mathbb{E}_{st}$ in the expectation value at steady state. Thus, we say that at steady state the scheme has weak order of convergence equal to γ if there exists a constant C such that

$$e_{weak} = \left| \mathbb{E}_{st}(y_i) - \mathbb{E}_{st}(y) \right| \quad \leq C\Delta t^\gamma. \tag{42}$$

Figure 5(b) shows simulation results that are consistent with a weak order of convergence equal to $\gamma = 1$, for both, the weak and strong noise cases. This order of convergence is equal to the weak order of convergence of the EM scheme in the same limits [45]. It also agrees with the weak order of convergence proved in [47] for the case of reflected stochastic differential equations or RSDEs with addictive noise. However, to the best of our knowledge, this is the first numerical calculation of the weak order of convergence of the EL scheme in a model involving a multiplicative noise term.

In Figs. 2 and 3 we showed that the RCLE numerically solved with the EL scheme reproduces very well the solution of the CFPE. We also numerically verified that the EL scheme, in which every new step requires the easy simulation of a Gaussian variable and a new independent exponential variable, yields the same rates of convergences as in the usual EM scheme (see Figs. 4 and 5). Thus, in the following sections we explore how accurately this scheme reproduces solutions of the CME.

3.1. Comparison of Euler-Lépingle scheme with steady state solutions of the chemical master equation

In this work we assume that the approximation of the CME by the CFPE and its corresponding SDEs only holds for larger N, and so for small N the diffusion process approximation may no longer accurately reproduce the dynamics of
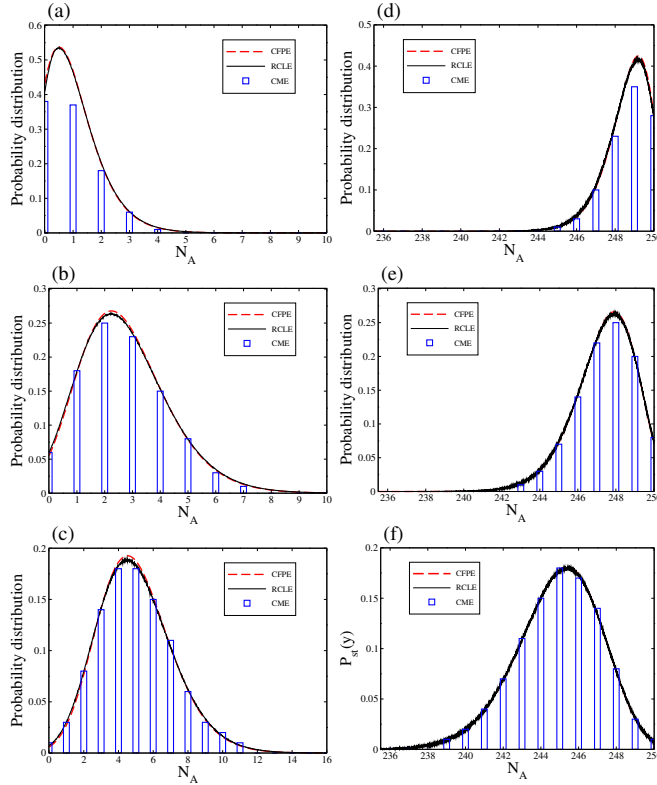
Figure 6: Comparisons of stationary discrete versus continuous distributions (CME versus CFPE/RCLE) for a fixed-size lattice while varying k_{ads}. The stationary discrete distribution functions were obtained by numerically solving the CME at steady state (bar plots). The continuous stationary distributions are obtained from the solution of the CFPE (red dashed lines – or gray on black/white prints) but also from the numerical integration of the RCLE using the EL scheme (solid black lines). Low boundary cases (a), (b), and (c) are with $k_{ads} = 1$, 3, and 6, respectively. In this situation, $k_{des} = 1000$, $k_{rxn1} = 0$, $k_{rxn2} = 10000$, $N = 1000$, and $\zeta = 4$ (square lattice). Upper boundary cases (d), (e), and (f) are with $k_{ads} = 2000$, 1000, and 500, respectively. In this situation, $k_{des} = 10$, $k_{rxn1} = 0$, $k_{rxn2} = 0.1$, $N = 250$, and $\zeta = 4$ (square lattice), For the EL scheme $t_{ini} = 0$, $t_{fin} = T = 5000$, and $\Delta t = 2^{-14}$. We use $\bar{\alpha} = 0.1$ and $\bar{\beta} = 0.9$. The discrepancy between the CFPE and RCLE becomes small as Δt decreases (not shown in figure).

the discrete-state Markov chain described the CME. Thus, in this section, we compare stationary discrete distributions obtained with the RCLE solved using the EL scheme or Eqs. 32-38 (black solid lines) and the CFPE or Eq. 22 (red
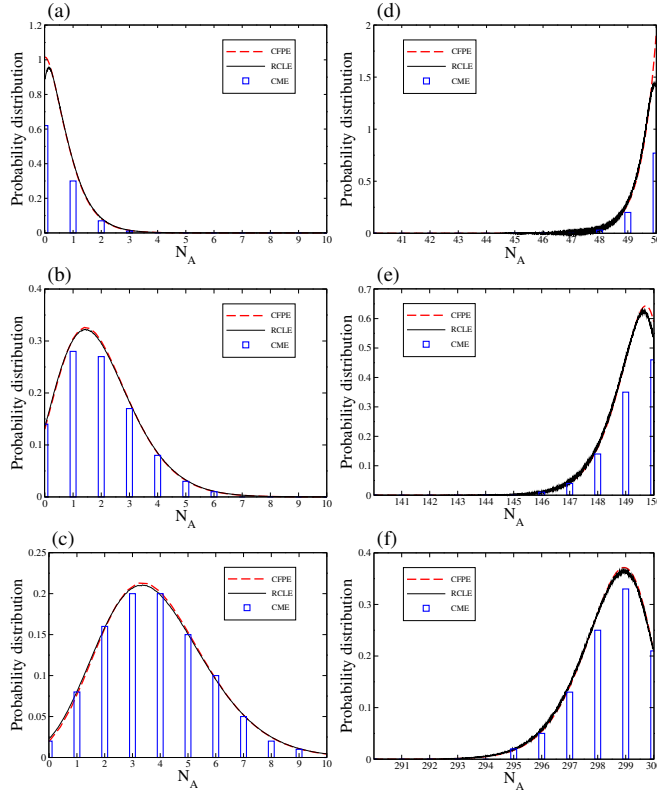
Figure 7: Comparisons of stationary discrete versus continuous distributions (CME versus CFPE/RCLE) for lattices of increasing size, N. The stationary discrete distribution functions were obtained by numerically solving the CME at steady state (bar plots). The continuous stationary distributions are obtained from the solution of the CFPE (red dashed lines – or gray on black/white prints) but also from the numerical integration of the RCLE using the EL scheme (solid black lines). Low boundary cases (a), (b), and (c) are with $N = 500$, 2000, and 4000, respectively. In this situation, $k_{ads} = 1$, $k_{des} = 1000$, $k_{rxn1} = 0$, $k_{rxn2} = 10000$, and $\zeta = 4$ (square lattice). Upper boundary cases (d), (e), and (f) are with $N = 50$, 150, and 300, respectively. In this situation, $k_{ads} = 2000$, $k_{des} = 10$, $k_{rxn1} = 0$, $k_{rxn2} = 0.1$, and $\zeta = 4$ (square lattice). For the EL scheme $t_{ini} = 0$, $t_{fin} = T = 5000$, and $\Delta t = 2^{-14}$. We use $\bar{\alpha} = 0.1$ and $\bar{\beta} = 0.9$. The discrepancy between the CFPE and RCLE becomes small as Δt decreases (not shown in figure).

dashed lines) with stationary distribution functions of the CME (bar plots). The CME is solved numerically at steady state by solving a set of $N + 1$ algebraic equations along with steady state equations for the reflective boundaries (see

Eqs. 5, 6, and 7). Figure 6 shows that for the low (upper) boundary the agreement between the three approaches becomes progressively better as the rate constant of adsorption k_a increases (decreases). The inability of the RCLE and CFPE to reproduce the CME results very close to the boundaries is due to the inherent continuous character of these two approaches. Figure 7 also shows the stationary discrete distribution of the CFPE (red dashed lines), the RCLE (black solid lines), and the CME (bar plots) for different system sizes N. It is evident that, as expected, the two continuous descriptions approach the CME as the system size increases and we move away from the boundaries.

3.2. Computational savings

In order to analyse the computational savings of the EL scheme, we compare the numerical solutions of the RCLE with individual stochastic realisations of N_A simulated using the so-called stochastic simulation algorithm (SSA) due to Gillespie (we refer to Sec. 2.1.1 for a presentation of this method).

Figures 8(a) and (b) summarise the results of these comparisons for the case of paths very close to the low barrier. Panel (a) shows transients of N_A obtained with the EL scheme (red-dashed line) and the SSA (black-solid line). In panel (b) we compare the mean simulation times of the EL scheme and the SSA. These are the computational times needed to simulate the system over a fixed time interval. In general the figure shows that, although for the parameter values considered, the SSA would be the more convenient choice when the system size N is small, the RCLE approach offers computational savings when the system size is large. Furthermore, it clearly shows that the computational time of the SSA increases with the system size N, while that of the RCLE remains more or less constant. This leads to a crossover between the two lines at some critical value of N. This is due to the fact that, while the computational cost of the EL scheme is only dependent on the Δt used, the SSA simulates each and every reaction event in the system. The number of these reaction events before reaching the end time of the simulation increases with the system size N (i.e.
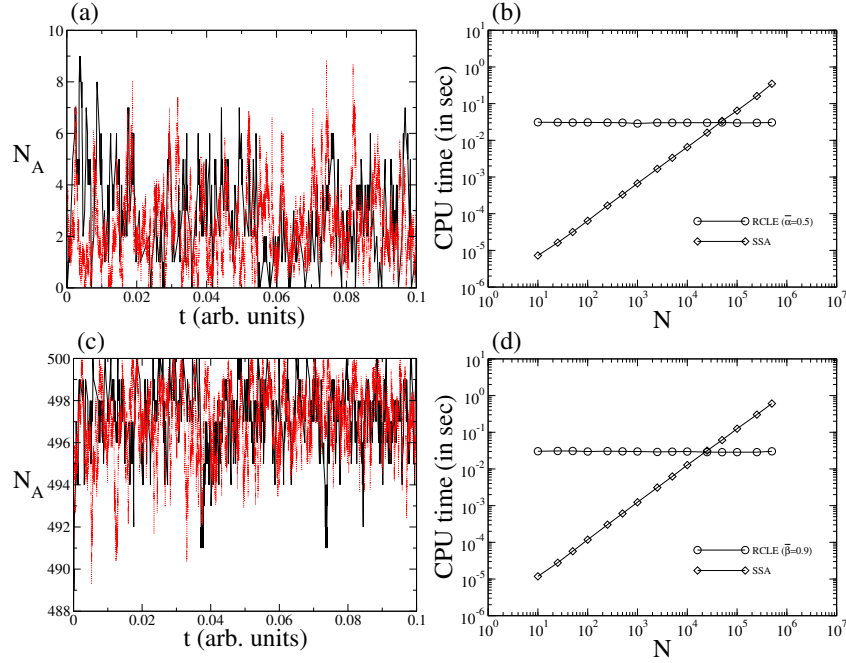
30

Figure 8: (a) and (b) show simulations for the low barrier case, with $k_{ads} = 6$, $k_{des} = 1000$, $k_{rxn1} = 0$, $k_{rxn2} = 10000$, and $\zeta = 4$ (square lattice). (a) Comparison of the transients obtained with the SSA (black solid line) and the RCLE integrated using the EL scheme (red dashed line – or gray on black/white prints). In these simulations $N = 500$, $\bar{\alpha} = 0.5$, and $\bar{\beta} = 0.9$. (b) Comparison of the mean computational times of the RCLE integrated using the EL scheme with the SSA for parameter values leading to a path closes to $y = 0$. Mean simulation times are obtained after 2000 independent realisations. In all cases simulations are stopped at $t = T = 1$ and for the EL scheme we considered $\Delta t = 2^{-17}$ and $\bar{\beta} = 0.9$. (c) and (d) show simulations for the upper barrier case, with $k_{ads} = 2000$, $k_{des} = 10$, $k_{rxn1} = 0$, $k_{rxn2} = 0.1$, and $\zeta = 4$ (square lattice). In this case, the description of panel (c) is similar to panel (a), but with $\bar{\alpha} = 0.1$ and $\bar{\beta} = 0.9$. Panel (d) is similar to panel (d) but with $\bar{\alpha} = 0.1$.

see Sec. 2.1.1 and reference [16] for details). Figures 8(c) and (d) show that something similar occurs for parameter values giving paths very close to the upper barrier.

Finally, it is interesting to note that, for our single species reaction model, it has been already demonstrated that simulating the CME is much faster than performing full on-lattice KMC simulations with fast diffusion [26]. Therefore

31

the computational savings of RCLE versus full KMC simulations can be tremendous for mesoscale systems (N of the order of a thousand or more sites).

4. Summary and conclusions

It is well-recognised that at the nano- and mesoscale, intrinsic noise or molecular fluctuations due to the erratic and incessant motion of atoms, has a strong impact on reactions and diffusion of matter. This is the case for chemical reactions occurring on catalytic surfaces [6, 7, 8, 9, 10, 11, 4].

The traditional tools used to theoretically investigate small catalytic systems have been the on-lattice KMC method, the so-called well-mixed CME, and its statistically equivalent SSA approach. However, it is also well known that the computational implementation of these approaches can be very expensive. Due to its widely recognised computational speedup, an attractive alternative to the aforementioned methods is the so-called CLE approach and its corresponding CFPE. This approach has already been implemented to model different aspects of surface reactions. However, the validity and accuracy of the solutions of the CLE have recently been questioned. In particular, it can be verified that in some cases numerical solution schemes can give rise to negative or even imaginary solutions. For many CLEs describing catalytic reaction systems this can happen when the noise term is finite, due to e.g an adsorption process, at coverage close to zero.

In this paper we have contributed detailed benchmarks of reflected CLE solution methods on a catalytic/chemical system, and have demonstrated the numerical accuracy of the Euler-Lépingle (EL scheme) showing that the numerical solutions thus obtained correctly reproduce physically realistic surface coverages. More specifically, we dealt with the issue of ensuring positivity of the solutions of the CLE, while also imposing that coverage fractions of species do not exceed the maximum value of unity (100%). To illustrate our results, we considered a simple reaction network consisting of a single species. We demon-

strated that the solutions are guaranteed to remain bounded in $[0, 1]$ provided appropriate reflection terms are added to the traditional CLE. In this way we introduced the reflected CLE or RCLE in line with the seminal Skorokhod's work [40, 41]. We showed that this formulation reproduces very well the diffusion process whose probability density function is described by the CFPE with reflected boundary conditions obtained after a second order truncation of the Kramers-Moyal expansion of the CME. We concluded that the RCLE constitutes a physically realistic and mathematically consistent modelling framework that can be effectively implemented to understand the role of intrinsic noise on the complex and non-linear kinetics of surface reactions. Furthermore, to numerically solve the RCLE we implemented the EL scheme. This is a numerical scheme that samples from the exact distribution of the reflected process at each time step of the numerical integration. It only requires the easy simulation of a Gaussian variable and an independent exponential variable [44]. We would like to note that the RCLE approach is more general than the also mathematically rigorous complex CLE method, in which however it is not clear how to impose the restriction that coverage cannot exceed 100% [34].

We then proceeded to calculate the error of the EL scheme. In particular, we numerically demonstrated that the strong order of convergence of the scheme increases from $1/2$ to 1 as the size of the surface increases (resulting in lower noise intensity). To our knowledge this is the first numerical verification of the strong order of convergence of this scheme in the small noise limit. Moreover, the strong order of convergence of $1/2$ that we obtained numerically coincides with the strong order of convergence already proved by Lépingle [44]. We also computed the weak order of convergence of the method at steady state conditions and showed that the order is always 1. For this scheme, a weak order of convergence equal to 1 has been already proven for the case of additive noise [47]. However no such convergence analysis has been carried out for a system with multiplicative noise (as is the case of the model implemented in this work). Thus, our computations suggest that, in contrast to the commonly used

33

projection method, the EL scheme has the same rate of convergence as in the traditional Euler-Maruyama scheme [45]. This high order of convergence makes the EL scheme very attractive for the computational integration of RCLEs.

We further compared the numerical simulations of the RCLE and the solutions of the CFPE with solutions of the CME and the corresponding SSA. We showed that in general the stationary probability distributions obtained by the integration of the RCLE using the EL scheme are in good agreement with the solutions of the CME equation. The observed deviations, very close to the barriers and for very small system sizes, are due to the discrete nature of the system which can not be properly captured by the continuous CFPE and RCLE approaches. Moreover, we demonstrated that for systems with a large number of sites (i.e large surfaces) the RCLE is more computationally efficient than the SSA.

Although we have focused on a single species reaction network containing some representative reaction steps, the same methodology can easily be extended to realistic surface reaction systems containing more than a single species. Moreover, the RCLE approach is an attractive methodology to numerically investigate a number of realistic catalytic surface reactions where the impact of intrinsic noise has been experimentally observed [6, 7, 8, 9, 9, 10, 11, 12]. Those catalytic systems can be analysed through reflected chemical Langevin equations like the one presented in this work, and integrated using the Euler-Lépingle scheme.

It is interesting to mention that the effects induced by extrinsic or parametric noise on macroscopic surface reactions have also been experimentally and theoretically explored [60, 61, 62]. Recently, several theoretical results concerning the impact of molecular noise on electrochemical reactions occurring on nano- and mesoscale electrodes have been also reported [13, 14]. The methods we have discussed in this work can be applied *mutatis mutandis* to these other catalytic systems. For the case of external noise on macroscopic surfaces, the analysis should include stochastic partial differential equations for reaction-diffusion pro-

34

cesses [63]. In the case of electrochemical reactions on nanoscale surfaces, one has to consider that the rate constants are themselves time-fluctuating quantities in addition to the normal chemical variables [14].

As a final note on the applicability and future extensions of this work, we would like to emphasise that the RCLE approach that we presented is valid in the so-called fast-diffusion limit and for mesoscopic system sizes. If diffusion rates are not sufficiently large or if the system size is very small, one should use the well-mixed CME or on-lattice CME with appropriate boundary conditions. A future interesting extension of our work is concerning the robustness/sensitivity of the LE scheme in the presence of model parametric uncertainty. We can identify two ways/approaches to do this: (i) considering ensemble of simulations over a distribution of rate constants, but with the parameter set fixed for each simulation of the ensemble, (ii) treating the rate constants as stochastic processes themselves. The first way would model parametric uncertainly when, for instance, we do not know the exact value for the activation energy of an event, but we know a range. The second way could be relevant when inherent fluctuations in the pressure or temperature of the system (or even species that are treated implicitly, within lumped parameters) make the kinetic constants themselves random functions of time. This is an interesting extension of our work that will be explored in the near future.

References

[1] G. Ertl, Angew. Chem., Int. Ed., 47 (2008) 3524–3535.

[2] R. Imbihl, New J. Phys., 5 (2003) 62.

35

[3] C. N. Satterfield, Heterogeneous Catalysis in Practice, New York: McGraw-Hill, 1980.

[4] V. Johanek, M. Laurin, A. W. Grant, B. Kasemo, C. R. Henry, J. Libuda, Science., 304 (2004) 1639–1644.

[5] Y. S. Lim, M. Berdau, M. Naschitzki, M. Ehsasi, J. H. Block, J. Catal., 149 (1994) 292–299.

[6] Y. Suchorski, J. Beben, R. Imbihl, E. W. James, D. J. Liu, J. W. Evans, Phys. Rev. B., 63 (2001) 165417.

[7] Y. Suchorski, J. Beben, E. W. James, J. W. Evans, R. Imbihl, Phys. Rev. Lett., 82 (1999) 1907–1910.

[8] P. Grosfils, P. Gaspard, T. V. de Bocarme, J. Chem. Phys., 143 (2015) 064705.

[9] Y. D. Decker, D. Bullara, C. Barroo, T. V. de Bocarmé, Nonlinear Dynamics of Reactive Nanosystems: Theory and Experiments., Springer Series in Materials Science, 2015.

[10] C. Barro, Y. D. Decker, T. V. de Bocarme, N. Kruse, Phys. Rev. Lett., 117 (2016) 144501.

[11] C. Barro, Y. D. Decker, T. V. de Bocarme, J. Phys. Chem. C., 121 (2017) 17235–17243.

[12] N. V. Peskov, M. M. Slinko, N. I. Jaeger, J. Chem. Phys., 116 (2002) 2098.

[13] V. García-Morales, K. Krischer, Proc. Natl. Acad. Sci. U.S.A., 107 (2010) 4528–4532.

[14] F. G. Cosi, K. Krischer, Eur. Phys. J. Special Topics., 226 (2017) 1997.

[15] S. Bozdech, K. Krischner, D. A. Crespo-Yapur, E. Savinova, A. Bonnefont, Faraday Discuss., 193 (2016) 187.

36

[16] D. T. Gillespie, Stochastic Chemical Kinetics.,, In Handbook of Materials Modeling, ed. S Yip, pp. 1735–52. Dordrecht: Springer, 2005.

[17] D. Schoerr, G. Sanguinetti, R. Grima, J. Phys. A: Math. Theor., 50 (2017) 093001.

[18] D. J. Liu, J. W. Evans, Prog. Surf. Sci., 88 (2013) 393.

[19] L. Kunz, F. M. Kuhn, O. Deutschmann, J. Chem. Phys., 143 (2015) 044108.

[20] H. Persson, P. Thormöhlen, V. P. Zhdanov, B. Kasemo, J. Vac. Sci. Technol. A., 17 (1999) 1721.

[21] M. Stamatakis, D. G. Vlachos, ACS Catal., 2 (2012) 2648–2663.

[22] W. Nielsen, M. d'Avezac, J. Hetherington, M. Stamatakis, J. Chem. Phys., 139 (2013) 224706.

[23] M. Tammaro, M. Sabella, J. W. Evans, J. Chem. Phys., 103 (1995) 10277.

[24] D. T. Gillespie, Physica A., 188 (1992) 404.

[25] M. Pineda, R. Imbihl, L. Schimansky-Geier, C. Zulicke, J. Chem. Phys., 124 (2006) 044701.

[26] M. Stamatakis, D. G. Vlachos, Comput. Chem. Eng., 35 (2011) 2602.

[27] D. T. Gillespie, A. Hellander, L. R. Petzold, J. Chem. Phys., 138 (2013) 170901.

[28] Y. Cao, D. Gillespie, L. Petzold, J. Comput. Phys., 206 (2005) 395.

[29] D. Gillespie, J. Chem. Phys., 113 (2000) 297.

[30] D. J. Liu, J. W. Evans, J. Chem. Phys., 117 (2002) 7319.

[31] C. Barro, Y. D. Decker, T. V. de Bocarmé, P. Gaspard, J. Phys. Chem. Lett. 6 (2015) 2189–2193.

[32] F. Liu, Y. Li, X. Sun, J. Chem. Phys., 140 (2014) 044715.

[33] Y. Go, Z. Hou, H. Xin, J. Phys. Chem. B., 108 (2004) 17796–17799.

[34] D. Schoerr, G. Sanguinetti, R. Grima, J. Chem. Phys., 141 (2014) 024103.

[35] L. Szpruch, D. J. Higham, Multiscale Model. Simul., 8 (2010) 605.

[36] J. Wilkie, Y. M. Wong, Chemical Physics., 353 (2008) 132–138.

[37] S. Dana, S. Raha, J. Comput. Phys., 230 (2011) 8813.

[38] Y. Niu, K. Burrage, L. Chen, J. Ther. Biol., 396 (2016) 90–1704.

[39] C. E. Dangerfield, D. Kay, K. Burrage, Phys. Rev. E., 85 (2012) 051907.

[40] A. V. Skorokhod, Theory Probab Appl., 6 (1961) 264.

[41] A. V. Skorokhod, Theory Probab Appl., 7 (1962) 3.

[42] T. Kawamura, Y. Saisho, Stochastic Models., 22 (2006) 273.

[43] D. R. Brillinger, Lecture Notes-Monograph Series., 41.

[44] D. Lépingle, Mathematics and Computers in Simulation., 38 (1995) 119.

[45] D. J. Higham, SIAM Rev., 43 (2001) 525.

[46] D. Ding, Y. Y. Zhang, Comput. Math. Appl., 55 (2008) 2413.

[47] M. Bossy, B. Jourdain, Ann. Probab., 30 (2002) 1797.

[48] C. G. Morale-Guio, L. A. Stern, X. Hu, Chem. Soc. Rev., 43 (2014) 6555.

[49] C. W. Gardiner, Handbook of Stochastic Methods., Springer, 3rd ed, 2003.

[50] H. Risken, The Fokker-Planck Equation: Methods of Solution and Applications., Springer, Second Edition, 1989.

[51] N. G. van Kampen, Stochastic Processes in Physics and Chemistry., North Holland, Third Edition, 2007.

[52] R. Grima, P. Thomas, A. V. Straube, J. Chem. Phys., 135 (2011) 084103.

[53] H. Grabert, P. Hänggi, I. Oppenheim, Physica A., 117 (1983) 300.

[54] W. Feller, Ann. of Math., 55 (1952) 468.

[55] A. Pilipenko, An Introduction to Stochastic Differential Equations with Reflection., volume 1. Universitätsverlag Potsdam,, 2014.

[56] R. J. Chitashvili, N. L. Lazrieva, Stochastics., 5 (1981) 225–309.

[57] L. Slominski, Stoch. Proc. Appl., 50 (1994) 197–220.

[58] D. Lépingle, C.R.A.S. Paris 316 (1993) 601.

[59] L. Kruk, J. Lehoczky, K. Ramanan, S. Shreve, Ann. Probab., 35 (2007) 1740.

[60] S. Wehner, P. Hoffmann, S. Schmeisser, H. R. Brand, J. Küppers, Phys. Rev. Lett., 95 (2005) 038301.

[61] M. Pineda, R. Toral, J. Chem. Phys., 130 (2009) 124704.

[62] J. Feng, W. Xu, Y. Xu, X. Wang, J. Kurths, Chaos., 27 (2017) 073105.

[63] E. Moro, Phys. Rev. E., 70 (2004) 045102(R).

Supplementary material.
The Euler-Lépingle scheme for the one dimensional case, with a low barrier at $y = 0$

M. Pineda[1] and M. Stamatakis[1]

[1]*Department of Chemical Engineering, University College London, Roberts Building, Torrington Place, London WC1E 7JE, United Kingdom*

If we are interested in the numerical simulation of

$$dy = a(y)dt + \sqrt{b(y)}dW + dk^0, \tag{1}$$

where $y(t) \geq 0$, $k^0(0) = 0$, and

$$k^0(t) = \int_0^t \mathbf{1}_{\{y(s)=0\}} dk^0(s), \tag{2}$$

on the interval $[0, \infty)$, we need to know how to simulate the reflecting process $k^0(t)$. Note that the indicator variable $\mathbf{1}_{\{y(s)=0\}}$ evaluates to 1 if $y(s) = 0$. The process $k^0(t)$ ensures that the modified process $y(t) = x(t) + k^0(t)$ remains in the interval $[0, \infty)$. For this one-dimensional case, with a lower barrier at zero, there is an expression for $k^0(t)$ [1–4]:

$$k^0(t) = \sup_{0 \leq s \leq t} \left[-x(s)\right]^+, \tag{3}$$

where $a^+ = a \vee 0$, $\vee$ stands for maximum, sup stands for supremum, and $x(s)$ is the unreflected process governed by

$$dy = a(y)dt + \sqrt{b(y)}dW. \tag{4}$$

That is, the smallest value that would have to be added to the unreflected process to ensure that it remains in $[0, \infty)$ in the time interval $[0, t]$ will be the maximum amount

by which the unreflected process crosses the zero boundary towards negative values up until time t. Using this explicit expression together with the so-called Skorokhod problem, Lépingle derived a numerical scheme based on the well-known Euler scheme to solve Eq. 1, for the simple situation in which a reflective plane exists at 0. For our system, this numerical scheme is outlined below:

$$y_{i+1} = y_i + a(y_i)\Delta t + \sqrt{b(y_i)}\Delta W_i + C_{i+1}, \tag{5}$$

where

$$C_{i+1} = (\Gamma_i^0 - y_i) \vee 0 \tag{6}$$

with

$$\Gamma_i^0 = \sup_{\tau_i \leq s \leq \tau_{i+1}} \{-a(y_i)(s - \tau_i) - \sqrt{b(y_i)}(W(s) - W(\tau_i))\}, \tag{7}$$

and $y_0 = x_0$. The time discretisation is given by

$$0 = \tau_0 < \tau_1 < \cdots < \tau_i < \cdots < \tau_M = T, \tag{8}$$

in the interval $[0, T]$, with $\Delta t = T/M = \tau_{i+1} - \tau_i$, $\tau_i = i\Delta t$, and $\Delta W_i = W(\tau_{i+1}) - W(\tau_i)$. The simulation of this scheme requires the simulation of the pair $(\Delta W_i, \Gamma_i^0)$ at each time step. This scheme is feasible because we already know how to simulate ΔW_i (increments of a Wiener process W), and Lépingle proposed a way to simulate Γ_i^0 based on the following result

Proposition 1 *Let $h = (h_1, \ldots, h_r)$ and d be a real number. We consider the r-dimensional Wiener process $W(t)$ and the random value*

$$\Psi(t) = \sup_{0 \leq \omega \leq t} (c\omega + h.W(\omega)). \tag{9}$$

We also consider the Gaussian random vector $U = (U_1, \ldots, U_r)$ with mean zero and covariance matrix $t\mathbb{I}_r$, where $\mathbb{I}_r$ is the r-dimensional identity matrix, and the exponential random variable ϑ with parameter $(2t)^{-1}$, with U and ϑ being independent.

2

Let

$$Z = \frac{1}{2}\left[ct + h.U + \sqrt{|h|^2\vartheta + (ct + h.U)^2}\right].$$ (10)

Then $(W(t), \Psi(t))$ and (U, Z) have the same distribution or law. See [1,2] for a proof.

Applied to our system, Lépingle proposal gives

$$\Gamma_i^0 = \frac{1}{2}\left(\Lambda_i + \sqrt{b(y_i)\vartheta_i + \Lambda_i^2}\right),$$ (11)

where

$$\Lambda_i = -a(y_i)\Delta t - \sqrt{b(y_i)}\Delta W_i,$$ (12)

where in the proposition t is replaced by $\Delta t = \tau_{i+1} - \tau_i$, $c = a(y_i)$, $h = \sqrt{b(y_i)}$, and $U = \Delta W_i = \mathcal{N}(0, \Delta t)$, where ΔW_i is a Gaussian random variable with mean zero and variance Δt. The new term ϑ_i is an exponential random variable with rate parameter $(2\Delta t)^{-1}$. Therefore, every new step requires the easy simulation of a Gaussian variable ΔW_i and a new independent exponential variable. This numerical method has a strong rate of convergence equal to 1/2 (similar to the EM scheme) [1]

References

[1] D. Lépingle, Mathematics and Computers in Simulation., 38 (1995) 119.

[2] D. Lépingle, C.R.A.S. Paris 316 (1993) 601.

[3] A. V. Skorokhod, Theory Probab Appl., 6 (1961) 264.

[4] A. V. Skorokhod, Theory Probab Appl., 7 (1962) 3.