

# Reaction time for trimolecular reactions in compartment-based reaction-diffusion models

Cite as: J. Chem. Phys. **148**, 204108 (2018); <https://doi.org/10.1063/1.5024927>

Submitted: 05 February 2018 . Accepted: 08 May 2018 . Published Online: 24 May 2018

Fei Li, Minghan Chen, Radek Erban , and Yang Cao 



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

**Communication: Transition-path velocity as an experimental measure of barrier crossing dynamics**

The Journal of Chemical Physics **148**, 201102 (2018); <https://doi.org/10.1063/1.5030427>

**Probing quantum coherence in ultrafast molecular processes: An ab initio approach to open quantum systems**

The Journal of Chemical Physics **148**, 204112 (2018); <https://doi.org/10.1063/1.5022976>

**Coupled-cluster treatment of molecular strong-field ionization**

The Journal of Chemical Physics **148**, 204102 (2018); <https://doi.org/10.1063/1.5028179>

The Journal  
of Chemical Physics

2018 EDITORS' CHOICE

READ NOW!

# Reaction time for trimolecular reactions in compartment-based reaction-diffusion models

Fei Li,<sup>1</sup> Minghan Chen,<sup>1</sup> Radek Erban,<sup>2</sup> and Yang Cao<sup>1,a)</sup>

<sup>1</sup>Department of Computer Science, Virginia Tech, Blacksburg, Virginia 24061, USA

<sup>2</sup>Mathematical Institute, University of Oxford, Oxford OX2 6GG, United Kingdom

(Received 5 February 2018; accepted 8 May 2018; published online 24 May 2018)

Trimolecular reaction models are investigated in the compartment-based (lattice-based) framework for stochastic reaction-diffusion modeling. The formulae for the first collision time and the mean reaction time are derived for the case where three molecules are present in the solution under periodic boundary conditions. For the case of reflecting boundary conditions, similar formulae are obtained using a computer-assisted approach. The accuracy of these formulae is further verified through comparison with numerical results. The presented derivation is based on the first passage time analysis of Montroll [J. Math. Phys. **10**, 753 (1969)]. Montroll's results for two-dimensional lattice-based random walks are adapted and applied to compartment-based models of trimolecular reactions, which are studied in one-dimensional or pseudo one-dimensional domains. *Published by AIP Publishing.*  
<https://doi.org/10.1063/1.5024927>

## I. INTRODUCTION

Trimolecular (or termolecular) reactions are important components of biochemical models which include oscillations,<sup>1</sup> multistable states,<sup>2</sup> and pattern formation.<sup>3</sup> Considering their reactant complexes, trimolecular reactions can be subdivided into the following three forms:



where  $U$ ,  $V$ , and  $W$  are distinct chemical species (reactants) and the symbol  $\emptyset$  denotes products. In what follows, we will assume that product complexes  $\emptyset$  do not include  $U$ ,  $V$ , and  $W$ . Let us denote by  $u$  the concentration of  $U$ . Then the conventional reaction-rate equation for trimolecular reaction (1) can be written as

$$\frac{du}{dt} = -k(t)u^3, \quad (4)$$

where  $k(t)$  denotes the reaction rate constant (in general, time-dependent).<sup>4</sup> Using mass-action kinetics, rate  $k(t)$  is assumed to be constant and Eq. (4) becomes an autonomous ordinary differential equation (ODE) with a cubic nonlinearity on its right-hand side. Cubic nonlinearities significantly enrich the dynamics of ODEs. For example, ODEs describing chemical systems with two chemical species which do not include cubic or higher nonlinearities cannot have any limit cycles.<sup>5</sup> On the other hand, it has been reported that by adding cubic nonlinearities to such systems, one can obtain chemical systems undergoing homoclinic<sup>5</sup> and saddle-node on invariant circle (SNIC) bifurcation,<sup>6</sup> i.e., oscillating solutions are present for some parameter regimes.

Motivated by the developments in systems biology,<sup>7,8</sup> there has been an increased interest in recent years in stochastic methods for simulating chemical reaction networks. Such approaches provide detailed information about individual reaction events. Considering well-mixed reactors, this problem is well understood. The method of choice is the Gillespie algorithm<sup>9</sup> or its equivalent formulations.<sup>10,11</sup> These methods describe stochastic chemical reaction networks as continuous-time discrete-space Markov chains. They are applicable to modeling intracellular reaction networks in relatively small domains which can be considered well-mixed by diffusion. If this assumption is not satisfied, then stochastic simulation algorithms for spatially distributed reaction-diffusion systems have to be applied.<sup>12,13</sup> The most common algorithms for spatial stochastic modeling in systems biology can be classified as either Brownian dynamics (molecular-based)<sup>14,15</sup> or compartment-based (lattice-based) approaches.<sup>16,17</sup> Molecular-based models describe a trajectory of each molecule involved in a reaction network as a standard Brownian motion. This can be justified as an approximation of interactions (non-reactive collisions) with surrounding molecules (heat bath) on sufficiently long time scales.<sup>18,19</sup> It is then often postulated that bimolecular or trimolecular reactions occur (with some probability) if the reactant molecules are sufficiently close.<sup>4,20–22</sup>

Brownian dynamics' treatment of bimolecular reactions is based on the theory of diffusion-limited reactions, which postulates that a bimolecular reaction occurs with a certain probability if two reactants are within the distance  $R$  (reaction radius) from each other. In some models, this probability is chosen to be less than one,<sup>20,23,24</sup> but in the rest of this overview, we focus on the case when the bimolecular reaction occurs with probability one when the distance of two reactants is equal to  $R$ . The properties of this model depend on the physical

<sup>a)</sup>Electronic mail: [ycao@cs.vt.edu](mailto:ycao@cs.vt.edu)

dimension of the reactor. Considering one-dimensional<sup>25</sup> and two-dimensional problems,<sup>26</sup> diffusion-limited reactions lead to mean-field models with time-dependent rate constants which converge to zero for large times. This qualitative property is in one spatial dimension shared by trimolecular reactions. It has been shown that the mean-field model (4) can be obtained with a time-dependent rate constant  $k(t)$  satisfying scenarios with large times,

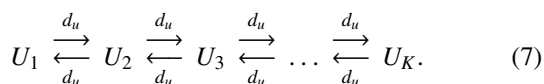
$$k(t) \approx C \sqrt{\frac{\log t}{t}}, \quad (5)$$

where  $C$  is a constant.<sup>4,27</sup> Using (5), we have  $k(t) \rightarrow 0$  as  $t \rightarrow \infty$ , i.e., the rate constant converges to zero for large times. Considering three-dimensional problems, the reaction radius of a diffusion-limited bimolecular reaction can be related with a (time-independent) reaction rate of the corresponding mean-field model.<sup>20</sup> Trimolecular reaction (1) can then be incorporated into three-dimensional Brownian dynamics simulations either directly<sup>22</sup> or as a pair of bimolecular reactions  $U + U \rightarrow Z$  and  $Z + U \rightarrow \emptyset$ , where  $Z$  denotes a dimer of  $U$ .<sup>28</sup>

Compartment-based models divide the simulation domain into compartments (voxels) and describe the state of the system by the numbers of molecules in each compartment. Compartments can be both regular (cubic lattice) or irregular (unstructured meshes).<sup>17</sup> Considering that the simulation domain is divided into cubes of side  $h$ , the diffusive movement of molecules of  $U$  is then modeled as jumps between neighbouring compartments with rate  $d_u = D_u/h^2$ , where  $D_u$  is the diffusion constant of the chemical species  $U$ . In this paper, we will consider compartment-based stochastic reaction-diffusion models of the trimolecular reactions (1)–(3) in a narrow domain  $[0, L] \times [0, h] \times [0, h]$  where  $L \gg h$ . Such domains are useful for modeling filopodia,<sup>29,30</sup> but they can also be viewed as a simplification of a real three-dimensional geometry when there is no variation in the  $y$  and  $z$  directions. Then the mean-field model of trimolecular reaction (1) can be formulated in terms of the spatially varying concentration  $u(x, t)$ , where  $x \in [0, L]$  and  $t \geq 0$ , which satisfies the partial differential equation (PDE)

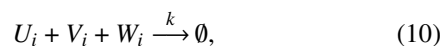
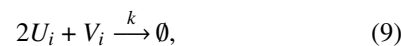
$$\frac{\partial u}{\partial t} = D_u \frac{\partial^2 u}{\partial x^2} - k u^3, \quad (6)$$

where  $k$  is the macroscopic rate constant of the trimolecular reaction. To formulate the compartment-based stochastic reaction-diffusion model, we divide the domain  $[0, L] \times [0, h] \times [0, h]$  into  $K = L/h$  cubic compartments  $[(i-1)h, ih] \times [0, h] \times [0, h]$ ,  $i = 1, 2, \dots, K$ . Denoting  $U_i$  as the number of molecules of  $U$  in the  $i$ th compartment, the diffusion of  $U$  can be written as the chain of “chemical reactions”<sup>12</sup>



Diffusion of  $V$  and  $W$ , which appear in trimolecular reactions (2) or (3), is described using the chains of “chemical reactions” of the form (7) with the jump rates given by  $d_v = D_v/h^2$  and  $d_w = D_w/h^2$ , where  $D_v$  and  $D_w$  are the diffusion constants of the chemical species  $V$  and  $W$ , respectively.

Trimolecular chemical reactions are localized in compartments, i.e., each of trimolecular reactions (1)–(3) is replaced by  $K$  reactions



where  $i = 1, 2, \dots, K$  and  $k$  denotes the macroscopic reaction rate constant with units  $[\text{m}^6 \text{s}^{-1}]$ . Compartment-based modeling postulates that each compartment is well-mixed. In particular, chemical reactions (7)–(10) can be all simulated using the Gillespie algorithm<sup>9</sup> (or other algorithms for well-mixed chemical systems<sup>10,11</sup>) and the system can be equivalently described using the reaction-diffusion master equation (RDME).<sup>12</sup> In particular, the probability that a trimolecular reaction occurs in time interval  $[t, t + \Delta t]$  in a compartment containing one triplet of reactants is equal to  $\alpha_1 \Delta t$ . Since the volume of one compartment is  $h^3$  and  $k$  has physical units  $[\text{m}^6 \text{s}^{-1}]$ , the rate of reaction per compartment is

$$\alpha_1 = \frac{k}{h^6} \quad (11)$$

and  $\Delta t$  is chosen sufficiently small so that  $\alpha_1 \Delta t \ll 1$ . The standard scaling of reaction rates (11) is considered in this paper when we investigate the dependence of trimolecular reactions on  $h$ . It has been previously reported for bimolecular reactions that the standard RDME scaling leads to large errors of bimolecular reactions.<sup>20</sup> One of the goals of the presented manuscript is to investigate the effect of  $h$  on trimolecular reactions.

The diffusive chain of reactions (7) is formulated using a narrow three-dimensional domain  $[0, L] \times [0, h] \times [0, h]$ . It can also be interpreted as a purely one-dimensional simulation. In this case, the simulated domain is a one-dimensional interval  $[0, L]$ , divided into  $K = L/h$  intervals  $[(i-1)h, ih]$ ,  $i = 1, 2, \dots, K$ , where  $U_i$  is the number of molecules of  $U$  in the  $i$ th interval. Trimolecular reactions (8)–(10) are then written with a one-dimensional rate constant  $k_{1D}$  which has physical units  $[\text{m}^2 \text{s}^{-1}]$ . Then the probability that a trimolecular reaction occurs in time interval  $[t, t + \Delta t]$  in an interval containing one triplet of reactants is equal to  $\alpha_2 \Delta t$  where

$$\alpha_2 = \frac{k_{1D}}{h^2} \quad (12)$$

and  $\Delta t$  is chosen sufficiently small so that  $\alpha_2 \Delta t \ll 1$ . In particular, two interpretations of the simulated diffusive reaction chain (7) considered in this paper differ by physical units of  $k$  and  $k_{1D}$  and by the corresponding scaling of the reaction rate with  $h$ : compare Eqs. (11) and (12). The simulated model can also be used to verify the corresponding Brownian dynamics result (5), provided that we appropriately relate  $h$  and  $R$  and postulate that the trimolecular reaction occurs (for sure) whenever a compartment contains one triplet of reactants.<sup>4</sup>

The rest of the paper is organized as follows. In Sec. II, we summarize recent results for bimolecular reactions in

two spatial dimensions for periodic boundary conditions and generalize them to reflecting (no-flux) boundary conditions. These results are useful for investigating compartment-based stochastic reaction-diffusion modeling of trimolecular reactions (1)–(3) as we will show in Sec. III. Computational experiments illustrating the presented results are given in Sec. IV.

## II. BIMOLECULAR REACTIONS IN TWO DIMENSIONS

In this paper, we focus on a special case where there is only one combination of reactants present in the system and we study the firing times of trimolecular reactions (1)–(3). A similar problem for a bimolecular reaction  $U + V \rightarrow \emptyset$  with one  $U$  molecule and one  $V$  molecule is investigated by Hellander *et al.*<sup>31</sup> for both two-dimensional and three-dimensional compartment-based models. Dividing the two-dimensional domain  $[0, L] \times [0, L]$  into square compartments of size  $h$ , the mean time until the two molecules react is<sup>31</sup>

$$\tau_{bimol} = \frac{h^2(1 + N_{steps}^1)}{k_b} + \frac{h^2 N_{steps}}{4(D_u + D_v)}, \quad (13)$$

where  $k_b$  is the bimolecular reaction rate with units  $[m^2 s^{-1}]$  and  $N_{steps}^1$  is the mean number of diffusive jumps until  $U$  and  $V$  are in the same compartment, given that they are initially one compartment apart.  $N_{steps}$  is the mean number of steps, until  $U$  diffuses to  $V$ 's location for the first time. The quantities  $N_{steps}$  and  $N_{steps}^1$  can be estimated using the following theorem.<sup>32</sup>

**Theorem 1** (Montroll, 1969). *Assume that the molecule  $U$  has a random, uniformly distributed, starting position on a finite two-dimensional square lattice with periodic boundary conditions. Then the following holds:*

$$N_{steps} = \pi^{-1} N \log(N) + 0.1951N + O(1), \quad (14)$$

where  $N = L^2/h^2$  is the number of lattice points (compartments) in the domain. Furthermore,  $N_{steps}^1 = N - 1$ .

By Theorem 1 and formula (13), we have<sup>31</sup>

$$\tau_{bimol} = \frac{L^2}{k_b} + \tau_{coll}, \quad (15)$$

where  $\tau_{coll}$  is the mean time for the first collision of molecules  $U$  and  $V$ , which can be, for two-dimensional bimolecular reactions with periodic boundary conditions, approximated by

$$\tau_{coll} \approx \frac{L^2 \log(L h^{-1})}{2\pi(D_u + D_v)} + 4.878 \times 10^{-2} \frac{L^2}{D_u + D_v}, \quad (16)$$

as  $h \rightarrow 0$ . Using formulae (15) and (16), we see that the reaction time for the bimolecular reaction  $U + V \rightarrow \emptyset$  tends to infinity when the compartment size  $h$  tends to zero.<sup>31</sup> In particular, formulae (15) and (16) imply that the number of bimolecular reaction firings will tend to zero in RDME simulations when  $h$  tends to zero. Thus there has to be a lower bound for the compartment size. This has also been shown using different methods in three-dimensional systems, and

improvements of algorithms for  $h$  close to the lower bound have been derived.<sup>20,31</sup>

Formulae (15) and (16) give a good approximation to the mean reaction time of the bimolecular reaction for two-dimensional domains, provided that periodic boundary conditions are used, as assumed in Theorem 1. However, the chain of chemical reactions (7) implicitly assumes reflecting boundary conditions. Such boundary conditions are commonly used in biological systems whenever the boundary of the computational domain corresponds to a physical boundary (e.g., cell membrane) in the modeled system.<sup>13</sup> In Fig. 1, we show that formula (16) is not an accurate approximation to the mean collision time for reflecting boundary conditions. Reflecting boundary conditions mean that a molecule remains in the same lattice point when it hits the boundary. We plot results for  $D_u = D_v$ , for  $D_u > D_v$ , and for  $D_u < D_v$  in Fig. 1. In each case, we see that formula (16) matches well with numerical results for periodic boundary conditions. In order to find a formula that matches with numerical experiment results for reflecting boundary conditions, we fix the coefficient of the first term in (16) and perform data fitting on the second coefficient. We obtain the following formula for bimolecular reactions with reflecting boundary conditions in two-dimensional compartment-based framework:

$$\tau_{coll} \approx \frac{L^2 \log(L h^{-1})}{2\pi(D_u + D_v)} + 0.2742 \frac{L^2}{D_u + D_v}. \quad (17)$$

The average times for the first collision of  $U$  and  $V$  given by (17) are plotted in Fig. 1 corresponding to different sets of diffusion rates. It can be seen that formula (17) matches well with numerical experiment results with reflecting boundary conditions. In Fig. 2, we show the mean collision time as a function of  $D_u$  based on both formula (17) and the numerical experiment data with reflecting boundary conditions.

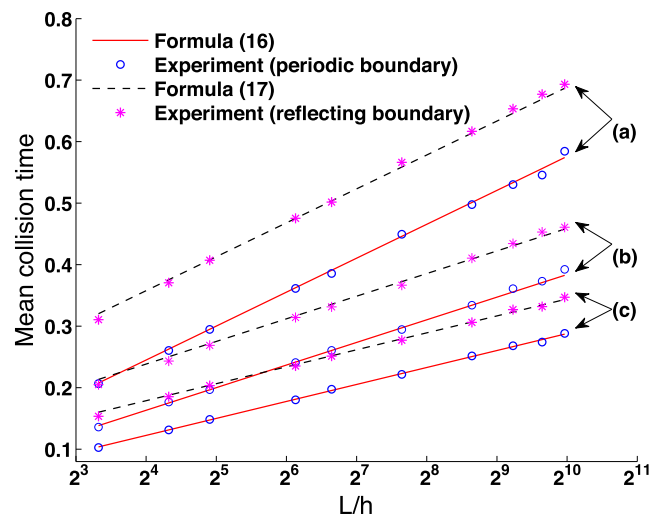


FIG. 1. Comparison of the original formula (16) and our revised formula (17) with the mean collision time from numerical simulations with periodic and reflecting boundary conditions. The two molecules diffuse freely in the 2D domain, with the initial position uniformly distributed in the computational domain. Three parameter sets: (a)  $D_u = 1$ ,  $D_v = 1$ ; (b)  $D_u = 2$ ,  $D_v = 1$ ; (c)  $D_u = 1$ ,  $D_v = 3$ .

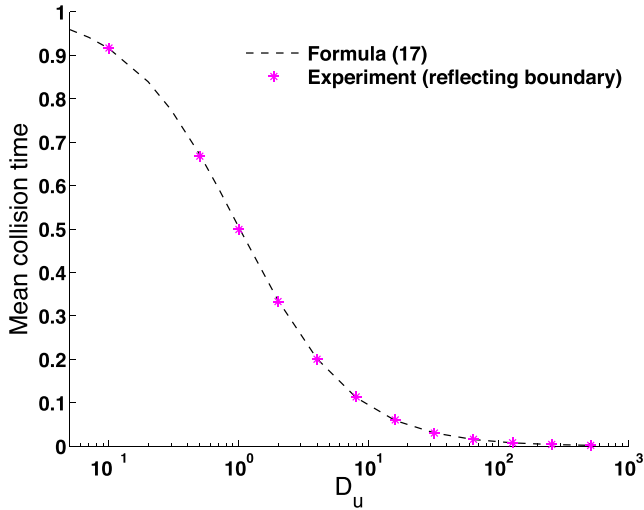


FIG. 2. Mean collision times with reflecting boundary conditions corresponding to different  $D_u$  values when  $D_v = 1$  is fixed. We use  $L = 1.0$  and  $h = 0.01$ .

### III. MEAN REACTION TIME FOR TRIMOLECULAR REACTION

In this paper, we will first consider periodic boundary conditions and generalize formula (15) to all cases of trimolecular reactions (1)–(3) in one spatial dimension, using both Eqs. (11) and (12) to scale reaction rates.

Both trimolecular reactions (1) and (2) are special cases of (3). Since we will focus on the simplified situation where there is only one molecule for each reactant of (3), we may consider (1) as the special case of (3) where diffusion rates for all three molecules are the same and (2) as the special case where at least two of the three molecules have the same diffusion rates. We denote the value of the mean reaction time by  $\tau_{\text{trimol}}$ . We will decompose  $\tau_{\text{trimol}}$  into two parts,

$$\tau_{\text{trimol}} = \tau_{\text{react}} + \tau_{\text{coll}}, \quad (18)$$

where  $\tau_{\text{react}}$  gives the mean time for reaction given that the molecules are initially located in the same compartment and  $\tau_{\text{coll}}$  is the mean time for the first collision, i.e., the average time for the system to arrive at a state where all molecules are in the same compartment, given that they were initially uniformly distributed. Note that the bimolecular formula (15) was written in the form of a similar decomposition like (18). We will call  $\tau_{\text{coll}}$  a collision time.

#### A. Special case where $D_u = D_v$ and $D_w = 0$

We will start with a simple case with  $D_u = D_v$  and  $D_w = 0$ . Then the only  $W$  molecule will be fixed at its initial location. Under the periodic boundary condition assumption, without loss of generality, we assume that this  $W$  molecule is located at the center of the interval  $[0, L]$  or, specifically, in the compartment  $\left[\frac{K+1}{2}\right]$ , where  $[x]$  represents the largest integer that is smaller than  $x$ . The  $U$  and  $V$  molecules diffuse according to (7) from their initial compartments. The reaction can fire only when both  $U$  and  $V$  molecules jump to the center since the

only  $W$  molecule is at the center. Let  $\tau_{\text{coll}}$  be the mean time for these two molecules ( $U$  and  $V$ ) to be both at the center for the first time, which requires that their compartment number is equal to  $\left[\frac{K+1}{2}\right]$ .

Instead of trying to develop a formula for the collision time of two molecules diffusing to a fixed compartment in the one-dimensional lattice, we consider an equivalent problem: imagine a  $Z$  molecule jumps with a diffusion rate  $D_u$  within a  $K \times K$  grid in the two-dimensional space and let its compartment index be  $z = (x_1, x_2)$ . Then the two independent random walks by the two  $U$  and  $V$  molecules in the one-dimensional lattice can be viewed equivalently as the random walk of the  $Z$  molecule in the two-dimensional square lattice with the diffusion rate  $D_u$ . Collision time  $\tau_{\text{coll}}$  is then equal to the mean time for the pseudo molecule  $Z$  to jump to the center for the first time, which is the case discussed in Sec. II. Therefore, formula (15) can also be applied to the trimolecular reaction (3) when  $D_w = 0$  and  $D_u = D_v$  and periodic boundary conditions are considered.

#### B. Special case where $D_u \neq D_v$ and $D_w = 0$

Formula (15) cannot be directly applied to (3) if  $D_u \neq D_v$  even if  $D_w = 0$ . We can still assume  $W$  is in the center and consider the equivalent problem of a  $Z$  molecule jumping with a diffusion rate  $D_u$  in the  $x$  axis and  $D_v$  in the  $y$  axis within a  $K \times K$  grid in the 2D space. The two independent random walks by the two  $U$  and  $V$  molecules in 1D space can be viewed equivalently as the random walk of the  $Z$  molecule in the 2D space with anisotropic diffusion rates, given by  $D_u$  and  $D_v$  in different directions. In Appendix A, we prove the following theorem:

**Theorem 2.** Assume that the molecule  $Z$  has a random, uniformly distributed, starting position in a 2D lattice and that it can move from the current compartment to its nearest neighbours only. Assume  $Z$  diffuses with the diffusion rate  $D_u$  in the  $x$  direction,  $D_v$  in the  $y$  direction, and  $D_v \geq D_u$ . Then the following holds:

$$\tau_{\text{coll}} = \frac{L^2}{2\pi\sqrt{D_u D_v}} \log\left(\frac{L}{h}\right) + \frac{L^2}{12D_v} + \frac{L^2}{4\pi\sqrt{D_u D_v}} \left[ 2\left(\gamma + \log\left(\frac{2}{\pi}\right)\right) - \log\left(1 + \frac{D_u}{D_v}\right) \right], \quad (19)$$

where

$$\gamma = \lim\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n\right) \approx 0.5772. \quad (20)$$

#### C. Collision time for the general trimolecular reaction

In this subsection, we consider the trimolecular reaction (3) with corresponding diffusion rates  $D_u$ ,  $D_v$ , and  $D_w$ . Without loss of generality, we assume  $D_u \geq D_v \geq D_w > 0$ .

We consider one pseudo molecule  $Z = (z_1, z_2)$ , where  $z_1$  and  $z_2$  are expressed in terms of locations  $x_u$ ,  $x_v$ , and  $x_w$  of three molecules by

$$z_1 = x_u - x_w, \quad \text{and} \quad z_2 = x_v - x_w. \quad (21)$$



When  $U$ ,  $V$ , and  $W$  diffuse with rates  $D_u$ ,  $D_v$ , and  $D_w$ , the pseudo molecule  $Z$  jumps on the 2D lattice corresponding to (21). When  $Z$  jumps to the origin,  $U$ ,  $V$ , and  $W$  will be in the same grid, and vice versa. Thus the collision time of the trimolecular reaction in 1D is again converted to the corresponding collision time of the bimolecular problem in 2D. The actual grid and jumps are illustrated in Fig. 3.

The difficulty of the mapping (21) is that the resulting domain is not square. So we cannot apply the theoretical results presented in Appendix A, which is for square domain cases. In order to apply an estimation on a square lattice, we need to further modify the mapping (21). We will take

$$z_1 = |x_u - x_w| \quad \text{and} \quad z_2 = |x_v - x_w|. \quad (22)$$

Then the molecule  $Z$  with coordinates  $(z_1, z_2)$  will jump in a  $K \times K$  two-dimensional square lattice, where  $K = L/h$ . Of course, the jumps at the boundary will be different from the illustration in Fig. 3, but that is just symmetric reflection. It does not change the validity of the formula derived in Appendix A, which is based on the assumption of periodic lattices. The domain resulting from the mapping (22) is shown in Fig. 4. For a trimolecular reaction in 1D domain with periodic boundary conditions, we have the following approximation formula for the mean jump time:

$$\tau_{coll} = \frac{L^2}{2\pi\hat{D}} \log\left(\frac{L}{h}\right) + \frac{L^2}{12(D_v + D_w)} + \frac{L^2}{4\pi\hat{D}} \left[ 2\left(\gamma + \log\left(\frac{2}{\pi}\right)\right) - \log(1 + \eta') \right], \quad (23)$$

where

$$\hat{D} = \sqrt{D_u D_v + D_u D_w + D_v D_w}, \quad (24)$$

and

$$\eta' = \frac{D_u^2}{\hat{D}^2}. \quad (25)$$

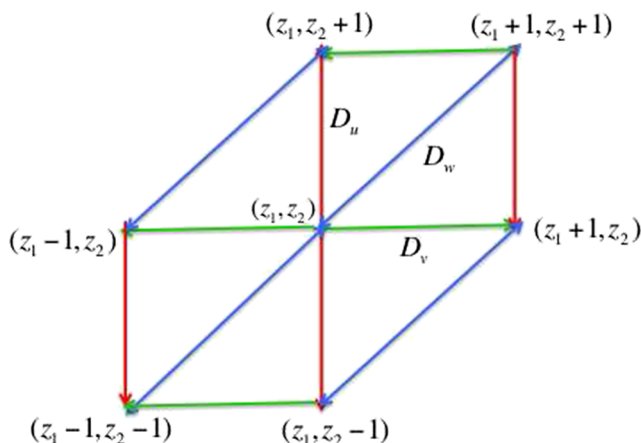


FIG. 3. Illustration of the 2D structure for the jumps of  $Z$  based on (21). When the  $U$  molecule jumps, the corresponding pseudo-molecule  $Z$  jumps (shown in red) in the direction  $(z_1, z_2) \rightarrow (z_1 \pm 1, z_2)$ . When the  $V$  molecule jumps, the corresponding pseudo-molecule  $Z$  jumps (shown in green) in the direction  $(z_1, z_2) \rightarrow (z_1, z_2 \pm 1)$ ; when the  $W$  molecule jumps, the corresponding pseudo-molecule  $Z$  jumps (shown in blue) in the direction  $(z_1, z_2) \rightarrow (z_1 \pm 1, z_2 \pm 1)$ . The whole domain has a similar shape and is not square.

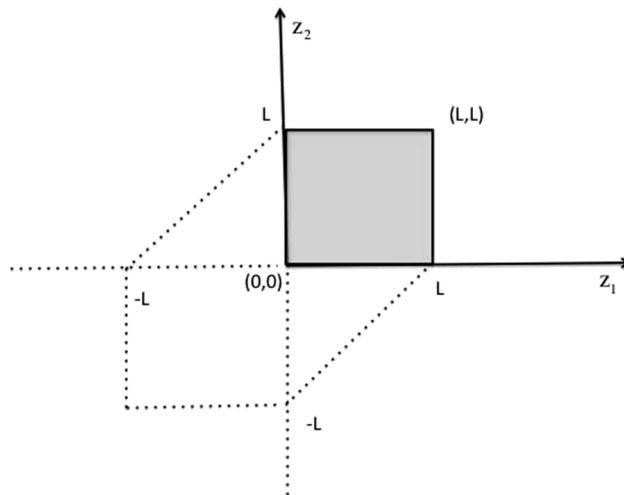


FIG. 4. Illustration of the 2D square domain resulting from (22). The original domain (dotted line) resulting from (21) is not square. With (22), the shaded domain is a square and the theoretical analysis in Appendix A can be applied.

*Remark.* Formula (23) also implies an estimation of the collision time of a bimolecular reaction in a 1D domain. In (23), if we let  $D_u \rightarrow \infty$ , the formula leads to an estimation of the mean collision time of a bimolecular reaction  $V + W \rightarrow \emptyset$  in a 1D domain  $[0, L]$  with periodic boundary conditions, which is independent of  $h$ :

$$\tau_{coll} = \frac{L^2}{12(D_v + D_w)}. \quad (26)$$

#### D. Formula under reflecting boundary conditions

As shown in Sec. IV, formula (23) matches very well with the result computed by stochastic simulations for the periodic boundary condition. However, when we consider a reflecting boundary condition, we see a mismatch. In order to find a formula that matches the reflecting boundary condition results, we performed numerical experiments and collected the mean collision time with reflecting boundary conditions. Then with data fitting, we obtain the formula that matches the numerical results corresponding to the reflecting boundary condition—see Appendix B for derivation of the coefficient of 0.140. Please note that although we derived the value of the coefficient, the collision time formula for the reflecting boundary condition is still based on data fitting with a function similar to Eq. (23) for the periodic boundary condition. Formula (27) gives an estimation of the mean first collision time,

$$\tau_{coll} = \frac{L^2}{2\pi\hat{D}} \log\left(\frac{L}{h}\right) + 0.140 \frac{L^2}{D_v + D_w} + \frac{L^2}{4\pi\hat{D}} \left[ 2\left(\gamma + \log\left(\frac{2}{\pi}\right)\right) - \log\left(0.125 + \frac{\eta'}{4}\right) \right], \quad (27)$$

where  $\hat{D}$  is defined in Eq. (24) and  $\eta'$  is defined in Eq. (25).

*Remark.* When  $D_u \rightarrow \infty$ , the collision time is

$$\tau_{coll} = 0.140 \frac{L^2}{D_v + D_w}. \quad (28)$$

### E. Mean reaction time

Formula (23) gives the estimation for the mean collision time  $\tau_{coll}$  of trimolecular reactions. For the first reaction time  $\tau_{trimol}$ , since our derivation is based on the corresponding analysis on 2D grids, an estimation similar to formula (13) can be applied. We have

$$\tau_{trimol} \approx \frac{h^2(1 + N_{steps}^1)}{k_{1D}} + \tau_{coll}, \quad (29)$$

as  $h \rightarrow 0$ , where  $k_{1D}$  is the reaction rate for the trimolecular reaction as defined in (12).

For biochemical reactions with reflecting boundary conditions, we have the following formula for the corresponding mean reaction time:

$$\tau_{trimol} = \frac{L^2}{k_{1D}} + \frac{L^2}{2\pi\tilde{D}} \log\left(\frac{L}{h}\right) + 0.140 \frac{L^2}{D_v + D_w} + \frac{L^2}{4\pi\tilde{D}} \left[ 2\left(\gamma + \log\left(\frac{2}{\pi}\right)\right) - \log\left(0.125 + \frac{\eta'}{4}\right) \right]. \quad (30)$$

Correspondingly, if we use the scaling in (11), (30) takes the form

$$\tau_{trimol} = \frac{L^2 h^4}{k} + \frac{L^2}{2\pi\tilde{D}} \log\left(\frac{L}{h}\right) + 0.140 \frac{L^2}{D_v + D_w} + \frac{L^2}{4\pi\tilde{D}} \left[ 2\left(\gamma + \log\left(\frac{2}{\pi}\right)\right) - \log\left(0.125 + \frac{\eta'}{4}\right) \right]. \quad (31)$$

We can see that as  $h \rightarrow 0$ , the reaction time will be dominated by the collision time in either case.

### IV. NUMERICAL RESULTS

We test the collision time of the general trimolecular reaction (3). Figure 5 shows the comparison of the numerical

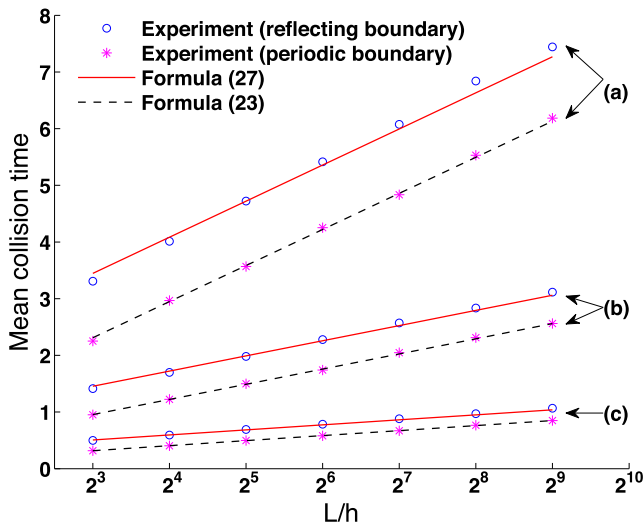


FIG. 5. The mean first collision time of three molecules in the 1D domain (for periodic and reflecting boundary conditions). The solid lines plot the mean collision time with periodic and reflecting boundary conditions corresponding to formula (23) and formula (27), while dashed lines show the numerical results by stochastic simulations. For the same color from top to bottom, the three sets of diffusion rate parameters are (a)  $D_u = D_v = D_w = 0.1$ ; (b)  $D_u = 0.5$ ,  $D_v = 0.2$ ,  $D_w = 0.1$ , and (c)  $D_u = 2.5$ ,  $D_v = 0.5$ ,  $D_w = 0.1$ .

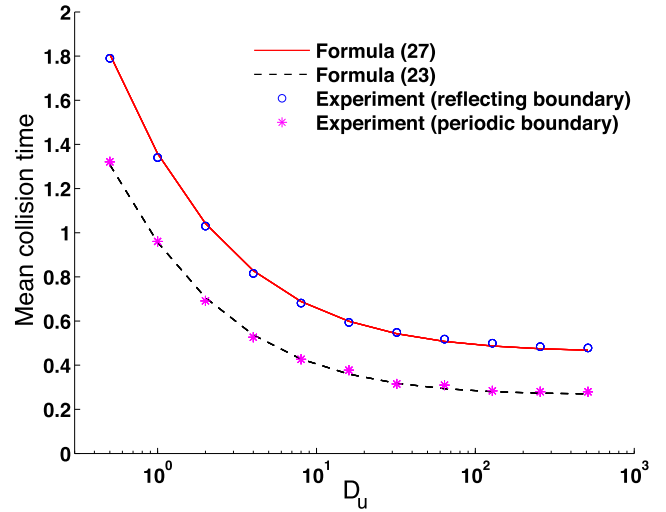


FIG. 6. The mean trimolecular collision time (for periodic and reflecting boundary conditions) against the diffusion rate  $D_u$ . Other parameters are given as  $D_v = 0.2$ ,  $D_w = 0.1$ , and  $N = 20$ . The 1D domain size is fixed at  $L = 1.0$ .

results of the mean first collision time with periodic and reflecting boundary conditions with the two formulae (23) and (27). Figure 5 demonstrates that formula (23) matches well with computational results obtained by stochastic simulations corresponding to periodic boundary conditions, justifying our analysis. It also shows that formula (27) matches well with simulation results corresponding to reflecting boundary conditions. Figure 6 presents the numerical results of the mean first collision time with periodic and reflecting boundary conditions corresponding to different  $D_u$  when we fix  $L$ ,  $D_v$ , and  $D_w$ . We can see that the two formulas match with the numerical results very well. We also present another comparison and analysis for the special case when  $D_u \rightarrow \infty$  in Appendix B.

For the mean reaction time comparison, we focus only on reflecting boundary conditions. In Fig. 7, we show the comparison of the mean reaction time with formula (30) corresponding

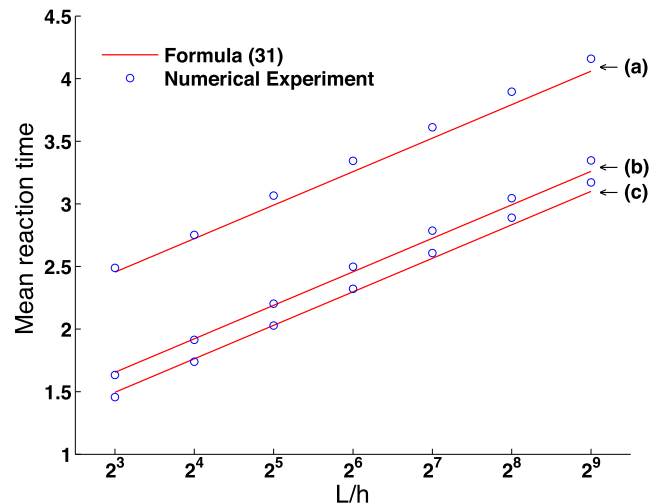


FIG. 7. The mean trimolecular reaction time for reflecting boundary conditions against  $\log(L/h)$ . The parameters for the plot are  $L = 1.0$ ,  $D_u = 0.5$ ,  $D_v = 0.2$ ,  $D_w = 0.1$ . For the same color, the reaction rates from top to bottom are (a)  $k_{1D} = 1.0$ , (b)  $k_{1D} = 5.0$ , (c)  $k_{1D} = 25.0$ , respectively.

to a fixed set of diffusion rates and different reaction rates  $k_{1D}$ .

## V. DISCUSSION

The important consequence of formula (23) and formula (27) is that when  $h \rightarrow 0$ ,  $\tau_{coll} \rightarrow \infty$ . Since the trimolecular reaction time  $\tau_{trimol} > \tau_{coll}$ , the reaction time also tends to infinity. In particular, our analysis confirms the observations made in Refs. 4 and 27 that the diffusion-limited three-body reactions do not recover mean-field mass-action results in 1D. The rate constant in (4) for such problems is time dependent, given by (5), and converges to zero as  $t \rightarrow \infty$ . In the limit  $h \rightarrow 0$ , the average collision time goes to infinity and the rate constant  $k(t)$  in (4) converges to zero for any finite time. It is given by (5), where  $C \equiv C(h)$  converges to zero as  $h \rightarrow 0$ . In recent years, this topic has been discussed for the bimolecular reaction in the 2D and 3D cases<sup>20,31,33</sup> as well. This limitation is a great challenge for spatial stochastic simulation. The work presented in this article provides an analysis of bimolecular and trimolecular reactions in that framework. Further work may be continued to derive an appropriate formula for the reaction rate of trimolecular reactions in compartment-based stochastic modeling and simulations. Moreover, in the current work, we do not have a satisfying theoretical analysis on the reaction time under reflecting boundary conditions. Formula (27) is derived based on an analysis of the collision time for bimolecular reactions in 1D and then data fitting. Further work to improve this analysis is expected.

## ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Research Council under the *European Community's* Seventh Framework Programme (No. FP7/2007-2013)/ERC Grant Agreement No. 239870. R. Erban would like to thank the Royal Society for a University Research Fellowship and the Leverhulme Trust for a Philip Leverhulme Prize. This work was carried out during the visit of R. Erban to the Isaac Newton Institute. This work was supported by EPSRC Grant No. EP/K032208/1. This work was partially supported by a grant from the Simons Foundation. Y. Cao's work is also partially supported by NSF Grant Nos. MCB-1613741 and CCF-1526666.

## APPENDIX A: FORMULA OF NONUNIFORM RANDOM WALK ON A 2D LATTICE

We consider the random walk problem on a square 2D lattice. The estimation procedure we present here is generalized from the original idea in Montroll.<sup>32</sup> Following the notations in Montroll,<sup>32</sup> let  $F_n(s)$  be the probability that a lattice walker starting from the origin arrives at the lattice point  $s$  for the first time at the  $n$ th step. Let  $P_n(s)$  be the probability that a walker starting from the origin arrives at  $s$  at the  $n$ th step (no matter how many times it may have visited the state  $s$  in the previous  $n - 1$  steps). The generating functions of these two probabilities are

$$F(s, z) = \sum_{n=1}^{\infty} z^n F_n(s)$$

and

$$P(s, z) = \sum_{n=0}^{\infty} z^n P_n(s),$$

respectively. Montroll has proved the relationship of the two generating functions

$$F(s, z) = [P(s, z) - \delta_{s,0}]/P(0, z),$$

where  $\delta_{s,0} = P_0(s)$  is the initial condition that the walker starts from the origin  $s = 0$ , represented as the Dirac delta function.

Assume that the origin point is a trapping point. As the walker has the same probability of being at any of  $N - 1$  nontrapping point at time  $t = 0$ , the generating function, for the probability that the walker is trapped at the origin after the  $n$ th step, is

$$\begin{aligned} G_s(z) &= \frac{1}{N-1} \sum_{s \neq 0} F(s, z) \\ &= \frac{1}{N-1} \left\{ \sum_s \frac{P(s, z) - \delta_{s,0}}{P(0, z)} - \frac{P(0, z) - \delta_{0,0}}{P(0, z)} \right\} \\ &= \frac{1}{N-1} \left\{ -1 + \sum_s \frac{P(s, z)}{P(0, z)} \right\} \\ &= \frac{1}{N-1} \left\{ -1 + \frac{\sum_n z^n \sum_s P_n(s)}{P(0, z)} \right\} \\ &= \frac{1}{N-1} \left\{ -1 + \frac{1}{(1-z)P(0, z)} \right\}, \end{aligned}$$

where  $N = m^2$  (with  $m = \frac{L}{h}$ ) is the total number of grids in the square 2D lattice. Note that  $\sum_s P_n(s) = 1$ . The average number of steps required to reach the origin for the first time is<sup>32</sup>

$$\begin{aligned} \langle n \rangle &= \partial G_s / \partial z|_{z=1} \\ &= \frac{1}{N-1} \frac{\partial}{\partial z} \left\{ \frac{1}{(1-z)P(0, z)} \right\}_{z=1}. \end{aligned} \quad (A1)$$

In order to find the formula for  $P(0, z)$ , define the structure function as

$$\lambda(\theta) = \sum_s p(s) \exp(i\theta s),$$

where  $p(s)$  is the probability that at any step a random walker makes a displacement  $s$  and  $\sum_s p(s) = 1$ . In this 2D square lattice, we consider a general problem. Suppose the random walk can jump left and right with the rate  $D_u/h^2$ , up and down with the rate  $D_v/h^2$ , and diagonally with the rate  $D_w/h^2$ . Let  $\sigma_1^2 = \frac{D_u}{D_u+D_v+D_w}$ ,  $\sigma_2^2 = \frac{D_v}{D_u+D_v+D_w}$ , and  $\sigma_3^2 = \frac{D_w}{D_u+D_v+D_w}$ . We thus have

$$\sigma_1^2 + \sigma_2^2 + \sigma_3^2 = 1.$$

There are some special cases. In a uniform diffusion case,  $\sigma_1 = \sigma_2 = \sigma_3 = \sqrt{3}/3$ . In the case  $D_w = 0$ ,  $\sigma_3 = 0$  and then when  $\sigma_1 = \sigma_2$ , that will be the simple case discussed in Montroll.<sup>32</sup> Here we will generalize Montroll's work and derive a general formula and then discuss different special cases. For the probability  $p(s)$ , we have  $p(1, 0) = p(-1, 0)$



$= \sigma_1^2/2$  and  $p(0, 1) = p(0, -1) = \sigma_2^2/2$ ,  $p(1, 1) = p(-1, -1) = \sigma_3^2/2$ , and

$$\lambda(\theta) = c_1\sigma_1^2 + c_2\sigma_2^2 + \sigma_3^2(c_1c_2 - s_1s_2),$$

where  $\theta = (\theta_1, \theta_2)$ ,  $c_i = \cos \theta_i$ , and  $s_i = \sin \theta_i$ . Then

$$P(0, z) = m^{-2} \sum_{k_1=0}^{m-1} \sum_{k_2=0}^{m-1} [1 - z\lambda(2\pi k_1/m, 2\pi k_2/m)]^{-1}.$$

Since the above calculation involves two sequential summations, we want to approximate  $P(0, z)$  for simplification. By Montroll's work,<sup>32</sup>

$$P(0, z) = \frac{1}{m} \sum_{k_1=0}^{m-1} [1 - c_1\sigma_1^2 z]^{-1} f(z, \theta_1),$$

where

$$f(z, \theta_1) = \frac{1}{m} \sum_{k_2=0}^{m-1} [1 - c_2w_1 - w_2s_2]^{-1},$$

with

$$w_1 = \frac{z(\sigma_2^2 + c_1\sigma_3^2)}{1 - zc_1\sigma_1^2}$$

and

$$w_2 = \frac{z(\sigma_3^2s_1)}{1 - zc_1\sigma_1^2}.$$

Following Appendix A of Montroll,<sup>32</sup> we define  $w_1 + iw_2 = \rho e^{i\phi}$ . Thus

$$\rho^2 = w_1^2 + w_2^2.$$

Then [Eq. (A7) in Montroll<sup>32</sup>]

$$f(z, \theta_1) = \frac{1}{\sqrt{1 - \rho^2}} \frac{1 - x^{2m}}{1 - 2x^m \cos m\phi + x^{2m}},$$

where  $x$  is the smaller root of the equation

$$x^2 - 2x/\rho + 1 = 0. \quad (\text{A2})$$

Now we rewrite  $P(0, z)$  as

$$P(0, z) = \frac{1}{m} \sum_{k_1=0}^{m-1} \left\{ \frac{(1 - zc_1\sigma_1^2)^{-1}}{\sqrt{1 - \rho^2}} \frac{1 - x^{2m}}{1 - 2x^m \cos m\phi + x^{2m}} \right\}. \quad (\text{A3})$$

The simplification of (A3) can be achieved in several steps. First, when  $k_1 = 0$ ,  $c_1 = 1$ ,  $s_1 = 0$ ,  $w_2 = 0$ , and  $\phi = 0$ , then

$$\rho = w_1 = z(1 - \sigma_1^2 z)^{-1}(\sigma_2^2 + \sigma_3^2) = \frac{(1 - \sigma_1^2)z}{1 - \sigma_1^2 z}.$$

The corresponding term in  $P(0, z)$  is given by

$$\begin{aligned} & \frac{1}{m} (1 - z\sigma_1^2)^{-1} \frac{1 + x^m}{1 - x^m} \left\{ 1 - \left( \frac{z(1 - \sigma_1^2)}{1 - z\sigma_1^2} \right)^2 \right\}^{-\frac{1}{2}} \\ &= \frac{1}{m} \frac{1 + x^m}{1 - x^m} \left\{ (1 - z\sigma_1^2)^2 - z^2(1 - \sigma_1^2)^2 \right\}^{-\frac{1}{2}} \\ &= \frac{1}{m} \frac{1 + x^m}{1 - x^m} (1 - z)^{-\frac{1}{2}} \left\{ 1 + z - 2z\sigma_1^2 \right\}^{-\frac{1}{2}} \\ &= \frac{1}{m} \frac{1 + x^m}{1 - x^m} \frac{1}{\alpha} \left\{ 2(1 - \sigma_1^2) \right\}^{-\frac{1}{2}} \left[ 1 - \beta^2(1 - 2\sigma_1^2) \right]^{-\frac{1}{2}}, \end{aligned}$$

where  $\alpha = \sqrt{1 - z}$  and  $\beta = \frac{\sigma_1^2}{\sqrt{2(1 - \sigma_1^2)}}$ . Now we estimate the  $x$  defined as in Eq. (A2):

$$\begin{aligned} x &= [1 - \sqrt{1 - \rho^2}]/\rho \\ &= \frac{1 - z\sigma_1^2 - \sqrt{1 - z}\sqrt{1 + z - 2\sigma_1^2 z}}{z(1 - \sigma_1^2)} \\ &= 1 - 2\beta + 2\sigma_1^2\beta^2 - (2\sigma_1^2 - 3)\beta^3 + \dots \end{aligned}$$

When  $m$  is large, we have the estimate

$$x^m = 1 - 2m\beta + 2m^2\beta^2 - 4/3m^3\beta^3 + \dots$$

Thus

$$(1 - x^m)^{-1} = \frac{1}{2m\beta} \left( 1 + m\beta + \frac{1}{3}m^2\beta^2 + O(\beta^3) \right).$$

Then the  $k_1 = 0$  term is given by

$$[m^2(1 - z)]^{-1} + \frac{1}{6(1 - \sigma_1^2)} + O(1 - z)^{1/2}.$$

For other values of  $k_1$  as  $z \rightarrow 1$ , the simplification is based on the term  $(1 - zc_1\sigma_1^2)\sqrt{1 - \rho^2}$ . We have

$$\begin{aligned} & (1 - c_1\sigma_1^2)\sqrt{1 - \rho^2} \\ &= \left[ (1 - c_1\sigma_1^2)^2 - (\sigma_2^2 + c_1\sigma_3^2)^2 - (\sigma_3^2s_1)^2 \right]^{1/2} \\ &= \sqrt{1 - c_1} \left( 1 - \sigma_2^4 - \sigma_3^4 - c_1\sigma_1^4 \right)^{1/2} \\ &= \sqrt{1 - c_1} \sqrt{2(\sigma_1^2\sigma_2^2 + \sigma_1^2\sigma_3^2 + \sigma_2^2\sigma_3^2) + \sigma_1^4(1 - c_1)} \\ &= 2\sqrt{\sigma_1^2\sigma_2^2 + \sigma_1^2\sigma_3^2 + \sigma_2^2\sigma_3^2} \sin(\pi k_1/m) \\ &\quad \times [1 + \eta \sin^2(\pi k_1/m)]^{1/2}, \end{aligned}$$

where

$$\eta = \frac{\sigma_1^4}{\sigma_1^2\sigma_2^2 + \sigma_1^2\sigma_3^2 + \sigma_2^2\sigma_3^2}.$$

Thus

$$\begin{aligned} \frac{(1 - c_1\sigma_1^2)^{-1}}{\sqrt{1 - \rho^2}} &= \frac{1}{2\sqrt{\sigma_1^2\sigma_2^2 + \sigma_1^2\sigma_3^2 + \sigma_2^2\sigma_3^2}} \times \frac{1}{\sin(\pi k_1/m)} \\ &\quad \times [1 + \eta \sin^2(\pi k_1/m)]^{-1/2}. \end{aligned}$$

Now for  $k_1 \neq 0$ , and let  $z \rightarrow 1$ , the term

$$\frac{(1 - c_1\sigma_1^2)^{-1}}{\sqrt{1 - \rho^2}} \frac{1 - x^{2m}}{1 - 2x^m \cos m\phi + x^{2m}}$$

in (A3) can be written as the sum of three terms, and we expand  $P(0, z)$  in (A3) to the second order of  $1 - z$  as

$$P(0, z) = [m^2(1 - z)]^{-1} + \frac{1}{2(1 - \sigma_1^2)} \left[ \frac{1}{3} + \frac{\Phi}{r} \right] + O(1 - z)^{1/2},$$

where

$$r = \frac{\sqrt{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2}}{1 - \sigma_1^2}, \quad (\text{A4})$$

and  $\Phi = S_1 + S_2 + S_3$  with

$$\begin{aligned} S_1 &= \frac{1}{m} \sum_{k=1}^{m-1} \frac{1}{\sin(\pi k/m)}, \\ S_2 &= \frac{1}{m} \sum_{k=1}^{m-1} \frac{[1 + \eta \sin^2(\pi k/m)]^{-1/2} - 1}{\sin(\pi k/m)}, \\ S_3 &= \frac{1}{m} \sum_{k=1}^{m-1} \frac{1}{\sin(\pi k/m)} \left\{ \frac{2x^m (\cos \phi m - x^m)}{1 - 2x^m \cos \phi m + x^{2m}} \right\} \\ &\quad \times [1 + \eta \sin^2(\pi k/m)]^{-1/2}. \end{aligned}$$

Following Appendix B in Montroll,<sup>32</sup> there are estimates for  $S_1, S_2$ :

$$\begin{aligned} S_1 &= \frac{2}{\pi} \left\{ \log m + [\gamma + \log(2/\pi)] - \frac{\pi^2}{72m^2} + \dots \right\}, \\ S_2 &= -\frac{1}{\pi} \log(1 + \eta) + \frac{\eta\pi}{12m^2} + \dots \end{aligned}$$

The estimation for  $S_3$  is complicated. But fortunately  $S_3$ 's contribution to  $P(0, z)$  can be considered small if  $r$ , defined in (A4), is not too small. Note that we can obtain an estimation (see the appendix in Montroll<sup>32</sup>)

$$x = \left[ 1 - \sqrt{1 - \rho^2} \right] / \rho \approx 1 - 2rs + 2r^2s^2 + \dots,$$

where  $s = \sin \pi k/m$  and  $r$  is given in (A4). Thus

$$x^m \approx e^{-2rsk} (1 + O(1/m^2)).$$

If  $r$  is close to zero,  $x^m$  will be close to one. Then  $\frac{x^m}{1-x^m}$  will be very large and so will  $S_3$ . But if  $r$  is large,  $x^m$  will be close to zero (except for the term  $k = 0$ ) and  $S_3$  will be small. In order to control the error from the  $S_3$  term, we always choose  $D_u, D_v$ , and  $D_w$  such that  $D_u \geq D_v \geq D_w$ . In this way,  $r \geq \left[ \frac{\sigma_1^2}{\sigma_2^2 + \sigma_3^2} \right]^{1/2} \geq \frac{\sqrt{2}}{2}$ , and  $S_3$  will remain relatively small, when  $m$  is large. Thus we will simply disregard the  $S_3$  term in our formula.

To sum it up, when we choose  $\sigma_1 \geq \sigma_2 \geq \sigma_3$  and considering  $N = m^2$ , we disregard  $S_3$  and have an estimate of  $P(0, z)$  as

$$\begin{aligned} P(0, z) &\approx [N(1 - z)]^{-1} + \{c_1 \log N + c_2 + c_3/N\} \\ &\quad + O(N^{-2}) + O((1 - z)^{1/2}), \end{aligned}$$

where

$$\begin{aligned} c_1 &= \frac{1}{2\pi\hat{\sigma}}, \\ c_2 &= \frac{1}{6(1 - \sigma_1^2)} + \frac{1}{2\pi\hat{\sigma}} [2(\gamma + \log(2/\pi)) - \log(1 + \eta)], \\ c_3 &= \frac{\pi}{24\hat{\sigma}} (\eta - 1/3), \end{aligned}$$

with  $\gamma$  defined in (20),

$$\hat{\sigma} = \sqrt{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2},$$

and

$$\eta = \frac{\sigma_1^4}{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2}.$$

According to (A1), we have

$$\langle n \rangle = c_1 N \log N + c_2 N + c_3 + O\left(\frac{1}{N}\right). \quad (\text{A5})$$

Now consider the special case  $D_w = 0$ . Then  $\sigma_3 = 0$  and the domain is really a square lattice. In this case, we have

$$\hat{\sigma} = \sigma_1 \sigma_2, \quad \eta = \frac{\sigma_1^2}{\sigma_2^2}, \quad r = \frac{\sigma_1}{\sigma_2}.$$

If we select  $D_u \geq D_v$  (thus  $\sigma_1 \geq \sigma_2$ ),  $r \geq 1$ , the  $S_3$  term will be relatively small. Then multiplying (A5) with the average time for each jump  $1/k$ , where  $k = 2(D_u + D_v)/h^2$ , applying  $N = (L/h)^2$ , and disregarding lower order terms, we obtain

$$\begin{aligned} \tau_{coll} &= \frac{L^2}{2\pi\sqrt{D_u D_v}} \log\left(\frac{L}{h}\right) + \frac{L^2}{12D_v} \\ &\quad + \frac{L^2}{4\pi\sqrt{D_u D_v}} \left[ 2(\gamma + \log(2/\pi)) - \log\left(1 + \frac{D_u}{D_v}\right) \right]. \quad (\text{A6}) \end{aligned}$$

If we assume further that  $D_u = D_v$ , formula (A6) is close to (15) except a small difference due to the  $S_3$  term. Note that in this case, formula (A6) is a rigorous estimate.

If  $D_w \neq 0$  and assume the 2D lattice is square, we let  $N = \frac{L^2}{h^2}$  and multiply (A5) with  $\frac{h^2}{2(D_u + D_v + D_w)}$ , we end up with the estimation of  $\tau_{coll}$  in formula (23).

## APPENDIX B: SPECIAL CASE: $D_u \rightarrow \infty$

When the diffusion rate of  $U$  approaches to infinity, the trimolecular system becomes a bimolecular collision model of  $V$  and  $W$ . Formula (28) gives the mean time for bimolecular collisions, when  $D_u \rightarrow \infty$ :

$$\lim_{D_u \rightarrow \infty} \tau_{coll} = 0.140 \frac{L^2}{D_v + D_w}. \quad (\text{B1})$$

In this appendix, we investigate the mean bimolecular collision time and derive the formula for the bimolecular collision time when the other two molecules have the same diffusion rate  $D_v = D_w$ .

Assume a 1D domain of length  $L^\theta$  and two molecules  $V$  and  $W$  diffuse freely with the rates  $D_v$  and  $D_w$ , respectively. The 1D diffusion model of two molecules is equivalent to the 2D model in which one molecule diffuses freely with the rates  $D_v$  and  $D_w$  in the two directions, independently. The first time for the two molecules to collide in the same position is equivalent to the first time when the molecule in the 2D domain comes across the diagonal line. With reflecting boundary conditions, the triangle domain divided by the diagonal line can be extended into a square domain with the length  $L = \sqrt{2}L^\theta$ .

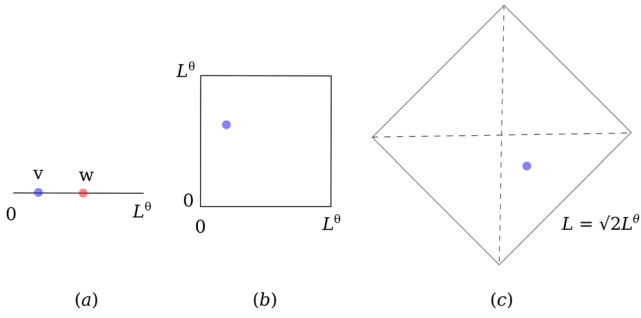


FIG. 8. The conversion of the 1D collision model to the first exit time model in a 2D domain. (a) The collision model of two molecules  $V$  and  $W$  freely diffusing in 1D. (b) The diffusion model of one molecule in the 2D domain. (c) The first exit time model of one molecule in the 2D square domain.

The first encounter time of two molecules on the 1D domain of size  $L^\theta$  is now converted to a first exit time of one molecule on a 2D domain of size  $L = \sqrt{2}L^\theta$ , shown in Fig. 8.

In the following, we derive the formula for the first exit time on the 2D domain. We note that it is already known<sup>34</sup> that the reaction time for a bimolecular reaction in the 1D domain does converge when  $h \rightarrow 0$ . Based on the above analysis, the first exit time on the 2D domain should also converge when  $h \rightarrow 0$ . Thus it is more convenient to analyze the first exit time based on continuous diffusion rather than a random walk in a 2D domain. That will be our strategy in the following.

For simplicity, we assume the diffusion rate in the two directions are the same;  $D = D_v = D_w$ . The Plank-Fokker equation for the diffusion in the 2D domain is given by

$$\frac{\partial}{\partial t} P(\mathbf{x}, t | \mathbf{x}_0, t_0) = D \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) P(\mathbf{x}, t | \mathbf{x}_0, t_0), \quad (\text{B2})$$

with  $P(x, t | x_0, t_0)$  being the state density function, defined as the probability density that the molecule stays in the position  $\mathbf{x}$  at time  $t$  given it starts from  $\mathbf{x}_0$  at time  $t_0$ .

Next, we define a probability function

$$G(\mathbf{x}_0, t; \Omega) \equiv \int_0^L \int_0^L dx dy P(\mathbf{x}, t | \mathbf{x}_0, t_0), \quad (\text{B3})$$

which describes the probability that the molecule stays in the domain  $\Omega = (0, L) \times (0, L)$  at time  $t$ , given it starts from  $\mathbf{x}_0$  at time  $t = 0$ . Then, we integrate the Plank-Fokker equation (B2) over the 2D interval  $\Omega$  and we have the equation for  $G$ :

$$\frac{\partial}{\partial t} G(x_0, t; a, b) = D \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) G(x_0, t; a, b). \quad (\text{B4})$$

The initial condition for the PDE (B4) is given as

$$G(\mathbf{x}_0, 0; \Omega) = 1.$$

The four boundaries of the square domain  $\Omega$  are all absorbing. Hence, we have the boundary conditions for PDE (B4) as

$$\begin{aligned} G((0, y), t; \Omega) &= 0; & G((L, y), t; \Omega) &= 0; \\ G((x, 0), t; \Omega) &= 0; & G((x, L), t; \Omega) &= 0. \end{aligned} \quad (\text{B5})$$

Following the definition of  $G(\mathbf{x}_0, t; \Omega)$  (B3),  $1 - G(\mathbf{x}_0, t; \Omega)$  gives the probability that the molecule exits  $\Omega$  before time  $t$ , which is exactly the distribution function of the first exit time  $T(\mathbf{x}_0; \Omega)$ . In addition, the density function of  $T(\mathbf{x}_0; \Omega)$  is given by

$$p(t, \mathbf{x}_0; \Omega) = \frac{\partial}{\partial t} [1 - G(\mathbf{x}_0, t; \Omega)] = -\frac{\partial}{\partial t} G(\mathbf{x}_0, t; \Omega), \quad (\text{B6})$$

and the  $n$ th moment  $T_n(\mathbf{x}_0; \Omega)$  of the random variable  $T(\mathbf{x}_0; \Omega)$  is therefore given by

$$T_n(\mathbf{x}_0; \Omega) = \int_0^\infty t^n \left[ -\frac{\partial}{\partial t} G(\mathbf{x}_0, t; \Omega) \right] dt, \quad (n \geq 0). \quad (\text{B7})$$

For  $n = 0$ , the formula gives

$$T_0(\mathbf{x}_0; \Omega) = -(G(\mathbf{x}_0, \infty; \Omega) - G(\mathbf{x}_0, 0; \Omega)) = 1.$$

Integrating (B7) by parts, we get the formula

$$\int_0^\infty t^{n-1} G(\mathbf{x}_0, t; \Omega) dt = \frac{1}{n} T_n(\mathbf{x}_0; \Omega) \quad (n \geq 1). \quad (\text{B8})$$

With Eqs. (B6) and (B8), we can formulate the moments of  $T_n(\mathbf{x}_0; \Omega)$  as coupled ordinary differential equations. Multiplying the Plank-Fokker equation (B2) by  $t^{n-1}$ , integrating the result over all  $t$ , and substituting from Eqs. (B6) and (B8), we have

$$-T_{n-1}(\mathbf{x}_0, \Omega) = D \left( \frac{\partial^2}{\partial x_0^2} + \frac{\partial^2}{\partial y_0^2} \right) \left( \frac{1}{n} T_n(\mathbf{x}_0; \Omega) \right).$$

The boundary conditions for these differential equations follow the simple derivation from (B5), and we have

$$\begin{aligned} T_n((0, y), t; \Omega) &= 0; & T_n((L, y), t; \Omega) &= 0; \\ T_n((x, 0), t; \Omega) &= 0; & T_n((x, L), t; \Omega) &= 0. \end{aligned}$$

With the equations ready, we can solve for the moments of the first passage time. Here we are only interested in the

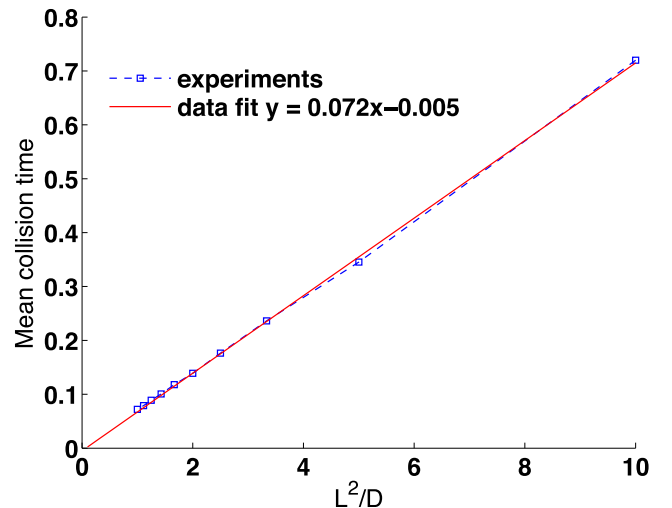


FIG. 9. The mean first collision time of two molecules in the 1D domain with reflecting boundary conditions, while the two molecules have the same diffusion rate,  $D = D_v = D_w$ . Other parameters are  $N = 64$  and  $L = 0.1$ .

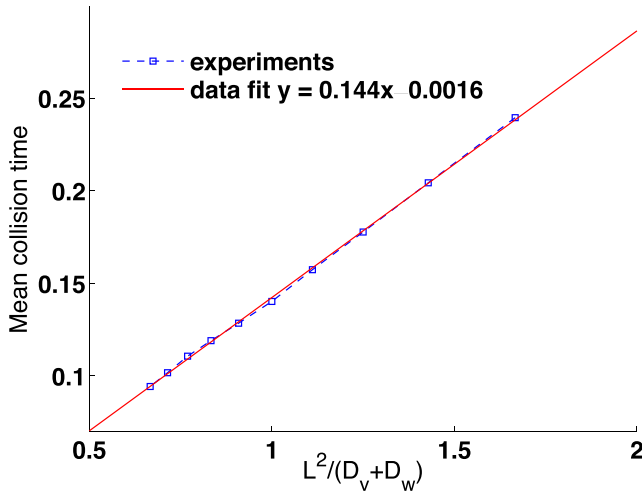


FIG. 10. The mean first collision time of two molecules in the 1D domain with reflecting boundary conditions, while the two molecules have different diffusion rates,  $D_v \neq D_w$ . Other parameters are  $N = 64$  and  $L = 0.1$ .

first moment, and the solution to the PDE of the first moment equation yields

$$T_1(x, y) = \frac{x(L-x)}{2D} - \frac{4L^2}{D\pi^3} \sum_{k=1, \text{odd}}^{\infty} \left\{ \frac{\sin(k\pi x/L)}{k^3 \sinh(k\pi)} \times (\sinh(k\pi y/L) + \sinh(k\pi(L-y)/L)) \right\}.$$

If initially the molecule is homogeneously presented in the square domain, we can calculate the mean first exit time as

$$\begin{aligned} \langle T_1 \rangle &= \frac{1}{L^2} \int_0^L \int_0^L T_1(x, y) dx dy \\ &= \frac{L^2}{12D} - \frac{16L^2}{D} \sum_{k=1, \text{odd}}^{\infty} \frac{\cosh(k\pi) - 1}{\pi^5 k^5 \sinh(k\pi)} \\ &\approx 0.0351 \frac{L^2}{D}. \end{aligned}$$

Therefore, the mean first time when the first two molecules encounter in the 1D domain of size  $L^\theta$  is exactly the same as the mean first exit time above, and the mean first encounter time is given by

$$\langle T_{\text{coll}} \rangle \approx 0.0702 \frac{(L^\theta)^2}{D}. \quad (\text{B9})$$

Figures 9 and 10 illustrate the numerical results of the mean first collision time for two molecules with the same diffusion rate and different diffusion rates for the general situations. The linear data fitting shows the excellent match with the formula (B9). Furthermore, although our derivation is only for the case  $D_v = D_w$ , the numerical results show that the first collision time for the general situation, where  $D_v \neq D_w$ , follows the similar formula. This formula is given as

$$\langle T_{\text{coll}} \rangle \approx 0.140 \frac{(L)^\theta}{D_v + D_w}.$$

For comparison purposes, Fig. 11 gives the numerical results of the mean first collision time, under periodic boundary conditions, for two molecules with different diffusion

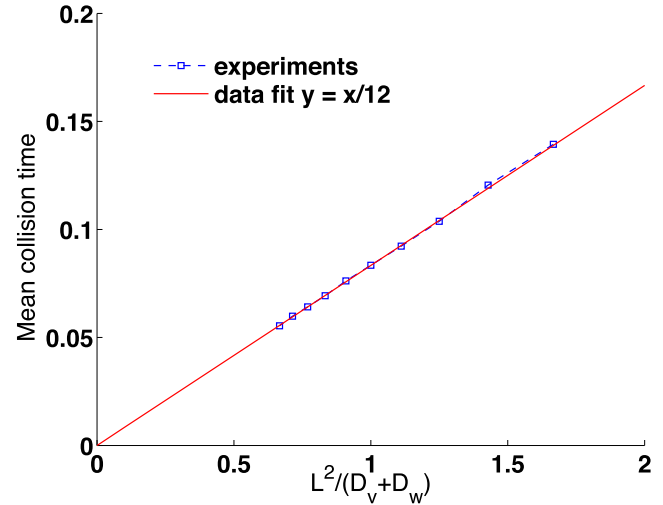


FIG. 11. The mean first collision time of two molecules in the 1D domain with periodic boundary conditions, while the two molecules have different diffusion rates,  $D_v \neq D_w$ . Other parameters are  $N = 128$  and  $L = 0.1$ .

rates ( $D_v \neq D_w$ ). We can see that the numerical results match with (26) very well.

- <sup>1</sup>J. Schnakenberg, *J. Theor. Biol.* **81**, 389 (1979).
- <sup>2</sup>F. Schlögl, *Z. Phys.* **253**(2), 147 (1972).
- <sup>3</sup>L. Qiao, R. Erban, C. Kelley, and I. Kevrekidis, *J. Chem. Phys.* **125**, 204108 (2006).
- <sup>4</sup>G. Oshanin, A. Stemmer, S. Luding, and A. Blumen, *Phys. Rev. E* **52**(6), 5800 (1995).
- <sup>5</sup>T. Plesa, T. Vejchodský, and R. Erban, *J. Math. Chem.* **54**(10), 1884–1915 (2016).
- <sup>6</sup>R. Erban, S. J. Chapman, I. Kevrekidis, and T. Vejchodsky, *SIAM J. Appl. Math.* **70**(3), 984 (2009).
- <sup>7</sup>J. Paulsson, O. Berg, and M. Ehrenberg, *Proc. Natl. Acad. Sci. U. S. A.* **97**(13), 7148 (2000).
- <sup>8</sup>M. Elowitz, A. Levine, E. Siggia, and P. Swain, *Science* **297**, 1183 (2002).
- <sup>9</sup>D. Gillespie, *J. Phys. Chem.* **81**(25), 2340 (1977).
- <sup>10</sup>M. Gibson and J. Bruck, *J. Phys. Chem. A* **104**, 1876 (2000).
- <sup>11</sup>Y. Cao, H. Li, and L. Petzold, *J. Chem. Phys.* **121**(9), 4059 (2004).
- <sup>12</sup>R. Erban, S. J. Chapman, and P. Maini, e-print [arXiv:0704.1908](https://arxiv.org/abs/0704.1908) (2007).
- <sup>13</sup>R. Erban and S. J. Chapman, *Phys. Biol.* **4**(1), 16 (2007).
- <sup>14</sup>S. Andrews and D. Bray, *Phys. Biol.* **1**, 137 (2004).
- <sup>15</sup>K. Takahashi, S. Tanase-Nicola, and P. ten Wolde, *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2473 (2010).
- <sup>16</sup>J. Hattne, D. Fange, and J. Elf, *Bioinformatics* **21**(12), 2923 (2005).
- <sup>17</sup>S. Engblom, L. Ferm, A. Hellander, and P. Lötstedt, *SIAM J. Sci. Comput.* **31**, 1774 (2009).
- <sup>18</sup>R. Erban, *Proc. R. Soc. London, Ser. A* **470**, 20140036 (2014).
- <sup>19</sup>R. Erban, *Proc. R. Soc. London, Ser. A* **472**(2186), 20150556 (2016).
- <sup>20</sup>R. Erban and S. J. Chapman, *Phys. Biol.* **6**(4), 046001 (2009).
- <sup>21</sup>I. Agbanusi and B. S. Isaacson, *Math. Biol.* **76**(4), 922 (2014).
- <sup>22</sup>M. Flegg, *SIAM J. Appl. Math.* **76**(4), 1403–1432 (2016).
- <sup>23</sup>J. Lipkova, K. Zygalakis, J. Chapman, and R. Erban, *SIAM J. Appl. Math.* **71**(3), 714–730 (2011).
- <sup>24</sup>S. Isaacson, A. Mauro, and J. Newby, *Phys. Rev. E* **94**, 042414 (2016).
- <sup>25</sup>D. Torney and H. McConnell, *J. Phys. Chem.* **87**(11), 1941 (1983).
- <sup>26</sup>D. Torney and H. McConnell, *Proc. R. Soc. London, Ser. A* **387**, 147 (1983).
- <sup>27</sup>D. ben Avraham, *Phys. Rev. Lett.* **71**(24), 3733 (1993).
- <sup>28</sup>D. Gillespie, *Markov Processes: An Introduction for Physical Scientists* (Academic Press, Inc., Harcourt Brace Jovanovich, 1992).
- <sup>29</sup>R. Erban, M. Flegg, and B. G. Papoian, *Math. Biol.* **76**(4), 799 (2014).
- <sup>30</sup>U. Dobramysl, G. Papoian, and R. Erban, *Biophys. J.* **110**, 2066 (2016).
- <sup>31</sup>S. Hellander, A. Hellander, and L. Petzold, *Phys. Rev. E* **85**, 042901 (2012).
- <sup>32</sup>E. Montroll, *J. Math. Phys.* **10**(4), 753 (1969).
- <sup>33</sup>S. A. Isaacson, *J. Chem. Phys.* **139**(5), 054101 (2013).
- <sup>34</sup>S. A. Isaacson, *SIAM J. Appl. Math.* **70**(1), 77–111 (2009).