

A stochastic cellular automaton simulation of chemical oscillations in small systems

Cite as: AIP Advances **8**, 125211 (2018); <https://doi.org/10.1063/1.5051550>

Submitted: 09 August 2018 • Accepted: 03 December 2018 • Published Online: 13 December 2018

Yue Hu and Jenny Xie



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Perspective: Stochastic algorithms for chemical kinetics](#)

The Journal of Chemical Physics **138**, 170901 (2013); <https://doi.org/10.1063/1.4801941>

[Numerical modeling of three dimensional Brusselator reaction diffusion system](#)

AIP Advances **9**, 015205 (2019); <https://doi.org/10.1063/1.5070093>

[A short introduction to the Lindblad master equation](#)

AIP Advances **10**, 025106 (2020); <https://doi.org/10.1063/1.5115323>



A stochastic cellular automaton simulation of chemical oscillations in small systems

Yue Hu^a and Jenny Xie

Department of Physics, Wellesley College, Wellesley, Massachusetts 02481, USA

(Received 9 August 2018; accepted 3 December 2018; published online 13 December 2018)

We propose a stochastic cellular automaton method to simulate chemical reactions in small systems. Unlike the standard Gillespie method, which simulates chemical reactions with a few thousand molecules reacting with each other but without spatial considerations, our systems are divided into independent cells, each containing only a few molecules. Our simulation of the Brusselator produces chemical oscillations that agree extremely well with solutions to deterministic rate equations, and we can see strong oscillations in systems with as few as 10 cells. We are able to study several factors that affect the robustness of these small chemical oscillators: system size, spatial distribution, and correlation of molecules. We have found that non-Poisson particle distributions can greatly suppress chemical oscillations and that chemical reactions can induce correlation between the spatial distributions of particles of different species and create large-scale inhomogeneity in particle concentrations. In addition, incomplete oscillations (misfirings) can appear among strong, regular oscillations when the system size is smaller than a certain threshold, and these misfirings are triggered by random events, with a probability that is related to the system size. Since these effects, resulting from several different physical causes, are difficult to accurately model by adding generic noise factors to deterministic rate equations, as is frequently done in theoretical studies, we argue that our stochastic cellular automaton method is a useful addition to the existing tools for studying small, inhomogeneous, and non-equilibrium reaction-diffusion systems, especially those of biological nature. © 2018 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/1.5051550>

I. INTRODUCTION

Oscillatory systems are ubiquitous in the natural world, ranging from simple physical systems, such as pendulums, to complicated biological systems, such as cellular cycles.^{1,2} What these oscillators have in common is a restoring mechanism that can pull the system back to a certain configuration—a fixed point—when the system moves away from it. The characteristic frequency of an oscillator depends on system parameters and, in the case of nonlinear oscillators, on initial conditions.

Chemical oscillations can occur in systems with numerous chemical species that undergo multi-step reactions. In well-stirred systems, the concentrations of the species can change periodically with time, as seen in the Belousov-Zhabotinsky reaction, one of the most famous oscillatory chemical reactions.^{3,4} Because the laws of thermodynamics dictate that a closed system eventually reaches a state of equilibrium, where the free energy is a minimum, an oscillatory chemical system is obviously not at equilibrium. In addition, these systems are nonlinear, and the restoring or feedback mechanism is often an autocatalytic reaction, in which a product is also a reactant.^{5,6}

^aAuthor to whom correspondence should be addressed. Electronic mail: yhu@wellesley.edu

The scale of biological oscillations varies widely from individual cells to entire ecosystems, with periods ranging from a fraction of a second to days and years. Much progress has been made in understanding the mechanisms responsible for various biological rhythms at the level of enzyme activity, protein synthesis, and gene regulation and expression.² Even though one might regard biological oscillators as more complicated versions of chemical oscillators, there is nevertheless a crucial difference in terms of the numbers of molecules involved, due to the difference in the sizes of the molecules. For example, the average number of proteins in an *E. coli* bacterium, which has a volume of about $1\ \mu\text{m}^3$, is about 3×10^6 .⁷ Consequently, for a proteome size of about 5000, the average number of any particular type of protein in an *E. coli* bacterium is less than 1000. By contrast, a 100-mM solution of electrolytes contains about 10^8 ions in the same volume. This difference in the numbers of molecules has a major impact on determining the appropriate theoretical tools for studying these systems. Generally speaking, understanding the robustness of biological rhythms requires a good understanding of fluctuations and correlations due to the small numbers of molecules in biological oscillators.²

The conventional method for studying the kinetics of chemical reactions involves using rate equations, which are ordinary differential equations (ODEs) that describe the rate of change of the concentration of each chemical species based on the chemical reactions involved. The rate equations assume homogeneous concentrations of all chemical species, and the solutions to the rate equations are deterministic for a given set of initial conditions. Applying rate equations to biological systems, however, is problematic, because the fluctuations that naturally arise from small particle numbers are not reflected in these deterministic equations and spatial homogeneity is difficult to achieve in a biological cell crowded with macromolecules.⁸

In the 1970s, Gillespie proposed a stochastic algorithm to simulate chemical reactions, and this algorithm has become the *de facto* standard in the simulation community.^{8,9} In the Gillespie algorithm, two random numbers are generated in each simulation step. One of these numbers is used to determine which chemical reaction will occur, with probabilities weighted by the reaction rate constants and particle concentrations. The other number is used to determine the duration of the time step. The simulation terminates when no more reactants remain. Gillespie proved that this algorithm is statistically fully equivalent to rate equations but with the added benefit of having the built-in impact of stochastic events. Similar to deterministic chemical rate equations, the Gillespie algorithm assumes well-mixed, uniformly distributed chemical species. A typical simulation usually involves a few thousand reactants of each species, and averaging a few simulation runs may also be necessary.⁸

Despite the enormous success of the Gillespie algorithm and several subsequent modifications that are computationally more efficient,^{10–14} the issues of how to treat spatial inhomogeneity and take into account diffusion of molecules remain an ongoing research effort.¹⁵ Methods of simulating reaction-diffusion in heterogeneous systems usually involve dividing systems into many subvolumes and allowing chemical reactions to occur in the subvolumes with diffusion of molecules between subvolumes.¹⁶ For example, in the binomial tau-leap spatial stochastic simulation algorithm,¹⁷ at each time step a chosen subvolume can undergo either a chemical reaction using a Gillespie-based method, or molecules in the subvolume can diffuse into neighboring subvolumes, depending on various parameters specified by the algorithm. In other words, simulation is performed sequentially in the subvolumes, one subvolume for each time step.

In cellular automaton simulations, on the other hand, systems are divided into multiple, independent units called “cells,” and events occur simultaneously in discrete time steps in each cell following rules that are applied uniformly to all cells. Adapting Gillespie’s method to cellular automata, however, is difficult because time steps in the Gillespie algorithm are determined stochastically and are uneven. This unevenness in time steps does not allow uniform advances in time for all cells, as is customary in a cellular automaton simulation.¹⁸ Nevertheless, researchers have developed several methods with different levels of complexity to study reaction-diffusion systems^{18–22} (especially surface reactions^{23–26}) using cellular automata. Direct comparisons of simulation results and solutions to differential equations, however, are uncommon,²⁷ except for simple reactions.^{22,28}

Our method uses standard cellular automaton techniques to divide the “world” into multiple, parallel “cells,” each occupied by chemical “agents,” and the status of the world is updated in discrete

time steps. During each time step, chemical reactions are allowed to take place in each cell independently following rules that are stochastic in nature but that are applied to all cells simultaneously. Chemical agents are also allowed to move between neighboring cells. Our method has elements similar to Gillespie's method and an earlier method by Nakanishi,²⁹ but with a few major differences in terms of the numbers of molecules involved in each reaction unit and how reactions and time steps are chosen. A major advantage of our method is that it can incorporate spatial inhomogeneity of particles to simulate reaction-diffusion systems when stochastic events are important, and our method is also easy to implement.

In the following sections, we will give a rationale for our method and test our algorithm using simple chemical reactions of various orders. We will then extend our method to more complicated chemical oscillators that involve multi-step chemical reactions. We will make direct comparisons of our simulation results with solutions of deterministic rate equations, which we regard as accurate descriptions of the chemical kinetics of macroscopic, homogenous experimental systems. We will also investigate the robustness of these simulated chemical oscillators when the system size is reduced and will report findings that are characteristic of small systems and that are not seen in larger systems. In addition, we will show that chemical oscillations are suppressed or distorted in the non-Poisson particle distributions that we have examined. Even though the scope of the current study is limited principally to homogeneous systems, we hope to further develop this method for application to spatially inhomogeneous reaction-diffusion systems — and particularly to networks of biochemical systems — where fluctuations due to small system size can play critical roles.

II. A STOCHASTIC CELLULAR AUTOMATON SIMULATION OF CHEMICAL REACTIONS

We will introduce our simulation method by examining standard chemical reaction rate equations and giving them statistical interpretations. We will then test our simulation algorithm using specific examples with increasing complexity. Our simulation methods for first- and second-order reactions are similar to those used by other authors,^{22,28} but our approach to multi-step reactions is different from the methods of other researchers, to the best of our knowledge.

Since our algorithm simulates chemical reactions on a molecular level, all concentrations used in this paper are number concentrations. We will refer to a reaction that involves a minimal set of reactant molecules as a “stoichiometric reaction.” For example, the stoichiometric reaction for the reaction $2A + B \rightarrow C$ involves a minimal reactant set of two A molecules and one B molecule. When molecule numbers exceed this minimal set, an important factor, as we will see later, is how many distinct reactant sets, or microstates, these molecules can form. Since the thermal de Broglie wavelength is less than an angstrom at room temperature for all molecules that are of interest here, we will treat the molecules as distinguishable, classical particles when enumerating distinct reactant sets, and we will denote the total number of possible reactant sets that can be formed from a given group of reactant molecules as the “multiplicity factor” (Ω), a concept also used in statistical mechanics.

A. Spatial distribution of molecules

We assume that our chemical system is divided into L cells. For computational simplicity, we assume that the cells lie in one dimension, but our method can easily be generalized to two or three dimensions. If we have N molecules of the same species, then the probability of finding a particular molecule in a cell is $1/L$ in a homogeneous system. One can easily prove that for large values of N and L , the number of molecules in the cells follows a Poisson distribution:³⁰

$$P(n) = \frac{\lambda^n}{n!} e^{-\lambda}, \quad (1)$$

where $P(n)$ is the probability of a cell having n molecules and $\lambda = N/L$ is the average number of molecules per cell. Note that this distribution depends only on n and λ and does not depend on N and L explicitly. For large values of λ , a Poisson distribution approaches a normal distribution. In this work, however, we are mostly interested in cases where λ is small.

To achieve a Poisson distribution in our simulation, we can pick a random cell from L cells and deposit a molecule in it and repeat the process as many times as we have molecules, or we can let molecules perform random walks (RW) multiple times. Our simulation and analysis show that when λ is small, a large number of RW steps are required for a non-Poisson distribution (for example, a flat distribution) to become Poisson. We will experiment with both methods in our work. The first method is computationally less time-consuming, but the second method mimics a diffusion process and is physically more realistic.

B. Simulation of single chemical reactions

1. Example 1. A first-order chemical reaction

We consider a simple first-order chemical reaction with a forward rate constant k_f and a backward rate constant k_b :



The rate equations for this reaction are

$$\frac{dA}{dt} = -\frac{dB}{dt} = -k_f A + k_b B, \quad (3)$$

where A and B are concentrations of chemical species A and B and are functions of time. (We will use upper-case, italicized letters for concentrations.) To give a statistical interpretation to the above rate equations, we rewrite the differential rate equation for A for the forward reaction ($dA/dt = -k_f A$) as a difference equation for a short time step τ :

$$\frac{\Delta A}{A} = -k_f \tau. \quad (4)$$

This equation states that during a time interval τ , the fractional reduction in the number of the molecules is equal to $k_f \tau$. Interpreting this fractional change as a probability, we conclude that statistically, the rate equations for the first-order reaction (2) are equivalent to the following statement:

During a small time interval τ , any molecule of species A has a probability $k_f \tau$ of becoming a molecule of species B , and any molecule of species B has a probability $k_b \tau$ of becoming a molecule of species A ($k_f \tau, k_b \tau \ll 1$; first-order reaction $A \rightleftharpoons B$).

Since this first-order reaction does not involve any A-B combinations, simulating this reaction is reasonably straightforward using a Monte Carlo method (please see the [Appendix](#) for details), and our results confirm previous work by other researchers.³¹

2. Example 2. A second-order chemical reaction (different reactant species)

The following second-order chemical reaction involves two reactants and one product and has a forward rate constant k_f and a backward rate constant k_b :



The rate equations for this reaction are:

$$\frac{dA}{dt} = \frac{dB}{dt} = -\frac{dC}{dt} = -k_f AB + k_b C. \quad (6)$$

For the initial conditions $A(0) = B(0) = A_0$ and $C(0) = 0$, the analytical solutions to the above rate equations are

$$C(t) = C_\infty \frac{1 - e^{-\alpha t}}{1 - \beta e^{-\alpha t}}, \quad A(t) = B(t) = A_0 - C(t), \quad (7)$$

where

$$\alpha = \sqrt{k_b(4A_0k_f + k_b)}, \quad \beta = \frac{k_b + 2A_0k_f - \alpha}{k_b + 2A_0k_f + \alpha},$$

and

$$C_\infty = \frac{k_b + 2A_0k_f - \alpha}{2k_f}.$$

We again rewrite the differential rate equation for A for the forward reaction ($dA/dt = -k_f AB$) as a difference equation for a small time step τ :

$$\frac{\Delta A}{A} = -Bk_f\tau.$$

Comparing the above expression with Eq. (4), we note the appearance of an additional factor of B on the right-hand side here. This indicates that compared to the first-order reaction (2), the probability of a molecule of species A participating in this second-order reaction is increased by a factor of B , the number of molecules of species B in a unit volume.

Another way to write this equation is

$$\Delta A = -ABk_f\tau. \quad (8)$$

The left-hand side of the above equation is equal to the number of times the stoichiometric reaction will occur in a unit volume (because each reaction consumes one A molecule), and the product AB on the right-hand side is numerically equal to the number of distinct A-B molecular pairs in the same volume. We therefore give the following statistical interpretation to the forward reaction in the second-order reaction (5):

In a volume V_1 , the probability that the stoichiometric reaction $A + B \xrightarrow{k_f} C$ will occur during a small time interval τ is equal to $k_f\tau/V_1$, and the maximum number of times that this reaction can occur is equal to the multiplicity factor, which is equal to the number of distinct A-B molecular pairs in the volume: $\Omega = ABV_1^2 = N_A N_B$, where $N_A = AV_1$ and $N_B = BV_1$ are the numbers of molecules of species A and B, respectively, in volume V_1 ($ABk_f\tau \ll A, B$).

Since rate constants have different units for reactions of different orders, we have included the volume explicitly in the above statement to ensure that the units are correct. In our simulations, however, for convenience, we set V_1 , the volume of each cell, to one unit unless otherwise noted.

Following the steps in the Appendix, we simulate reaction (5), and Fig. 1 shows both the simulation results and the analytical results given by Eq. (7). The agreement is very good.

3. Example 3. A second-order chemical reaction (same reactant species)

The following second-order chemical reaction involves reactants of the same species and one product:



Assuming the forward rate constant is k_f and the backward rate constant is k_b , the rate equations for this reaction are:

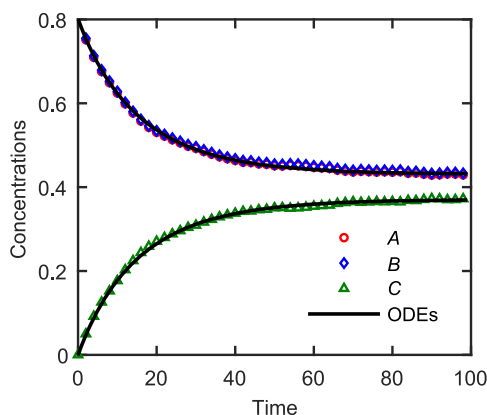


FIG. 1. Simulation results of the average concentrations A (open circles), B (open diamonds), and C (open triangles) are plotted as functions of time for the second-order reaction $A + B \rightleftharpoons C$. The solid lines are analytical solutions (Eq. (7)) to the rate equations (Eq. (6)). Simulation parameters: $L = 10000$, $k_f = 0.04$, $k_b = 0.02$, and $\tau = 1$. (The initial values of A and B are the same, resulting in their almost identical time dependence in the simulation.)

$$\frac{1}{2} \frac{dA}{dt} = -\frac{dC}{dt} = -k_f A^2 + k_b C. \quad (10)$$

For the initial conditions $A(0) = A_0$ and $C(0) = 0$, the analytical solutions to the above rate equations are

$$C(t) = C_\infty \frac{1 - e^{-\gamma t}}{1 - \delta e^{-\gamma t}}, \quad A(t) = A_0 - 2C(t), \quad (11)$$

where

$$\gamma = \sqrt{k_b(8A_0k_f + k_b)}, \quad \delta = \frac{k_b + 4A_0k_f - \gamma}{k_b + 4A_0k_f + \gamma},$$

and

$$C_\infty = \frac{k_b + 4A_0k_f - \gamma}{8k_f}.$$

We now try to give a statistical interpretation for the forward reaction in the rate equation $((1/2)dA/dt = -k_f A^2)$. Suppose that in the i th cell, there are $N_{A,i}$ molecules of species A. If we follow arguments similar to those presented in Example 2, we would then conclude that the multiplicity factor should be $N_{A,i}^2$. However, the number of distinguishable pairs in this cell is actually $N_{A,i}(N_{A,i} - 1)/2$. It seems as if two issues need careful consideration: Is enumerating particle pairs the correct statistical interpretation for rate equations when stoichiometric coefficients are larger than 1, and does the order of the molecules in a pair matter? For example, if there are two A molecules in the cell ($N_{A,i} = 2$), should the multiplicity factor Ω_i be 1, 2, or even 4? We will answer this question by running computer experiments with the values of Ω_i set to $N_{A,i}(N_{A,i} - 1)/2$, $N_{A,i}(N_{A,i} - 1)$, and $N_{A,i}^2$, ($i = 1, 2, \dots, L$).

Figure 2 shows simulation results of reaction (9) compared to the analytical results given by Eq. (11). We can see that only $\Omega_i = N_{A,i}(N_{A,i} - 1)$ produces results that agree with analytical solutions. This means that if, for example, we have two A molecules ($N_{A,i} = 2$) reacting with each other, then the multiplicity factor is $\Omega_i = 2$. In other words, the order of the molecules in the pair *does* matter. In the previous example for reactants of different species, the multiplicity factor is $\Omega_i = 1$ for a molecule A reacting with a molecule B ($N_{A,i} = 1$ and $N_{B,i} = 1$); the order of the molecules in the pair *does not* matter in this case. This inconsistency in the ways we count A-B pairs and A-A pairs, however, is simply the result of how the rate equations are written and is not due to some conflict with fundamental physical laws. Instead of using a consistent counting method and introducing new rate constants that are defined differently from the ones in conventional rate equations (as Gillespie did^{8,9}), we choose to use the conventional rate constants but give them different statistical interpretations depending on whether the reactant species are the same or different. (We draw readers' attention to the subtle

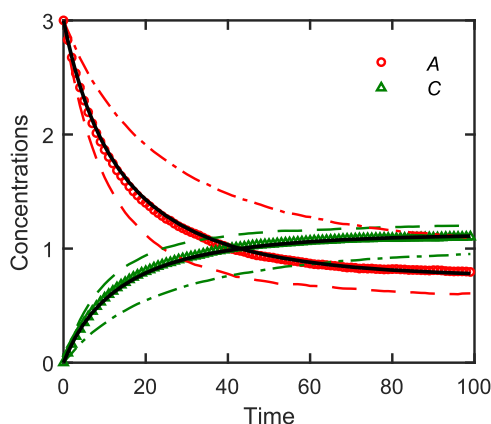


FIG. 2. Simulation results of the average concentrations A (open circles) and C (open triangles) are plotted as functions of time when using the multiplicity factor $\Omega_i = N_{A,i}(N_{A,i} - 1)$ for the forward reaction in $2A \rightleftharpoons C$. The solid lines are analytical solutions (Eq. (11)) to the rate equations (Eq. (10)); the agreement is excellent. The dashed lines are simulation results for $\Omega_i = N_{A,i}^2$, and the dash-dotted lines are simulation results for $\Omega_i = N_{A,i}(N_{A,i} - 1)/2$; neither of these agrees with the analytical solutions. Simulation parameters: $L = 10000$, $k_f = 0.01$, $k_b = 0.005$, and $\tau = 1$.

difference in how molecules are counted in a reactant set in simulations and to the existence of several kinds of chemical rate constants used in the literature; confusion can arise if these are not handled correctly.)

4. Example 4. A third-order chemical reaction

For a third-order reaction, we consider



Assuming the forward rate constant is k_f and the backward rate constant is k_b , the rate equations for this reaction are:

$$\frac{1}{2} \frac{dA}{dt} = \frac{dB}{dt} = -\frac{dC}{dt} = -k_f A^2 B + k_b C. \quad (13)$$

We will solve the above differential equations numerically. To simulate this reaction, we use the multiplicity factor $\Omega_i = N_{A,i}(N_{A,i} - 1)N_{B,i}$ for the forward reaction in the i th cell ($i = 1, 2, \dots, L$). In Fig. 3, we plot simulation results and numerical solutions to Eq. (13); we again see excellent agreement between the two. Furthermore, if we set the volume of a cell to $V_1 = 2$ and double the initial molecule numbers (thus keeping the initial concentrations unchanged), we see no change in Fig. 3, indicating that the volume normalization used in calculating the reaction probability (step 4 in the Appendix) is appropriate.

Another useful result of the rate equations (13) is that at equilibrium (where $dA/dt = dB/dt = dC/dt = 0$), the following product-to-reactant concentration ratio is a constant:

$$\frac{C_\infty}{A_\infty^2 B_\infty} = \frac{k_f}{k_b} = K. \quad (14)$$

Here K is the equilibrium constant, and this is called the law of mass action in chemistry.

To test this, we run our simulation using randomly generated initial concentrations for each chemical species. Equilibrium concentrations A_∞ , B_∞ , and C_∞ are obtained by averaging the concentrations over many time steps after the system has reached equilibrium. In Fig. 4, we plot simulation results of C_∞ versus $A_\infty^2 B_\infty$, which are in good agreement with the law of mass action.

It is worth noting that the concentrations in the simulations in Fig. 4 are so low that it is much less likely to find two A molecules in a cell than it is to find one. Since it takes a minimum of two A molecules in a cell for the forward reaction in reaction (12) to occur, this means that the forward reaction cannot occur in most cells. The calculation of average concentrations, however, uses molecules in all cells, including these “inactive” cells. Nevertheless, the statistics still work out extremely well for our simulation. Furthermore, rate equations assume that chemical concentrations can make arbitrarily small changes over arbitrarily small time intervals. In contrast, particle numbers

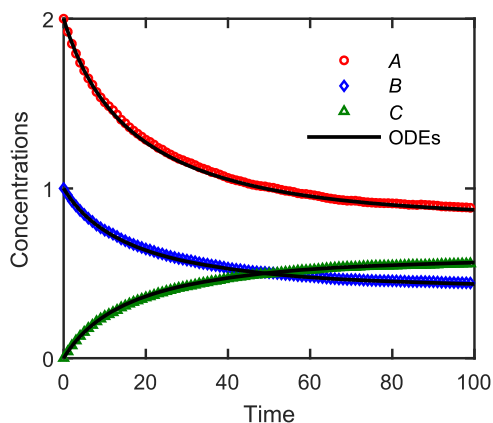


FIG. 3. Simulation results of the average concentrations A (open circles), B (open diamonds), and C (open triangles) are plotted as functions of time for the third-order reaction $2A + B \rightleftharpoons C$. The solid lines are numerical solutions to the rate Eq. (13). Simulation parameters: $L = 10000$, $k_f = 0.01$, $k_b = 0.005$, and $\tau = 1$.

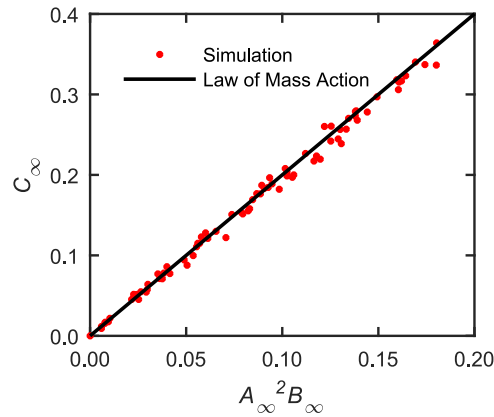


FIG. 4. Simulation results of the equilibrium product concentration C_∞ are plotted as a function of the product of equilibrium reactant concentrations $A_\infty^2 B_\infty$ for the third-order chemical reaction $2A + B \rightleftharpoons C$. The initial concentrations for species A, B, and C are generated randomly between 0 and 0.5. The solid line is the prediction of the law of mass action (Eq. (14)). Simulation parameters: $L = 2000$, $k_f = 0.01$, $k_b = 0.005$, and $\tau = 1$. For each set of initial values, 1000 time steps of simulation are taken, and data from the last 300 time steps are used to calculate the average equilibrium concentrations.

in our simulations are integers, and time steps are discrete. It is reassuring that if we count the particles correctly using appropriate multiplicity factors, our simulations do produce results that agree well with those obtained from continuous equations.

C. Simulation of multiple chemical reactions

In the simple chemical reactions discussed in the previous section, each chemical species can only participate in one reaction, and the reactants and products are of different chemical species. In this section, we extend our simulation to more complicated systems in which a single chemical species can participate in multiple chemical reactions and product molecules can contain the same species as reactant molecules (autocatalytic reactions).

The example we use is the Brusselator, a name given by Tyson³² to a theoretical system of chemical oscillators developed by Lefever, Nicolis, and Prigogine in Brussels.³³ A Brusselator consists of the following four chemical reactions:



We can see that in reaction (15), molecules X are produced by precursors A, but X can also decay to C through reaction (18). In the presence of B, X can produce Y through reaction (16). Reaction (17) is autocatalytic, because X is both a reactant and a product. In this reaction scheme, X is considered an “activator,” because it produces, or “activates,” Y, and Y is considered an “inhibitor,” because it consumes, or “inhibits,” X. For the purposes of our discussion here, we ignore backward reactions and assume that some external agents keep the concentrations of reactants A and B constant.

We can write individual rate equations for each reaction in the Brusselator and combine these to obtain the following overall rate equations for X and Y:

$$\frac{dX}{dt} = k_1 A - (k_2 B + k_4) X + k_3 X^2 Y, \quad (19)$$

and

$$\frac{dY}{dt} = k_2 B X - k_3 X^2 Y. \quad (20)$$

These coupled nonlinear differential equations for X and Y can be solved numerically for given rate constants, values of A and B , and initial conditions.

Before doing this, however, it is helpful to review some general properties of the Brusselator. First, we note that it is possible for X and Y to remain constant with time ($dX/dt = dY/dt = 0$). Setting Eqs. (19) and (20) to zero, we find these steady-state values:

$$X_{ss} = \frac{k_1}{k_4}A, \quad Y_{ss} = \frac{k_2k_4}{k_1k_3} \frac{B}{A}. \quad (21)$$

Whether this steady-state point (also call a fixed point) is stable or not can be analyzed using standard techniques for dynamical systems.¹ Linearizing the system near the fixed point, we find the Jacobian of the system at the fixed point (which involves computing the partial derivatives of dX/dt and dY/dt in Eqs. (19) and (20) with respect to X and Y and evaluating at the fixed point):

$$J = \begin{pmatrix} k_2B - k_4 & k_3(k_1^2/k_4^2)A^2 \\ -k_2B & -k_3(k_1^2/k_4^2)A^2 \end{pmatrix}. \quad (22)$$

Since the determinant of the Jacobian here is positive, the fixed point here is unstable when the trace of the Jacobian is positive. In other words, we expect to see chemical oscillations in the Brusselator when

$$B > \frac{k_4}{k_2} + \frac{k_3k_1^2}{k_2k_4^2}A^2. \quad (23)$$

For a first simulation of the oscillations of a Brusselator, we choose the parameters $A = 3$, $B = 11$, and $k_1 = k_2 = k_3 = k_4 = 1$. Numerical solutions to the rate equations (19) and (20) using these parameters are shown in Fig. 5(a). We can see that both $X(t)$ and $Y(t)$ have periodic sharp changes over very narrow time durations, which can be interesting — and challenging — to simulate.

We can also see in Fig. 5(a) that during most of the time in each cycle, the concentration of the inhibitors (Y) increases steadily with time, indicating that reaction (16) plays a dominant role. The

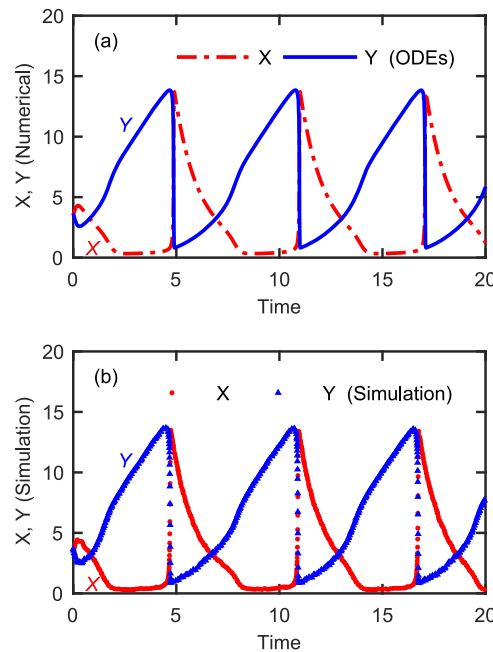


FIG. 5. (a) Numerical solutions to the Brusselator (Eqs. (19) and (20)) with parameters $A = 3$, $B = 11$, and $k_1 = k_2 = k_3 = k_4 = 1$. The initial conditions differ from the steady-state values by a small amount: $X(0) = X_{ss} + 0.5$ and $Y(0) = Y_{ss}$. The dash-dotted line (red) is the activator concentration (X), and the solid line (blue) is the inhibitor concentration (Y), both as functions of time. (b) Simulation results for the Brusselator using the same parameters as in (a), with X represented by circles (red) and Y represented by triangles (blue); every fifth data point is shown. Additional simulation parameters: $L = 1000$ and $\tau = 0.001$. Particles are redistributed randomly among the cells between simulation steps.

concentration of the activators (X), in the meantime, decreases steadily to a very low value and stays there for a while. When the concentration Y reaches a high enough level, however, the autocatalytic reaction (17) starts to occur at a rapid rate, and the system sees a sharp increase in X and an equally sharp decrease in Y within a very narrow time window. The cycle then repeats.

We note that an activator X in the Brusselator can participate in several different reactions. In a real chemical reaction, molecules move about randomly in space, and chemical reactions may occur when molecules come into contact with each other. If there are multiple chemical reactions, there is no specific order in which the reactions should occur. After some experimentation, we have found that the following method produces results that compare well with numerical solutions to rate equations (also see Fig. 6):

1. We randomly deposit molecules of species X and Y into L cells to achieve the desired initial concentrations. The numbers of molecules of species A and B , however, are the same in all cells and are kept constant throughout the entire simulation process.
2. During each time step of interval τ , we simulate chemical reactions in each cell independently. In each cell, we randomly pick a set of four reactions from reactions (15)–(18) and simulate the selected reactions *consecutively*; some of these reactions may be duplicates by chance. For each of these reactions, the rate constant and the multiplicity factor for the particular reaction in question are used to calculate how many times the stoichiometric reaction should happen, and molecule numbers are updated accordingly. (Please see the [Appendix](#) for details.)
3. After all of the selected reactions are simulated in each of the cells, the average concentrations are calculated using $X = (\sum_{i=1}^L N_{X,i}) / (LV_1)$ and $Y = (\sum_{i=1}^L N_{Y,i}) / (LV_1)$, where $N_{X,i}$ and $N_{Y,i}$ are the numbers of X and Y molecules in the i th cell, respectively. These molecules are then redistributed randomly to achieve a Poisson distribution. Steps 2 and 3 are repeated many times until a desired total simulation time has been reached.

Simulation results using the same parameters as those in Fig. 5(a) are shown in Fig. 5(b); the agreement between simulation and numerical results is excellent. The reason that we are able to use four sequential reactions to simulate four simultaneous reactions is because we choose τ such that the probability of any single reaction occurring ($k\tau/V_1^{n-1}$) is so small that the chance of multiple reactions actually taking place in the same cell is almost negligible. In fact, when we choose four fixed reactions in the order (15)–(18), we see no significant difference in the simulation results compared to those presented in Fig. 5(b). Simulating a fixed set of reactions would certainly reduce computation time, but we opted not to do this because of its potential to introduce bias into systems by preferring some reactions to others.

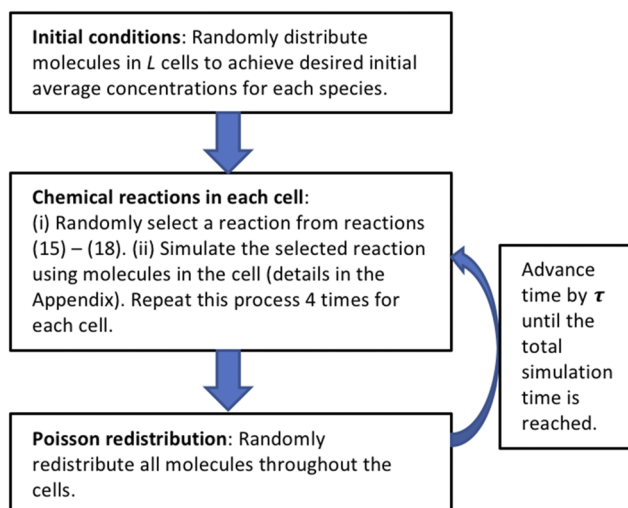


FIG. 6. A graphic illustration of the algorithm used for simulating the Brusselator.

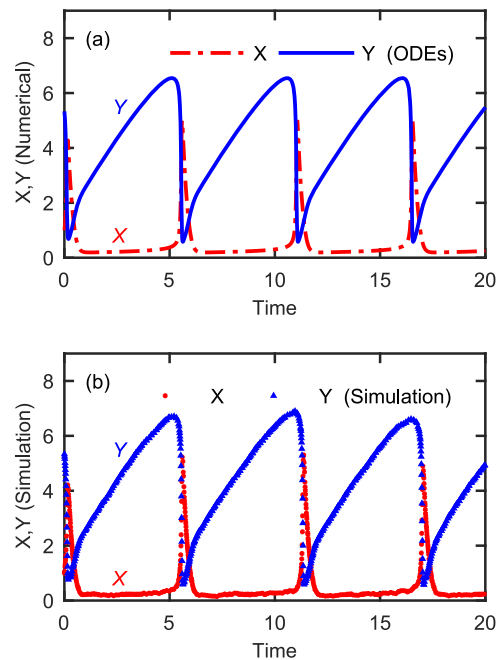


FIG. 7. (a) Numerical solutions to the Brusselator (Eqs. (19) and (20)) with parameters $A = 2$, $B = 4$, $k_1 = 1$, $k_2 = 2$, $k_3 = 3$, and $k_4 = 4$. The dash-dotted line (red) is the activator concentration (X), and the solid line (blue) is the inhibitor concentration (Y), both as functions of time. (b) Simulation results for the Brusselator using the same parameters as in (a), with X represented by circles (red) and Y represented by triangles (blue); every 10th data point is shown. Additional simulation parameters: $L = 1000$ and $\tau = 0.001$. Particles are redistributed randomly among the cells between simulation steps.

Figure 7 shows (a) numerical solutions and (b) simulation results for a Brusselator with unequal reaction rates. We again see very good agreement between simulation results and numerical solutions. It is worth pointing out that our method of choosing reactions randomly with equal probabilities during each time step works even when the reaction rates are different. In Gillespie's method, on the other hand, reaction rate constants affect which reactions are more likely to be chosen. In addition, a chosen reaction will occur in Gillespie's method in a time interval determined by another random number, whereas in our simulation, whether a chosen reaction will occur — and how many reactant molecules will participate in the reaction — is determined by the Monte Carlo simulation in step 4 in the [Appendix](#). In a sense, our method is less efficient than Gillespie's method, because in many cells, no reaction actually occurs in a given time step. But the advantage of our method is that we can use it to explore spatial parameters by associating different cells with different spatial coordinates, whereas Gillespie's method does not have any “built-in” spatial variables.

III. APPLICATION TO SMALL SYSTEMS AND DISCUSSION

A. The effects of particle distributions

We have shown that our stochastic cellular automaton simulation method can produce results that compare well with numerical solutions to deterministic rate equations. Now we want to address some specific aspects of our simulation.

Firstly, how important is it to maintain a Poisson particle distribution when simulating a homogeneous system? We explore this issue by conducting several computer experiments. Naively, we may expect a flat particle distribution to also accurately mimic a well-stirred chemical system (allowing a maximum variation of 1 among particle numbers in different cells, of course, because particle numbers are integers and average concentrations can be non-integers). However, when we change the Poisson distribution in our simulation in Fig. 5(b) to a flat distribution (but leave all other simulation parameters unchanged), we obtain the results shown in Fig. 8(a), where no chemical oscillations

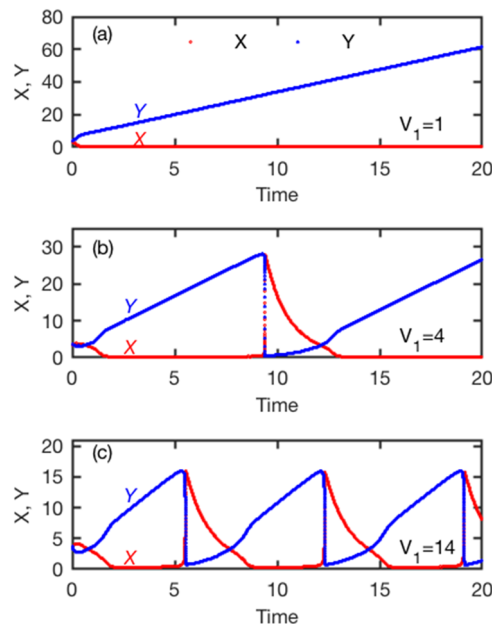


FIG. 8. Simulation results for the Brusselator using a flat distribution of molecules with different cell volumes (V_1). The initial concentrations of molecules are the same in all three figures here, but the numbers of molecules in each cell are proportionally larger for larger values of V_1 . (a) When $V_1 = 1$, there are no chemical oscillations, because the autocatalytic reaction cannot occur after the number of X molecules falls below 2 per cell. For (b) $V_1 = 4$ and (c) $V_1 = 14$, chemical oscillations do occur, but with periods and amplitudes larger than those obtained by solving rate equations numerically (Fig. 5(a)). We do not see such V_1 dependence, however, in simulations using the Poisson distribution.

are observed. The reason is not hard to understand: The autocatalytic reaction (17) requires two X molecules. By forcing a flat distribution, however, no cell can have more than one X molecule after the average concentration X drops below 1, and the autocatalytic reaction (17) can no longer occur. As a result, Y continues to grow due to reaction (16) while X remains constant at a very low value, and chemical oscillations cannot occur.

The problem of not having enough X molecules in a cell for the autocatalytic reaction can be solved, of course, by assuming a larger cell volume V_1 , thereby having more molecules in the cells to maintain the same concentrations. When we set $V_1 = 4$, we indeed see chemical oscillations in our simulation with a flat distribution (Fig. 8(b)), but the amplitude and period are significantly larger than the expected values from numerical solutions to rate equations (Fig. 5(a)). When we further increase V_1 to 14 (Fig. 8(c)), the oscillation amplitude and period are closer to the expected values, although they are still a bit larger. If we use a Poisson distribution with these increased values of V_1 , however, we do not see such V_1 dependence and continue to obtain consistent simulation results that all agree with numerical solutions to rate equations. This is an indication that the Poisson distribution is appropriate for capturing the essence of a homogeneous system from a statistical point of view, even when the numbers of molecules in each cell are very low. In other words, the statistics work correctly when the *probability* of finding a molecule in any cell is the same for all cells (Poisson distribution), but this does not necessarily mean that every cell actually has nearly the same number of molecules (flat distribution). The distinction between these two distributions disappears when large numbers of molecules are present in each cell, which explains why the simulation error using flat distributions decreases when V_1 becomes larger.

We have also tried to redistribute particles after each simulation step using lognormal distributions. A lognormal distribution is asymmetric, with a long tail towards the high-number end. This long tail ensures that the autocatalytic reaction, unlike the simulation using a flat distribution discussed in the previous paragraph, will occur in some cells during the period when X is small. Figure 9 shows the results using lognormal distributions with different variances, and these results deviate significantly from those using a Poisson distribution (Fig. 5(b)), even though all other simulation parameters

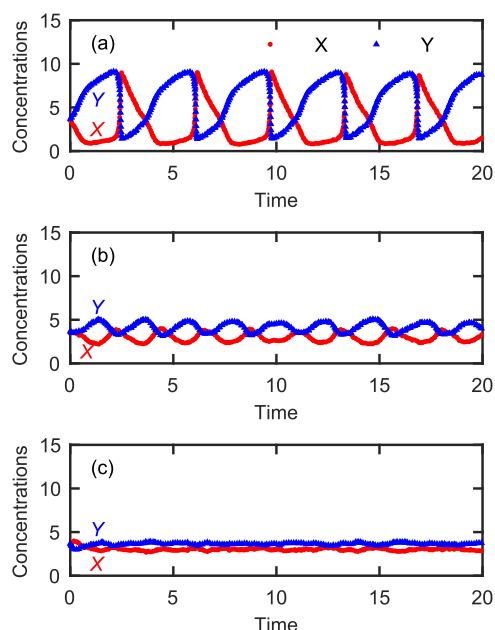


FIG. 9. Simulation results for the Brusselator using a lognormal redistribution of particles after each time step. Simulation parameters are identical to those in Fig. 5(b), and every 10th simulation data point is shown. (a) The variance of the lognormal distribution is set to 1. The amplitude and period of chemical oscillations are smaller than those in Fig. 5(b). (b) The variance of the lognormal distribution is set to 2. Chemical oscillations are further suppressed. (c) The variance of the lognormal distribution is set to 3. Chemical oscillations are destroyed.

are the same. We can see that for a variance of 1, stable chemical oscillations are produced with a lognormal distribution, but the period and the amplitude of the oscillations are smaller than those in Fig. 5(b), where the distribution is Poisson. When the variance is set to 2, chemical oscillations are still visible, but the amplitude and period are even smaller and more irregular. A variance of 3, however, almost completely destroys chemical oscillations, although the fluctuations in X and Y seem out of phase. Thus, our computer experiments indicate that of the particle distributions we have tested, only the Poisson distribution produces accurate chemical oscillations in our cellular automaton simulation of small systems. Interestingly enough, however, using a lognormal redistribution after each simulation step does not alter simulation results of reactions that reach equilibrium (reactions in Section II B).

B. Random walks between simulated chemical reactions

There are, of course, other ways to achieve Poisson distributions besides reassigning particle positions after each simulated reaction step. In a real chemical system, molecules undergo diffusion motion, which can be simulated using random walks (RWs). For each RW step of our simulation, every particle has the same 1/3 probability to move to the cell on the left, to move to the cell on the right, or to stay in the same cell; periodic boundary conditions are applied.³⁴

Figure 10 shows simulation results when particles are allowed to undergo 5 and 500 RW steps at the end of each chemical reaction time step. After completing the RW steps for each time step, we collect a set $\{N_{X,i}\}$ of X molecule counts in each of the L cells and a set $\{N_{Y,i}\}$ of Y molecule counts and calculate the Pearson correlation coefficient between the two sets. From Fig. 10, we can see that chemical reactions cause X and Y particle numbers to become correlated, but more RW steps help to reduce this correlation and produce better chemical oscillations. With 5 RW steps, the average X - Y correlation coefficient in Fig. 10(a) is -0.44, and chemical oscillations are greatly suppressed. In Fig. 10(b), we can see that even though 500 RW steps produce simulation results very close to those obtained by randomly reassigning particle positions (Fig. 5(b)), pronounced negative X - Y correlations are still evident during the sharp population reversal periods, with peak values of about -0.8.

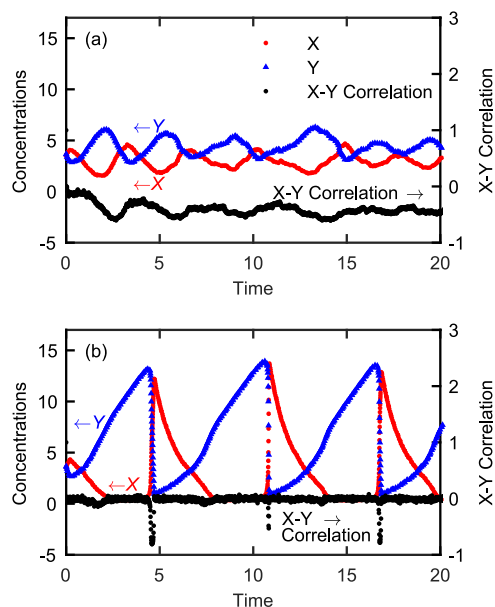


FIG. 10. Simulation results for the Brusselator using random walk (RW) redistribution of particles: (a) 5 RW steps and (b) 500 RW steps after each time step. The correlation coefficient between the numbers of X and Y molecules in each cell is also plotted. Every 10th simulation data point is shown. All simulation parameters are the same as those in Fig. 5(b). Chemical reactions can cause a strong X-Y correlation that can suppress chemical oscillations. This X-Y correlation can be reduced by taking more RW steps for the particle redistribution. The correlation, however, remains strong during the sharp population reversal periods, even with 500 RW steps.

This negative X-Y correlation indicates that in cells with fewer X molecules, there are more Y molecules, and vice versa. After reaction (16), a cell will have less X and more Y, and after reaction (17), a cell will have less Y and more X, both causing negative X-Y correlations. RW steps tend to randomize particle distributions and therefore reduce X-Y correlations. During the sharp population reversal periods, however, this correlation seems difficult to erase, even with 500 RW steps. A closer examination of the distributions of X and Y molecules during this transition period reveals the presence of spatial inhomogeneities with characteristic length scales that can reach a few hundred cells. A few hundred RW steps are obviously insufficient for erasing X-Y correlations under these circumstances.

Our simulation results indicate that a strong correlation between the inhibitor and activator populations plays a role in suppressing chemical oscillations. This, however, is not the only factor. Chemical oscillations are greatly suppressed by lognormal redistributions of particles (Section III A), even though the X-Y correlation is also statistically zero there. Since simulating diffusion by RW is more computationally time-consuming, we argue that reassigning particle positions using a Poisson distribution is a more efficient way to simulate a homogeneous system.

C. The robustness of chemical oscillations in small systems

After gaining some insight into factors that suppress chemical oscillations in the Brusselator in our simulation, we now examine how robust the oscillations are when the size of the system is reduced. Figure 11 shows results of simulations using $L = 10$, 100, and 1000 cells. Simulation parameters used in Fig. 11(c) are the same as those in Fig. 5(b), but the total simulation time is longer. We can see that with 1000 cells, simulation results are almost identical to solutions to rate equations. It is quite remarkable that the oscillations remain steady and strong even with 100 cells, although some peaks decay prematurely, resulting in incomplete cycles (“misfirings”), as seen in Fig. 11(b). These misfirings, however, do not seem to affect the height or the duration of the subsequent full peaks. With only 10 cells, we still see strong oscillations, but the period and amplitude become more irregular (Fig. 11(a)).

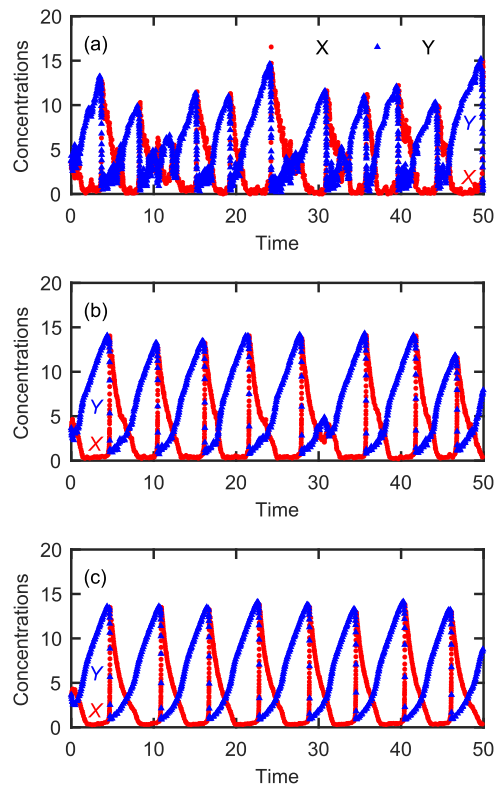


FIG. 11. The dependence of chemical oscillations on the number of cells used in the Brusselator simulation: (a) $L = 10$, (b) $L = 100$, and (c) $L = 1000$ cells. Simulation parameters are $A = 3$, $B = 11$, $k_1 = k_2 = k_3 = k_4 = 1$, and $\tau = 0.001$. When $L = 10$, large fluctuations are seen in oscillation amplitudes and periods. When $L = 100$ and 1000 , oscillations are more stable and match solutions to rate equations, although occasionally some peaks do not reach their expected values, resulting in misfirings (as seen in (b)).

Because of computational power limitations, computer simulations are often carried out on small systems, and it is a common practice to average simulation results of small systems to gain better statistics, with the expectation that the averaged results will more closely resemble the behavior of larger systems. Let $\langle X \rangle$ and $\langle Y \rangle$ be the average values of X and Y , respectively, over many simulations of the same system. Figure 12 shows $\langle X \rangle$ and $\langle Y \rangle$ from 1000 simulations for (a) a 10-cell system and (b) a 100-cell system. Both $\langle X \rangle$ and $\langle Y \rangle$ decay to constant values after a few oscillations, and the functional forms of the curves have little resemblance to those in Fig. 11. Other researchers have reported similar damped chemical oscillations in small systems after averaging over several thousand simulations.³⁵ We argue, however, that the meaning of the mathematical exercise of averaging for chemical oscillators needs careful examination: The damping of the oscillations is caused by destructive superposition due to fluctuations in the periods of the oscillations and the phase shifts that occur after misfirings, and not by the lack of oscillations, which are strong at all times, as seen in Fig. 11.

To further investigate the misfirings, we ran a simulation for $L=100$ cells for a total time duration of 4000 units and found 198 misfirings. We then divided the total time duration into 160 smaller consecutive intervals of $\Delta t = 25$ units each and counted the number of misfirings in each of these intervals. Figure 13 is a histogram of the results, with the misfiring counts in an individual time interval on the horizontal axis and the number of times these counts occur over the total time duration on the vertical axis. Since we have an average of $198/160 = 1.24$ misfirings per 25-unit time interval, we can also use Eq. (1) to calculate the predicted frequency of misfirings according to Poisson statistics. We see in Fig. 13 that our simulation data fit very well to a Poisson distribution, indicating that these misfirings are random events. This also means that we can find the probability of misfirings, but cannot predict precisely when they will actually occur. It is not a coincidence that it takes about 20

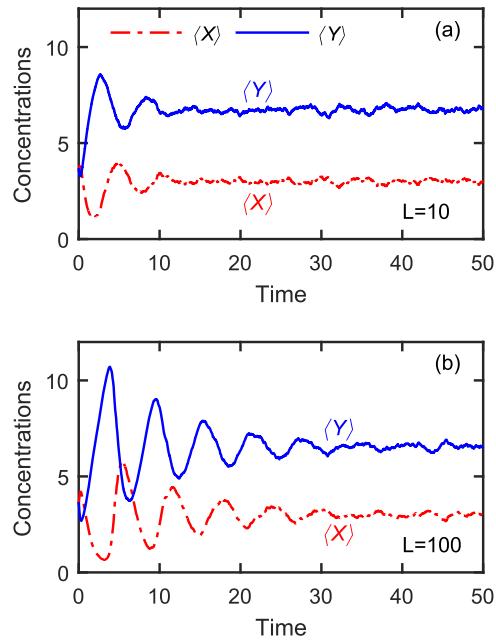


FIG. 12. The average of concentrations over 1000 Brusselator simulation runs for (a) $L = 10$ and (b) $L = 100$ cells. The solid lines (blue) represent the average concentration of the inhibitors $\langle Y \rangle$, and the dash-dotted lines (red) represent the average concentration of the activators $\langle X \rangle$. The decay in the averages is due to the temporal fluctuations in the phase of the oscillations and the misfirings, and is not due to the lack of oscillations in each individual simulation run (see Fig. 11). Additional simulation parameters are $A = 3$, $B = 11$, $k_1 = k_2 = k_3 = k_4 = 1$, and $\tau = 0.001$.

units of elapsed time for the averaged oscillations in Fig. 12(b) to die out, because we expect about one misfiring during that amount of time for a system of $L = 100$ cells.

The inset in Fig. 13 shows that the frequency of misfirings decreases when the system size L increases. The linear fit to the data seems to indicate that there should be no misfirings for L greater than ~ 1400 , and indeed we do not see misfirings in our simulations of 2000-cell systems.

Instead of averaging over many simulation runs, a more appropriate method for analyzing the behavior of chemical oscillators in small systems is to compile histograms of amplitudes and periods, as shown in Fig. 14. The results for (a) $L = 10$, (b) 100, and (c) 1000 cells are shown in the first,

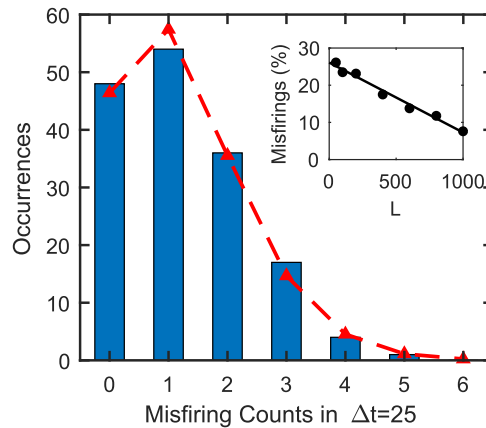


FIG. 13. A histogram of the number of misfirings in $\Delta t = 25$ intervals for $L = 100$ cells. During a total simulation time of 4000, 198 misfirings are observed. The predictions by Poisson statistics (Eq. (1)) are also plotted (red solid triangles). The inset shows the percentage of misfirings as a function of system size L . Additional simulation parameters are $A = 3$, $B = 11$, $k_1 = k_2 = k_3 = k_4 = 1$, and $\tau = 0.001$.

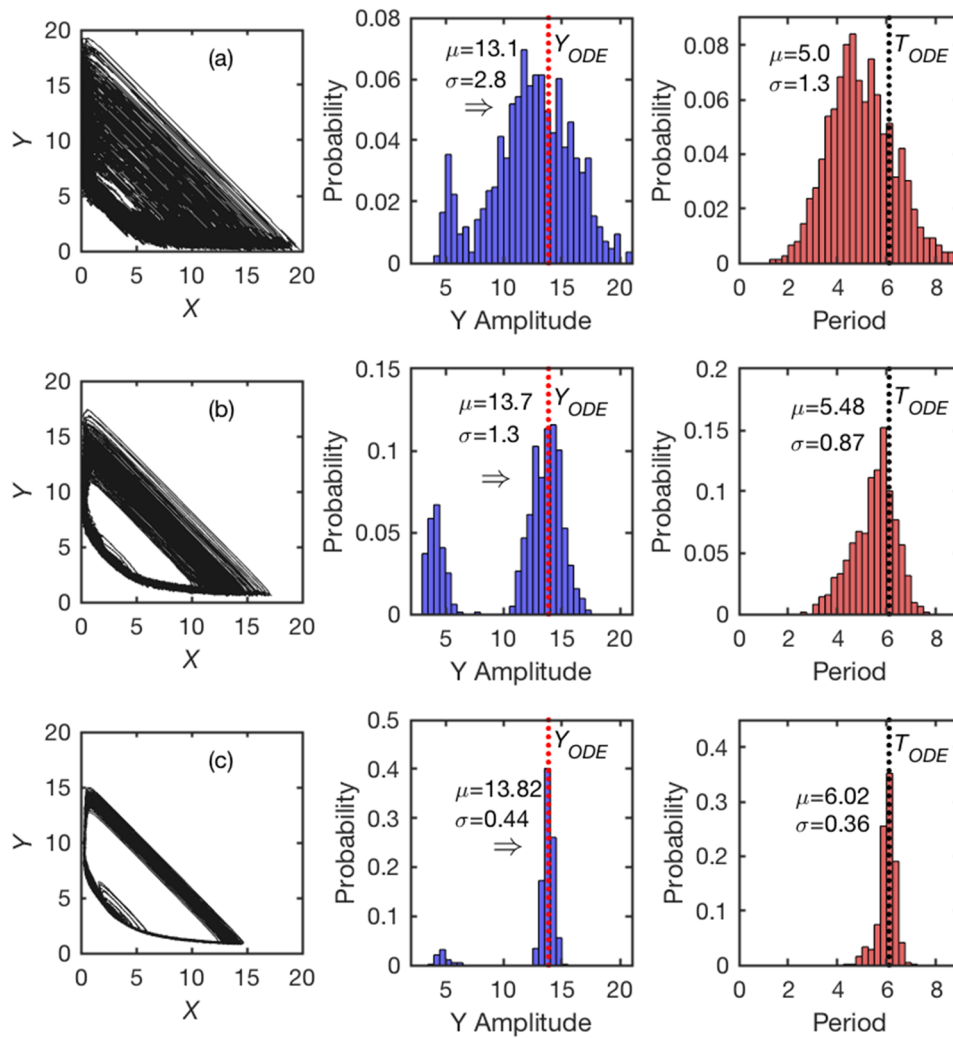


FIG. 14. Effects of Brusselator system size on fluctuations in amplitudes and periods. The top (a), middle (b), and bottom (c) rows are for $L = 10$, 100, and 1000 cells, respectively. The three figures in the left column show the average inhibitor concentration Y versus the average activator concentration X during a time period of 1000 units. The three figures in the middle column are histograms of the amplitudes of Y ; the clusters on the left of these figures are due to misfirings. The three figures in the right column are histograms of the periods (excluding misfirings). The mean values and the standard deviations of full cycles in the simulations are listed in these figures. The vertical dotted lines indicate the amplitude ($Y_{ODE} = 13.84$) and period ($T_{ODE} = 6.10$) of numerical solutions. Additional simulation parameters are $A = 3$, $B = 11$, $k_1 = k_2 = k_3 = k_4 = 1$, and $\tau = 0.001$.

second, and third rows, respectively. In the Y amplitude histograms, we can see two clearly distinct clusters, with the smaller clusters on the left due to misfirings.³⁶ The period histograms are plotted in the column on the right, with the periods associated with misfirings excluded. The means and standard deviations of the amplitudes and periods for full cycles are listed in the figures, as well. The amplitude ($Y_{ODE} = 13.84$) and period ($T_{ODE} = 6.10$) obtained by numerically solving the Brusselator rate equations (Eqs. (19) and (20)) are indicated in the histograms using dotted vertical lines for comparison. In the column on the left, we have included trajectories of Y -vs.- X in phase space, which also show larger fluctuations in smaller systems.

When $L = 1000$ cells, the mean values for amplitudes (13.82) and periods (6.02) produced by our simulations are almost identical to Y_{ODE} and T_{ODE} . When $L = 10$ cells, even though the oscillations look irregular in Fig. 11(a), the mean values for amplitudes (13.1) and periods (5.0) are still reasonably close to Y_{ODE} and T_{ODE} . When $L = 100$ cells, the oscillations are very robust despite the occasional misfirings. It is interesting that the misfirings for all three of the different cell numbers

have a characteristic amplitude of about 5. It is also interesting that the period histograms for $L = 100$ and $L = 1000$ in the column on the right in Fig. 14 look asymmetric, as if they might be a superposition of two Gaussians of different widths. In the future, we would like to investigate this further to gain a better theoretical understanding of this asymmetry.

Another interesting observation is that during the phase of the oscillation when the number of inhibitors (Y) increases steadily and the number of activators (X) is low, the average concentration of X is only about 0.34. According to the Poisson distribution (Eq. (1)), the probability of finding a cell with two X molecules (the minimum number for the autocatalytic reaction to occur) is only about 4%. This means that statistically speaking, only four cells out of 100 have enough X molecules to support the autocatalytic reaction during this phase. As we have discussed earlier, the sharp chemical population reversals and chemical oscillations cannot occur without the autocatalytic reaction. It is quite remarkable that chemical oscillations are still robust in our simulations of such small systems.

IV. CONCLUSION

We have developed a stochastic cellular automaton method to simulate multi-step chemical reactions, and our simulation results agree extremely well with analytical and numerical solutions to deterministic rate equations. We have conducted detailed analyses of the effects of fluctuations and particle spatial distributions on chemical oscillations in small systems, obtaining oscillations in the Brusselator even when average molecule numbers per cell in simulations fall significantly below 1 and when systems contain as few as 100 cells. This is quite remarkable considering that rate equations are usually written for systems with particle concentrations measured in moles per liter.

Of the different particle distribution schemes we have studied, we have found that only the Poisson distribution produces correct chemical oscillations in small systems. In addition, we have seen that chemical reactions cause the numbers of particles of different chemical species in each cell to become correlated, which also suppresses oscillations. Random walks help to reduce these correlations, but this is particularly difficult to do during the sharp activator-inhibitor population reversal period due to the development of large-scale particle inhomogeneity. Since our goal is to simulate homogeneous systems in the present work, we opt to redistribute particles after simulated reactions in each time step to reduce computation time. This approach is valid as long as diffusion of chemicals in the system is fast or the system is well stirred.

We have observed incomplete oscillations (misfirings) in our simulations of small systems, which are triggered by stochastic events but which have no effect on oscillations that reach full height at later times. Since aberrations such as misfirings are common in oscillations in biological systems, we believe our method provides a useful tool to further investigate these aberrations.

Our work has shown that system sizes play an important role in the behavior of chemical oscillations in small systems and that the characteristics of the oscillations are not “scalable,” meaning that one cannot produce behaviors of larger systems by averaging over smaller systems. This means that one needs to be very careful when choosing theoretical tools to study biochemical oscillators, where particle numbers are small and particle distributions are inhomogeneous. We further argue that cellular automaton simulations, which can incorporate these kinds of inhomogeneity, have advantages over the more traditional rate equations and Gillespie’s method, where spatial variations are not taken into account.

Compared to other simulation methods, our method is conceptually straightforward and has a number of notable differences and advantages. For example, when Gillespie’s algorithm is extended to reaction-diffusion systems, uneven time steps for all subvolumes need to be sorted, and only one reaction occurs in one subvolume in each time step. In our method, we use fixed time steps, and as a result, we do not need to sort time steps, and multiple chemical reactions are able to take place simultaneously in all cells at each time step. By taking advantage of this independence of chemical reactions in each cell, we can vectorize computer code to improve computational efficiency,³⁷ even though code optimization has not been the main focus of the current work. This cellular automaton approach is also very convenient for allowing us to take spatial inhomogeneity into consideration.

In addition, our simulation method offers concrete interpretations of the meanings of chemical rate equations on a molecule-by-molecule basis, which is helpful to students of chemistry, biology, and physics.

We want to point out that despite the fact that our simulation method can produce results that agree well with numerical solutions to chemical rate equations, numerical methods have an advantage in computational efficiency over any Monte Carlo method, ours included. The particular merit of computer simulation is revealed when fluctuations due to small system size are important, because numerical solutions to rate equations are intrinsically deterministic and are therefore not appropriate for stochastic systems.

In the future, we hope to expand our work to study the effects of particle spatial inhomogeneity and diffusion in greater depth. We would also like to explore other reaction-diffusion systems, including systems with more complex oscillators that can be used to model real biological systems. We believe that our method is a useful tool for studying the effects of diffusion, particle distribution, and stochastic events in small systems, especially those of biological nature.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for sample computer code used in our simulations, please see our MATLAB that produces a simulation of the third-order reaction $2A + B \rightleftharpoons C$ (Fig. 3). The numerical solutions are obtained using the MATLAB command `ode45`, and it takes about 3.6 seconds to run this simulation on a MacBook Pro using the parameters specified in the caption of Fig. 3.

ACKNOWLEDGMENTS

We thank S. Fraden, M. F. Hagen, M. M. Norton, and I. Hunter of Brandeis University for many helpful discussions on oscillatory chemical systems and on simulation methods. Additionally, Y. H. is grateful to the Physics Department of Brandeis University for its hospitality during her sabbatical leave.

APPENDIX: SUMMARY OF THE SIMULATION ALGORITHM

1. Initial configuration

We divide our chemical system into L individual cells and deposit molecules into the cells to achieve desired particle distributions and densities. Chemical reactions are simulated in each cell independently using only molecules present in the given cell as reactants. We assume that all cells have the same cell volume V_1 . (See Fig. 15.)

2. Selection of chemical reactions to simulate in a time step

If there are M different possible chemical reactions (e. g., $M = 4$ for the Brusselator), we randomly select a set of M reactions to simulate in each cell in a time step. The reaction sets in different cells are selected independently, and some of the selected reactions in the same cell may be duplicates by chance. We then simulate each of these reactions sequentially in each cell using steps 3 and 4 below, updating particle counts after each reaction.

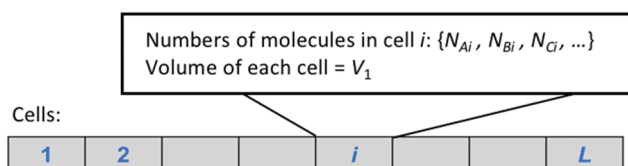


FIG. 15. An illustration of the simulation space.

3. Calculation of multiplicity factors

The multiplicity factor Ω is a product of factors related to the numbers of reactant molecules of different species and is calculated for each reaction in each cell separately. Suppose a reactant species A has a stoichiometric coefficient of 1 and the number of A molecules in the cell is N_A , then the contribution of species A to Ω is simply N_A . Suppose another reactant species B has a stoichiometric coefficient $\nu > 1$ and the number of B molecules in the cell is N_B , then species B's contribution to Ω is $N_B!/(N_B - \nu)!$. For example, the autocatalytic reaction in the Brusselator has stoichiometric coefficients 2 and 1 for reactants X and Y, respectively. The multiplicity factor for this reaction in cell i is $\Omega_i = N_{X,i}(N_{X,i} - 1)N_{Y,i}$, where $N_{X,i}$ and $N_{Y,i}$ are the numbers of molecules of species X and Y in the cell, respectively.

4. Monte Carlo simulation of a reaction

During a time step of duration τ , the probability of the occurrence of a single stoichiometric reaction is equal to $k\tau/V_1^{n-1}$, where k is the rate constant, n is the order of this reaction, and V_1 is the volume of a reaction cell. To determine the number of occurrences of this reaction in the cell, we generate Ω random numbers between 0 and 1, where Ω is the multiplicity factor for this reaction in this cell. We use N_r to denote the number of these random numbers having values less than $k\tau/V_1^{n-1}$. In our simulation algorithm, N_r is the number of times that a stoichiometric reaction will occur, as long as there are enough reactant molecules available; N_r needs to be adjusted otherwise. For example, suppose we have a cell with three X molecules and five Y molecules. Using Step 3, we find that for the autocatalytic reaction in the Brusselator, the multiplicity factor for this cell is $\Omega = 30$. If the Monte Carlo simulation in this step yields $N_r = 1$, then we need two X molecules and one Y molecule as reactants, and the numbers of X and Y molecules after the reaction (which increases the X count by 1 and decreases the Y count by 1) will become 4 and 4, respectively. If, however, the simulation yields $N_r = 2$, then we would need four X molecules and two Y molecules for this stoichiometric reaction to occur twice. Since we only have three X molecules available, we need to reduce N_r to 1. In general, we should choose τ small enough, so that the number of times that N_r needs to be adjusted this way in a simulation run is statistically insignificant.

5. Redistribution of molecules

After simulating all of the selected chemical reactions in all of the cells, simulation time is advanced by a step τ , and particles are allowed to redistribute among different cells. This redistribution can be achieved by reassigning particle positions or by allowing particles to perform random walks using periodic boundary conditions. Steps 2-5 are then repeated for many time steps.

- ¹ S. H. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*, 2nd edn (CRC Press, New York, 2015).
- ² A. Goldbeter, *Adv. Chem. Phys.* **135**, 253 (2007).
- ³ A. M. Zhabotinsky, *Proc. Acad. Sci. USSR* **157**, 392 (1964).
- ⁴ A. N. Zaikin and A. M. Zhabotinsky, *Nature* **225**, 535 (1970).
- ⁵ S. K. Scott, *Acc. Chem. Res.* **20**, 186 (1987).
- ⁶ I. R. Epstein and J. A. Pojman, *An introduction to nonlinear chemical dynamics: oscillations, waves, patterns, and chaos* (Oxford University Press, New York, 1998).
- ⁷ R. Milo and R. Phillips, *Cell biology by the numbers* (Garland Science, New York, NY, 2016).
- ⁸ D. T. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).
- ⁹ D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- ¹⁰ M. A. Gibson and J. Bruck, *J. Phys. Chem. A* **104**, 1876 (2000).
- ¹¹ D. T. Gillespie, *J. Chem. Phys.* **115**, 1716 (2001).
- ¹² Y. Cao, H. Li, and L. Petzold, *J. Chem. Phys.* **121**, 4059 (2004).
- ¹³ J. M. McCollum, G. D. Peterson, C. D. Cox, M. L. Simpson, and N. F. Samatova, *Comput. Biol. Chem.* **30**, 39 (2006).
- ¹⁴ R. Ramaswamy, N. Gonzalez-Segredo, and I. F. Sbalzarini, *J. Chem. Phys.* **130**, 244104 (2009).
- ¹⁵ D. T. Gillespie, A. Hellander, and L. R. Petzold, *J. Chem. Phys.* **138**, 170901 (2013).
- ¹⁶ J. Elf and M. Ehrenberg, *Syst. Biol.* **1**, 230 (2004).
- ¹⁷ T. T. Marquez-Lago and K. Burrage, *J. Chem. Phys.* **127**, 104101 (2007).
- ¹⁸ V. K. Vanag, *Physics-Uspekhi* **42**, 413 (1999).
- ¹⁹ D. Scalise and R. Schulman, in *DNA Computing and Molecular Programming*, edited by S. Murata, and S. Kobayashi (Springer, Kyoto, 2014), pp. 67–83.
- ²⁰ F. Takabatake, I. Kawamata, K. Sugawara, and S. Murata, *New Generat. Comput.* **35**, 213 (2017).
- ²¹ D. H. Zanette, *Phys. Rev. A* **46**, 7573 (1992).

- ²² T. Karapiperis and B. Blankleider, *Physica D* **78**, 30 (1994).
- ²³ R. J. Gelten, A. P. J. Jansen, R. A. van Santen, J. J. Lukkien, J. P. L. Segers, and P. A. J. Hilbers, *J. Chem. Phys.* **108**, 5921 (1998).
- ²⁴ H. C. Kang and W. H. Weinberg, *J. Chem. Phys.* **100**, 1630 (1994).
- ²⁵ H. C. Kang and W. H. Weinberg, *Surf. Sci.* **299**, 755 (1994).
- ²⁶ K. Odagiri and K. Takatsuka, *Phys. Rev. E* **79**, 056219 (2009).
- ²⁷ K. Odagiri and K. Takatsuka, *Phys. Rev. E* **79**, 026202 (2009).
- ²⁸ L. B. Kier, P. G. Seybold, and C. K. Cheng, *Cellular automata modeling of chemical systems: a textbook and laboratory manual* (Springer, Dordrecht, 2005).
- ²⁹ T. Nakanishi, *J. Phys. Soc. Japan* **32**, 1313 (1972).
- ³⁰ G. Grimmett and D. Stirzaker, *Probability and random processes*, 2nd edn (Oxford University Press, New York, 1992).
- ³¹ P. G. Seybold, L. B. Kier, and C. K. Cheng, *J. Chem. Inf. Comput. Sci.* **37**, 386 (1997).
- ³² J. J. Tyson, *J. Chem. Phys.* **58**, 3919 (1973).
- ³³ R. Lefever, G. Nicolis, and I. Prigogine, *J. Chem. Phys.* **47**, 1045 (1967).
- ³⁴ If remaining stationary is not an option, then particles in even-numbered cells, for example, can never end up in odd-numbered cells after an even number of steps in a 1D RW simulation. Since our goal is to achieve a Poisson distribution as quickly as possible, we allow some particles to remain stationary to ensure that particles in a given cell can redistribute to all cells.
- ³⁵ P. Dini, C. L. Nehaniv, A. Egri-Nagy, and M. J. Schilstra, *Biosystems* **112**, 145 (2013).
- ³⁶ Misfirings are identified by using an amplitude that separates the two clusters in the amplitude histograms in Fig. 14. In this work, peaks with amplitudes less than 7 are considered misfirings. Only histograms of the Y amplitudes are shown in Fig. 14; histograms of the X amplitudes are statistically identical to those of Y and are not shown.
- ³⁷ The computer code used in our simulation is written in MATLAB. Please see the [supplementary material](#) section for an example of the code.