

Chapter 2

Effects of Spatial Confinements on Biochemical Reactions

2.1 Chemical Reactions of Low Copy Number

2.1.1 System-size Resonance of a Reduced Model of Belousov–Zhabotinsky Reaction

Introduction

Belousov–Zhabotinsky (BZ) reaction is arguably the most famous example of oscillatory chemical reactions and nonlinear chemistry in general. Variants of BZ reactions are still attracting the interests of scientists even 50 years after its first publication [8]. Nonlinear phenomena may be more sensitive to the stochastic nature of chemical reactions. It is thus natural to subject BZ reactions to low copy number conditions and investigate the influence of noise on the behavior of the system. In addition, the recent development of microfluidic techniques [23] offers unprecedented access to femtoliter-sized reactors where the stochastic effect is expected to be more pronounced than in bulk conditions. This advancement enables the direct comparison between theoretical prediction and experimental observation.

Only a subset of nonlinear chemical reactions can be conveniently used in a microfluidic device for interrogation. Due to the complexity of world-to-chip integration, the content of a microreactor is

normally isolated from the environment. For example, a continuously stirred tank reactor with constant supply and removal of chemicals is not trivial to scale down to femtoliter range. This limits our choice of candidate systems to batch reactions. Furthermore, a convenient method, preferably based on an optical signal, should be available to record the kinetics of the reaction. Finally, it is extremely useful if the reaction mechanism of the reaction is known. On the basis of these consideration, we will focus on a variant of BZ reaction whose core elements are bromate and 1,4-cyclohexanedione (bromate–CHD system). It supports oscillation in bulk mode. Its reaction mechanism has been elucidated at various levels [148–151]. And optical signal, from fluorescence [150] or absorbance [42], is available to facilitate experimental observation. In addition, during the reaction CHD does not release carbon dioxide, which tends to form bubbles and disturbs the system. Figure 2.1 shows the spatial pattern formed by a bromate–CHD system in a Petri dish.

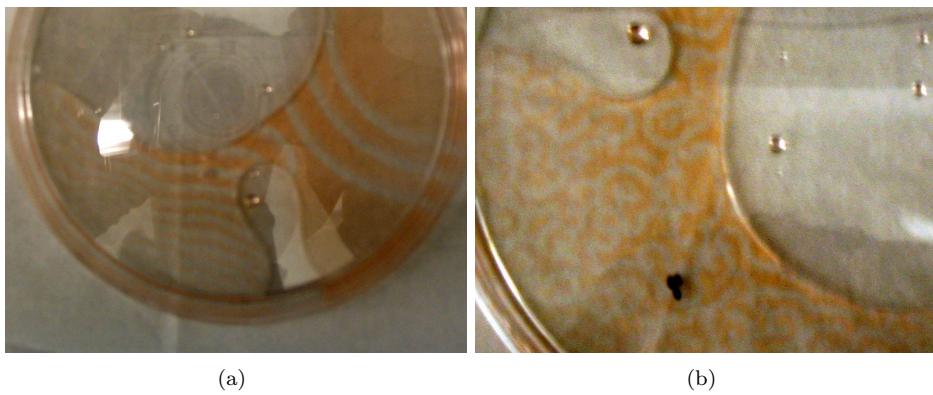


Figure 2.1. Ribbon (a) and spiral (b) patterns formed by bromate–CHD system in a Petri dish. The initial concentrations were $[CHD] = 0.1\text{ M}$, $[KBrO_3] = 0.1\text{ M}$, and $[H_2SO_4] = 1.0\text{ M}$. Ferroin (0.1 mM) was used as indicator, and 0.01 M KBr was added to reduce induction time.

Bromate–CHD system was used in microfluidic channels to generate obstacle-mediated spiral waves [42–44, 90]. The same system was also the subject of a recent experiment showing the effect of stirring rate on the reaction kinetics [172]. The effect of reactor size on this system is yet to be explored in theory or experiments. In this section, we report the effect of stochastic chemical reactions on the dynamical behavior of this system by using Gillespie algorithm and a reduced model of bromate–CHD system (Table 2.1). In particular, we studied the effect of system volume when

Table 2.1. Reduced model of bromate–CHD system [149]

Reaction ^a	Reaction rate expression	Rate constant
$X + Y \rightarrow \text{sink}$	$k_1[X][Y]$	$k_1 = 3.225 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$
$Y \rightarrow X$	$k_2[Y]$	$k_2 = 0.158 \text{ s}^{-1}$
$2X \rightarrow \text{sink}$	$k_3[X]^2$	$k_3 = 4.386 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$
$X + Z \rightarrow 2X$	$k_4[Z][X]^{1/2}$	$k_4 = 86.7 \text{ M}^{-3/2} \text{ s}^{-1}$
source $\rightarrow fZ + Y$	k_5	$k_5 = 1.32 \times 10^{-6} \text{ M s}^{-1}$
$Z \rightarrow X$	$k_6[Z]$	$k_6 = 2.04 \times 10^{-3} \text{ s}^{-1}$

^a $X \equiv \text{HBrO}_2$, $Y \equiv \text{bromide}$ and $Z \equiv 1,4\text{-dihydroxybenzene}$

the system was configured close to the bifurcation point and observed the existence of an optimal volume to support oscillation. The simulation results are expected to motivate the corresponding experiments in microreactors defined using microfluidic techniques (chapters 3 and 4).

Deterministic model

The full model of bromate–CHD system consists of 17 species and 25 reactions [148]. As shown in Table 2.1, it can be reduced to a three-variable model [149] using quasi-steady-state assumption and pool-component approximation. The bifurcation parameter f is essentially the ratio of 1,4-dihydroxybenzene production rate and bromide production rate, and it can be tuned by the initial species concentration.

It is straightforward to obtain the bifurcation diagram according to the deterministic model (Table 2.1). The concentration of X as a function of time reaches either a steady state or an oscillatory state after a short transient period (figure 2.3, black). By plotting the maximum and minimum values of $[X]$ as a function of f , we obtained a phase diagram (figure 2.2) illustrating the Hopf bifurcation between the two possible dynamical behaviors.

Stochastic simulation and system-size resonance

In this section, Gillespie algorithm [40, 109] is used to study the influence of stochasticity of chemical reactions on the system dynamics when $f \approx 1.05$, close to the second bifurcation point.

As shown in figure 2.3(a), the deterministic limit-cycle behavior is largely preserved in all stochastic simulations. The deviation from perfect periodic kinetics is bigger when the system volume is

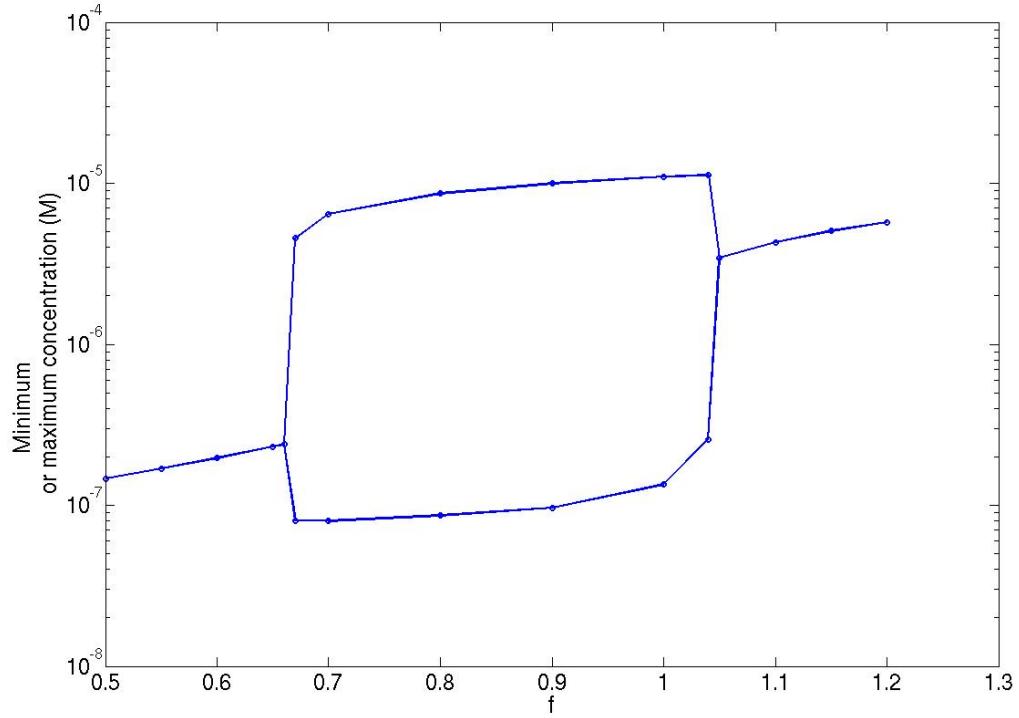


Figure 2.2. Bifurcation diagram based on deterministic simulation of the reduced model. The minimum and maximum concentration of X are plotted as a function of f .

decreased, in agreement with theoretical consideration in a more general context [37, 38]. This observation holds true independently of the value of control parameter f as long as system volume is not too small (compare red curves with black ones in figure 2.3). However, when more noise is present due to the decrease of system volume, the behavior of the system may show qualitatively different feature (compare green or blue curves with black ones in figure 2.3(b) and figure 2.3(c)). In particular, the system may be excited from steady state to limit cycle transiently and exhibit pronounced oscillation.

To quantify the periodicity in these kinetic traces, we consider the power spectral density (PSD) of figure 2.3. The PSD of a time series is defined as the magnitude square of its Fourier transform. All time series being of finite length, a Hamming window is used to reduce the contribution of spurious side peaks. The simulation data has a sampling rate of 1 Hz and around 10^5 points, sufficient to get a converging estimate of PSD. It is divided by a Welch window of 1024 with 50%

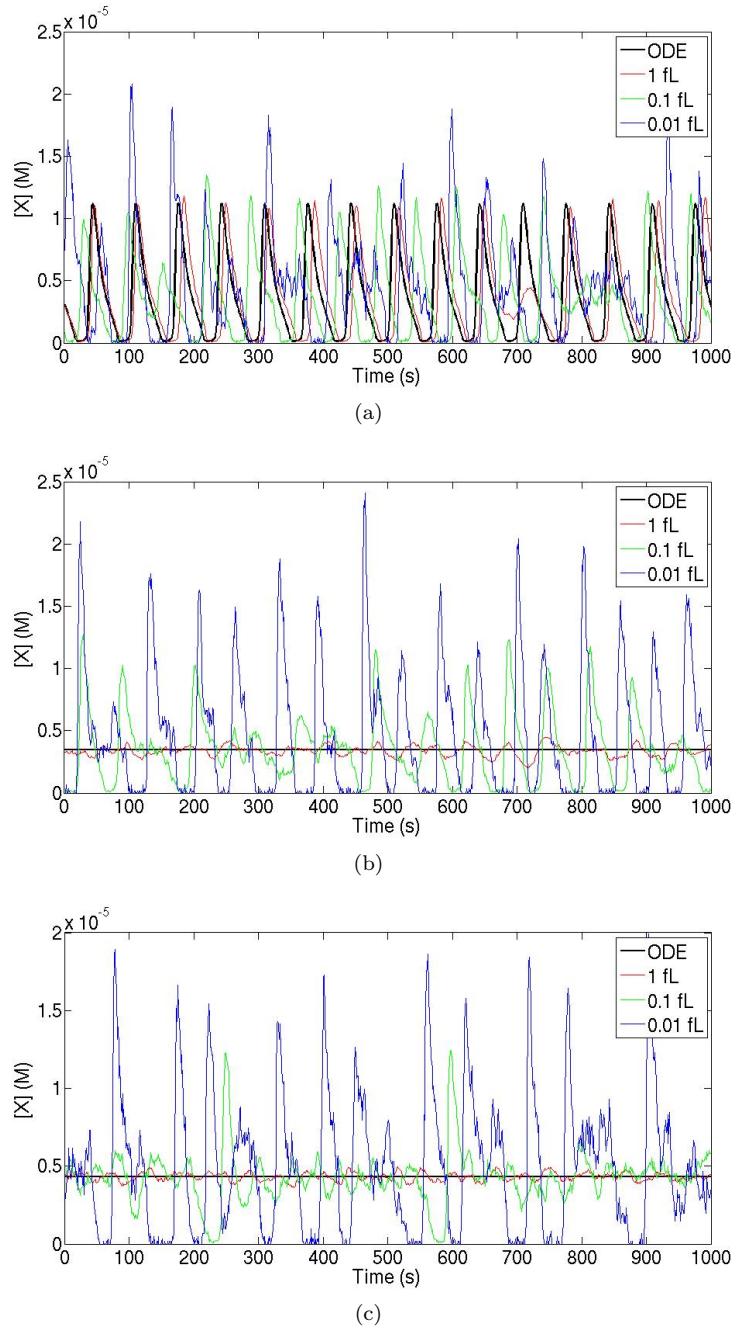


Figure 2.3. The concentration of X from deterministic (black) or stochastic simulation with system volume of 1 fL (red), 0.1 fL (green), and 0.01 fL (blue) at $f = 1.03$ (a), 1.05 (b), and 1.10 (c).

overlap to smooth the final PSD. A signal-to-noise ratio (SNR) is then obtained from the peak of PSD using $SNR \equiv \frac{P(B)}{P(A)} \frac{f_B}{f_C - f_B}$. Here $P(B)$ is the PSD at the signal peak located at f_B , $P(A)$ is the minimum PSD to the left of the peak, and f_C is the frequency which satisfies $P(C) = P(B)/e^{1/2}$ (figure 2.4(b)). SNR reflects the periodicity of the time series in a quantitative way. For example, at $f = 1.05$, the deterministic simulation does not show any periodicity (figure 2.3(b), black), so its PSD (figure 2.4(a)) has no nonzero peaks. With the introduction of stochastic noise, the system undergoes significant periodic fluctuation (figure 2.3(b), green) and there is a clear peak in its PSD (figure 2.4(b)) at 0.015 Hz, the limit-cycle frequency for $f = 1.04$.

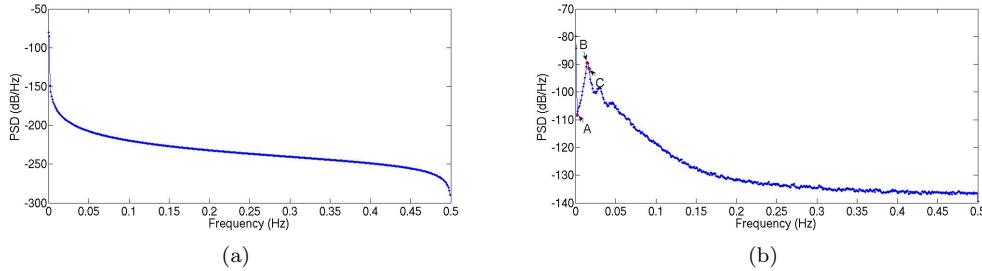


Figure 2.4. The PSD of $[X](t)$ from deterministic (a) or stochastic (b) simulation with $V = 0.1$ fL and $f = 1.05$.

Figure 2.5 summarizes the relation between SNR and system volume at different values of the bifurcation parameter. When the system is inside the limit-cycle region (figure 2.5, red and green), the destructive effect of stochastic noise is reflected in the decrease of SNR with decreasing volume. The sensitivity of SNR to system volume is bigger when f is further away from the bifurcation point mostly due to bigger SNR at large volumes, which is consistent with the intuition that limit cycle is more stable in that region. When the system is slightly outside the limit-cycle region (figure 2.5, blue and black), the noise induces periodic fluctuations with a characteristic frequency close to the neighboring limit-cycle value (Table 2.2). This influence tends to be more significant when system volume is reduced to introduce more noise (figure 2.5, black). When f is very close to the bifurcation point, there exists an optimal volume where SNR is maximized (figure 2.5, blue). This phenomenon has been reported previously in other systems [60–62, 64, 120, 125], and associated with stochastic

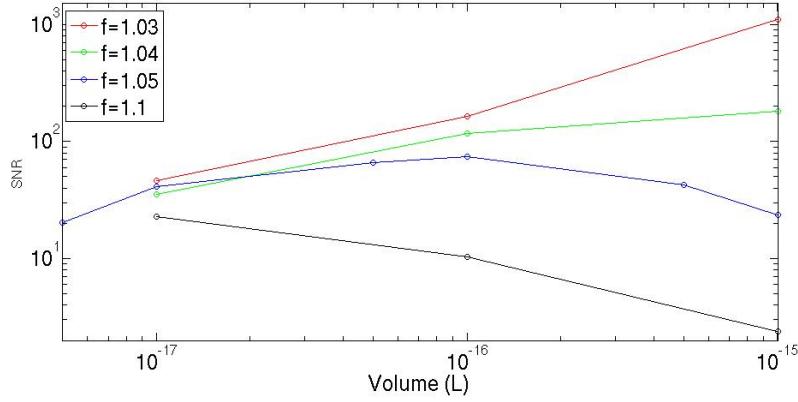


Figure 2.5. SNR as a function of system volume at different values of the bifurcation parameter f by spectral analysis of time series $[X](t)$ from stochastic simulations.

resonance [32] where a system driven by external periodic force is more responsive to noise of a certain magnitude. In the context of chemical reactions, noise is inversely proportional to the square root of system volume [39, 166], so the optimal noise strength translates to an optimal volume. From a mathematical point of view, SNR for small system volumes tends to be less dependent on the value of f , whereas for big system volume, it closely follows the deterministic results. As f increases and crosses the bifurcation point, SNR for big volumes increases significantly, and at some particular f , it must be very close to the SNR for small volumes. At this point, because SNR assumes similar values at big and small volumes, there must exist at least one extremum at intermediate volumes as long as SNR is dependent on volume.

Table 2.2. The peak frequency of PSD as a function of volume and bifurcation parameter

Volume (L)	Peak frequency (Hz)			
	$f = 1.04$ (deterministic)	$f = 1.04$ (stochastic)	$f = 1.05$ (stochastic)	$f = 1.1$ (stochastic)
5×10^{-18}	0.015	—	0.016	—
1×10^{-17}	0.015	0.016	0.016	0.016
5×10^{-17}	0.015	—	0.016	—
1×10^{-16}	0.015	0.016	0.015	0.016
5×10^{-16}	0.015	—	0.015	—
1×10^{-15}	0.015	0.015	0.015	0.017

Finally, even when the system volume is big enough to prevent the appearance of significant

fluctuation around the steady state (figure 2.3(b) and 2.3(c), red), there is still a well-defined peak in the PSD with $\text{SNR} > 1$. PSD being the Fourier transform of the correlation function, a peak in PSD implies that the system relaxes to the steady state in an oscillatory fashion [37]. A high SNR means the relaxation time is longer than the oscillation period. Considering that the peak frequency is still similar to the neighboring limit-cycle frequency, one could extract some information regarding the limit cycle region from a spectral analysis of the fluctuation around the steady state.

Conclusion

In this section, we analyze a reduced model of bromate–CHD system which supports oscillatory kinetics in bulk under appropriate conditions. Comparison between deterministic and stochastic simulations reveals the effect of random noise on the dynamical features of the system. In general, noise adds a perturbation to the deterministic trajectory. When the system is configured very close to the bifurcation point, chemical reaction noise may transiently push the system over the bifurcation point, and one would observe limit-cycle behavior when steady state is predicted by the deterministic model. Even when the noise is not strong enough to change the dynamics qualitatively, it induces fluctuation whose power spectrum contained valuable information on the nearby limit-cycle region. Moreover, when the bifurcation parameter assumes some particular value very close to the bifurcation point, there exists an optimal system volume for the excitation of steady state by stochastic noise. This phenomena is closely related to stochastic resonance.

This study may lead to the experimental verification of the conclusions drawn above from numerical investigations. With the advancement of microfluidic techniques, it is straightforward to observe the bromate–CHD reaction in femtoliter reactors with tunable volume [68, 116] and study the influence of chemical noise on reaction kinetics.

2.1.2 Stochastic Defocusing of Peroxidase–Oxidase System in Response to Initial Conditions

Introduction

Chemical reactions are in nature a random process, but the stochasticity does not manifest itself in traditional bulk experiments dictated by the ensemble average of the behavior of a large number of molecules [166]. Chemical reactions essential to biological systems are conveniently studied in terms of the time evolution of the components' concentrations, and lots of useful information has been extracted and used to construct complicated metabolic [5] and genetic [31, 132] oscillators *in vivo*. Real cellular environment, however, differs from *in vitro* experiments in two fundamental aspects. First, in many cases only a small number of molecules are involved in biochemical interactions [110]. For example, in a highly compartmentalized cell, reactions often take place in a small volume on the order of femtoliters, so the copy numbers of involved molecules are much smaller than that in bulk experiments. Small copy numbers are also characteristic of genetic materials [89]. In addition, cellular environment is extremely crowded, so diffusion of molecules is much slower than in dilute solution [99].

An intuitive way to address the first problem is to study biochemical reactions in microreactors. With the development of “lab on a chip,” fabrication of microreactors have been more and more accessible to chemists and biologists [23, 66]. Much effort has been put into understanding the physics of fluid at micrometer scale [133], and integrating and parallelizing processes in microfluidic chips [93]. High throughput, low reagent consumption, and fast turnout rate are among the most often cited advantages of microfluidics. More relevant to this work, it is possible to fabricate microreactors comparable in size to real cellular compartments and thus more relevant to the investigation of biochemical reactions in their native conditions. The deterministic approach of reaction kinetics based on law of mass action should give way to its stochastic counterpart when it is used to predict or interpret experimental results in microreactors. For a well mixed microreactor, the stochasticity can be accounted for by using Gillespie's stochastic simulation algorithm or its variants [40]. Although, as mentioned above, diffusion in cellular environment is hindered due to molecular crowding, there

is some evidence that only minor modification is required to take crowdedness into account [41]. So for this work, Gillespie algorithm [40, 109] is used to simulate biochemical reactions in a confined volume of femtoliter scale.

The inherent randomness of chemical reactions adds noise to the deterministic evolution of a biochemical system. The effect is more significant when a nonlinear system is of interest. Noise-induced phenomena have been explored in both abstract [37, 59] and realistic chemical reaction networks [6, 47, 171]. Peroxidase–oxidase (PO) oscillator [103] is a well-studied nonlinear system of biological importance. A recent review summarizes the current models for PO system [12, 56, 119], and the list of relevant reactants are still being updated. Although an abstract model was proposed around 30 years ago [103], realistic models have been the interest of different groups to explain various dynamical behavior observed in PO systems in appropriate part of its parameter space [9, 11, 12, 104, 119]. By taking stochastic effects into consideration, it is possible to predict the behavior of the PO system in a confined space and compare it against experimental results.

We compared the asymptotic behavior of the PO system as a function of initial H_2O_2 concentration [56] using stochastic and deterministic simulation methods. The results demonstrated the effect of the randomness of chemical reactions even when the system volume is as big as 100 fL. It may be compared to stochastic defocusing [72] reported for other systems. This work is expected to prompt future experimental investigations to quantitatively test the prediction of numerical simulation.

Deterministic model

PO system can manifest oscillatory or steady-state behavior under different conditions. For example, it was observed that initial concentration of H_2O_2 determined the asymptotic behavior of the PO system (i.e., steady state or periodic oscillation). To be particular, deterministic simulation of the BFSO model [56] showed that when $[\text{H}_2\text{O}_2]_0 < c_0$, the system oscillated with a defined period, and when $[\text{H}_2\text{O}_2]_0 > c_0$, steady state was eventually achieved. The critical concentration c_0 was determined to be $\approx 0.29998 \mu\text{M}$ from a reaction network (BFSO model) composed of 14 reactions and 11 species (Table 2.3). We take this as a case study to explore the stochastic effect of chemical

Table 2.3. BFSO model for PO system [56]

Reaction	Rate constant	Unit
$\text{NADH} + \text{O}_2 + \text{H}^+ \rightarrow \text{NAD}^+ + \text{H}_2\text{O}_2$	10.0	$\text{M}^{-1} \text{s}^{-1}$
$\text{H}_2\text{O}_2 + \text{Per}^{3+} \rightarrow \text{compound I}$	1.8×10^7	$\text{M}^{-1} \text{s}^{-1}$
$\text{compound I} + \text{NADH} \rightarrow \text{compound II} + \text{NAD}\bullet$	4.0×10^5	$\text{M}^{-1} \text{s}^{-1}$
$\text{compound II} + \text{NADH} \rightarrow \text{Per}^{3+} + \text{NAD}\bullet$	2.6×10^5	$\text{M}^{-1} \text{s}^{-1}$
$\text{NAD}\bullet + \text{O}_2 \rightarrow \text{NAD}^+ + \text{O}_2^-$	2.0×10^7	$\text{M}^{-1} \text{s}^{-1}$
$\text{O}_2^- + \text{Per}^{3+} \rightarrow \text{compound III}$	1.7×10^7	$\text{M}^{-1} \text{s}^{-1}$
$2\text{O}_2^- \xrightarrow{2\text{H}^+} \text{H}_2\text{O}_2 + \text{O}_2$	2.0×10^7	$\text{M}^{-1} \text{s}^{-1}$
$\text{compound III} + \text{NAD}\bullet \rightarrow \text{compound I} + \text{NAD}^+$	6.0×10^7	$\text{M}^{-1} \text{s}^{-1}$
$2\text{NAD}\bullet \rightarrow \text{NAD}_2$	5.6×10^7	$\text{M}^{-1} \text{s}^{-1}$
$\text{Per}^{3+} + \text{NAD}\bullet \rightarrow \text{Per}^{2+} + \text{NAD}^+$	1.8×10^6	$\text{M}^{-1} \text{s}^{-1}$
$\text{Per}^{2+} + \text{O}_2 \rightarrow \text{compound III}$	1.0×10^5	$\text{M}^{-1} \text{s}^{-1}$
$\rightarrow \text{NADH}$	7×10^{-8}	M s^{-1}
$\rightarrow \text{O}_2$	5.28×10^{-8}	M s^{-1}
$\text{O}_2 \rightarrow$	4.4×10^{-3}	s^{-1}

reactions.

The transition is quantified by plotting the probability of limit cycle behavior with respect to $[\text{H}_2\text{O}_2]_0$. As expected, deterministic simulation predicts that the transition from limit cycle to steady state is discontinuous, and the system's behavior is not well defined if $[\text{H}_2\text{O}_2]$ starts exactly from c_0 . The behavior of the system starting away from c_0 ends up to be either steady state or periodic oscillation after a short transient period. As is shown in figure 2.6 (red lines), deterministic simulation shows that PO system will assume either steady state or oscillatory state depending on the initial concentration of H_2O_2 even if it is very close to c_0 (figure 2.6(c)).

Stochastic simulation and result

The transition around c_0 should be sharp but continuous in reality. In this case stochastic simulation is invoked to resolve the discontinuity by taking stochastic effect into consideration.

Intuitively, the concentrations of involved species in the system are fluctuating around the deterministic results in stochastic simulations as shown in figure 2.6. However, when the system is initiated using $[\text{H}_2\text{O}_2]_0 \approx c_0$, the system's fate lost its predictability. Either steady state or oscillation may be the ultimate choice (compare blue and green lines in figure 2.6). Interestingly, no transition between the two choices was observed within a realistic period of time (data not shown).

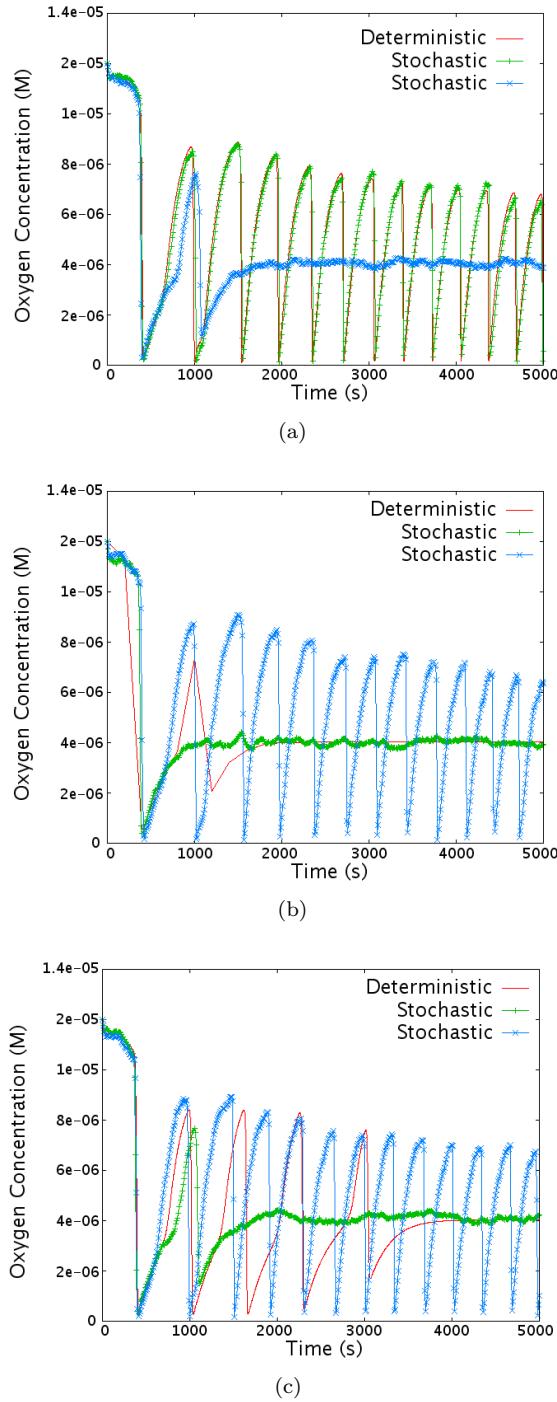


Figure 2.6. Oxygen concentration as a function of time for $[H_2O_2]_0 = 0.28 \mu\text{M}$ (a), $0.32 \mu\text{M}$ (b), or $0.30 \mu\text{M}$ (c). Red: Deterministic simulation results; green and blue: stochastic simulation with qualitatively different results (system volume is 10^{-15} L).

The response of the system to initial conditions is then quantified by running the simulation a sufficient number of times and calculating the probability of finding the system in either of the two attractors.

Among N independent simulations of the system, the number n of instances of oscillation obeys a binomial distribution of mean Np and variance $Np(1-p)$, where p is the probability of the system assuming oscillatory behavior. With a confidence of 95%, N should satisfy

$$\frac{2\sqrt{Np(1-p)}}{Np} < 0.05$$

to achieve an error $\leq 5\%$ in p . In particular, $N > 1600$ if $p \approx .5$. Therefore, more than 1600 simulations of the BFSO model for each initial $[\text{H}_2\text{O}_2]_0$ were used to calculate p and its error using

$$\begin{aligned} p &= \frac{n}{N}, \\ \Delta p &= \frac{2\sqrt{Np(1-p)}}{N}, \end{aligned}$$

respectively.

Figure 2.7 shows the convergence of p to a constant when the number of simulation is increased. In these simulations, $[\text{H}_2\text{O}_2]_0$ is chosen to be 0.30 μM , and 1600 simulation runs were carried out with different length of simulation time. The probability for the system to be in either oscillatory or steady state was estimated after the completion of each simulation and plotted against the number of finished simulation runs. Clearly, once a sufficient number of simulations are realized, we would be able to get a very accurate estimate of p . In addition, even if different simulation time is used, the estimated p is within statistical error of one another, suggesting that 5000 s is enough for the system to relax to an attractor.

The purpose of this work is to explore the effect of system size on the behavior of the system when $[\text{H}_2\text{O}_2]_0$ is changed in the critical region. Simulations were carried out using different $[\text{H}_2\text{O}_2]_0$, and the probability p of oscillatory state was calculated as above. From the results summarized in figure. 2.8, one readily sees the effect of stochastic nature of chemical reactions: the transition from

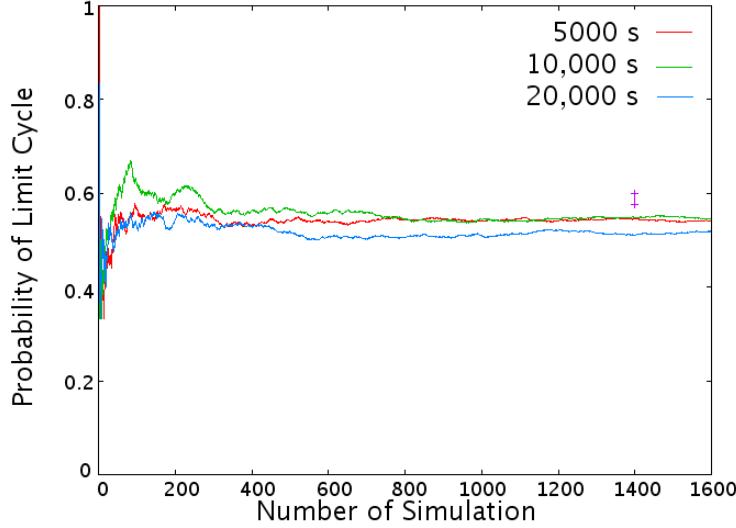


Figure 2.7. Convergence of p when the number of simulation runs is increased. System volume was 10^{-15} L, and simulation time was 5000 (red), 10,000 (green), or 20,000 (blue) s. After 1600 simulation runs, p is within statistical error (purple bar) of one another.

oscillatory behavior to steady state is a *continuous* function of $[H_2O_2]_0$. This transition approaches to step function as system volume is increased. However, even when $V = 10^{-13}$ L, the probability of oscillatory dynamics for $[H_2O_2]_0$ close to c_0 is still about 0.5, far away from possible deterministic values of 0 or 1.

It is intuitive to understand the results with a much simplified picture of a particle moving in a double-well potential. Here the valley of the double well corresponds to either oscillatory behavior or steady state. $[H_2O_2]_0$ determines the initial position of a ball moving under the influence of both the potential and a random force. The potential, originated from the reaction network, exerts a deterministic force on the ball's movement; the random force is due to the stochastic nature of chemical reactions, and its magnitude is inversely proportional to the square root of the system volume [39]. In our particular PO system, the barrier between the wells is so high compared with random force that no transition between wells is observed once the system settles in either well. But if the system starts very close to the top of the barrier, random force may be strong enough to perturb it to either side with almost equal probability. Even when the random noise is decreased by increasing system volume, the fate of the system is undetermined as long as it is close enough to

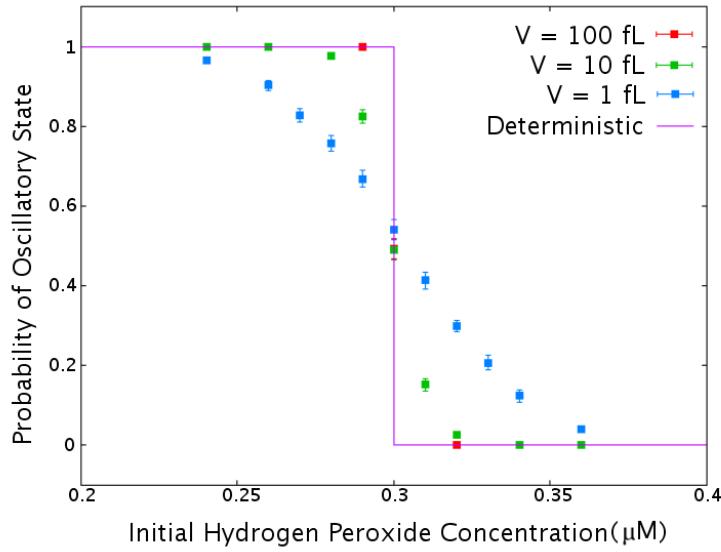


Figure 2.8. The probability of observing oscillatory behavior in PO system as a function of initial H₂O₂ concentration. System volume was 10⁻¹³ L (red), 10⁻¹⁴ L (green), or 10⁻¹⁵ L (blue). Deterministic result was also shown as a line.

the top of the barrier; in fact, p is independent of system volume ranging from 1 fL to 100 fL when $[H_2O_2]_0 = 0.30 \mu M$. It turns out that p is very close to 0.5 under this condition, and this suggests that the potential is quite symmetric at the barrier top. The symmetric feature of the potential is also reflected in the symmetric shape of the curves in figure 2.8 in smaller volumes around the point $(c_0, 0.5)$.

Conclusion and discussion

In summary, we carried out stochastic as well as deterministic simulations of PO system using realistic reaction models. The behavior of the system, controlled by the initial H₂O₂ concentration, exhibits a continuous transition when reaction noise is accounted for. Its deviation from the discontinuous transition predicted by deterministic simulation is more obvious when system volume is smaller, but very close to the critical H₂O₂ concentration, all stochastic results show significant deviation from deterministic prediction independent of system volume.

The experimental verification of the above conclusion is expected to be a direct corroboration of

master-equation approach to chemical reaction kinetics. From a practical point of view, this system is easier to handle for the control parameter is the initial concentration of a reagent. This saves us from maintaining the concentrations of certain chemicals during the course of the experiment as normally required by some other oscillators. The fate of the system after a short transient period may assume either one of the two qualitatively different choices, leaving little uncertainty in the interpretation of the data. With the advent of microfluidics, such experiments may be carried out with high throughput to gain enough statistics for comparison with numerical results.

The smoothing of a discontinuous transition is very similar to stochastic defocusing where the steepness of the system's response curve to the control parameter is reduced by stochastic noise in chemical reactions [72]. It would be interesting to determine whether this is detrimental or beneficial to living organisms and how nature avoids or employs this effect.

2.2 The Effect of Diffusion on Confined Enzymatic Reaction Cascades

2.2.1 Introduction

In vitro studies on biological systems have attracted more and more interests among biologists and chemists by offering a typical reductionists' way of tackling complex problems pertaining to life sciences [1]. Practically, chemists in this area are working in a biological setting by leveraging their expertise on chemical reactions. Just as the property of a molecule can, at least in principle, be deduced from the constituent atoms and bonds, the interesting behaviors of complex biological systems can also be explained by studying their basic elements, that is, chemical reactions and their communications.

Most of the chemical reactions in cells are catalyzed by different enzymes with a very high specificity and efficiency. Usually, enzymes are regarded as factories converting raw materials to end products. These factories are not working independent of one another; rather, they function as a whole entity to generate the multitude of life sciences. To wire these enzymes together, na-

ture has to tune the communication among them appropriately. This communication is mediated through the coupling of enzymatic reactions. Two reactions are coupled by sharing the same type of molecule in their reaction schemes. The shared molecule may be small metabolites or macromolecules, including enzymes themselves. These communications usually lead to a very large and correlated reaction network whose dynamics cannot be predicted simply by intuition. Numerical simulation using stochastic algorithm has been successful in solving the dynamics of complicated reaction networks [40, 108].

Compared with bulk chemical reactions, the complexity of biological system is more evident when spatial constraints are taken into consideration. In this sense, diffusion is influential in determining the dynamics of the reaction network. In section 2.1, biochemical reaction networks are studied under the assumption that all species are well mixed. In living cells, however, molecules interact with each other in a crowded and confined environment [99], where diffusion is closely coupled with reactive collisions. In biological systems, due to the coexistence of molecules with extremely different diffusion constants, the significance of diffusion may exhibit itself evidently. So far, people have studied the effect of diffusion on different biological systems and showed that diffusion is indispensable in the explanation [126] and prediction [35] of some complex behaviors such as synchronization and pattern formation.

The effect of diffusion may be modeled directly as discussed below in section 2.3, but it is instructive to explore its influence on the basis of some scaling analysis. Recently, a group of authors investigated the effect of spatial limitation or compartmentalization starting from timescale analysis of diffusion process and using stochastic simulations of a group of enzymes [80, 96, 136–139]. Their system (termed molecular network) is unique in that law of mass action is not valid, and nonlinear dynamics emerge. We will summarize this theory and then make some extension in different experimental contexts in this section. The following section is devoted to an outline of molecular network. In section 2.2.3, the effect of free diffusion of small molecules on one or several molecular networks are described. In section 2.2.4, a possible method is proposed to enslave the dynamics of a group of enzymes by a strong external injection of signal molecules. The feasibility of this scheme depends largely on the legality of the model in section 2.2.

2.2.2 Molecular Network

Consider a cellular compartment of size L in which an enzyme is converting substrates into products.

The following timescale analysis [58] will reveal the existence of two different dynamic regimes, namely, local or global. First of all, because the size of enzyme is much larger than that of small molecules (substrates or products), one can neglect the diffusion of enzymes. In mathematical language, the diffusion constant of enzyme is much smaller than that of small molecules, so only the latter contributes significantly to the relative diffusion constant D . Then one has a good estimate of the time needed for a small molecule to diffuse throughout the compartment (t_{mix}),

$$t_{\text{mix}} \approx L^2/D. \quad (2.1)$$

In another word, it takes about t_{mix} for a molecule to forget its initial location.

The substrate molecule has to collide with the active site of an enzyme to start a reaction, so another important characteristic time is the traffic time, t_{traffic} , the small molecule spends on diffusion before meeting the active site of an enzyme (figure 2.9). According to reaction-diffusion theory [57], in three-dimensional space, if there is only one enzyme in the system,

$$t_{\text{traffic}} \approx L^3/DR = t_{\text{mix}} \cdot \frac{L}{R}, \quad (2.2)$$

where R is the size of the active site. The last expression in equation (2.2) means it always takes longer for a small molecule to collide with the active site of an enzyme than to transverse the cell volume. Equivalently, the small molecule has forgot its initial position before it finds an enzyme. For a real biological system, there are usually many enzymes (say, N). Then it takes a shorter period of time (t_{transit}) for the molecule to dock an active site. If the enzymes are randomly distributed in the volume,

$$t_{\text{transit}} = \frac{1}{N} t_{\text{traffic}} = \frac{L^3}{NDR} = t_{\text{mix}} \cdot \frac{L}{NR}. \quad (2.3)$$

Now the relation between t_{transit} and t_{mix} can be changed qualitatively by tuning N around a

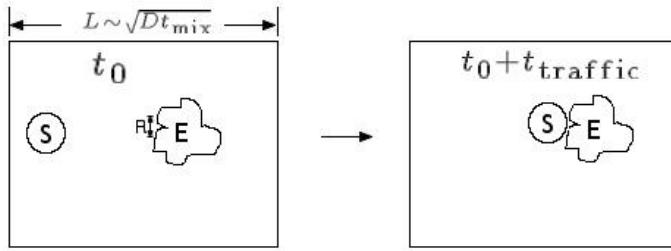


Figure 2.9. The physical meaning of t_{traffic} . In a small volume of dimension $L \approx \sqrt{Dt_{\text{mix}}}$ containing only one enzyme, the small substrate molecule spends t_{traffic} in diffusion before colliding with the active site (of dimension R) of enzyme E.

critical number

$$N_{\text{crit}} = L/R.$$

If $N \ll N_{\text{crit}}$, $t_{\text{transit}} \gg t_{\text{mix}}$ which means the substrate molecule has lost its memory of initial condition before docking an enzyme's active site. In another word, it is equally possible for the substrate to react with any enzyme in the system. So the enzyme can feel the *global* change of the number of substrate molecules. In the regime where $N \gg N_{\text{crit}}$, that $t_{\text{transit}} \ll t_{\text{mix}}$ indicates the substrate molecule finds a docking site before traveling a significant portion of the volume. Alternatively, the substrate molecule only interacts *locally* with nearby enzymes.

To have a feeling of the order of magnitude, consider a volume of micrometer size containing enzymes with nanometer-sized active sites, then $N_{\text{crit}} = 1000$ and the critical concentration $c_{\text{crit}} \approx 10^{-6} \text{ M}$. With this number of enzymes, if the diffusion constant for small molecules is about $10^{-5} \text{ cm}^2/\text{s}$, then $t_{\text{mix}} = t_{\text{transit}} \approx 1 \text{ ms}$ and $t_{\text{traffic}} = 1 \text{ s}$.

Finally, the characteristic time of an enzyme reaction, $t_{\text{turnover}} \approx 1/k_{\text{cat}}$, should be taken into consideration. If t_{turnover} is smaller than both t_{transit} and t_{mix} , the system is equivalent to a diffusion-controlled reactor in the local dynamics regime or a well-stirred reactor in the global dynamics regime. These two scenarios are familiar to chemists. On the other hand, if enzyme dynamics is the slowest process in the system, in the global regime, during the enzyme turnover cycle, the change in the number of the substrates can be experienced by every enzyme with equal probability, because t_{mix} is the smallest timescale in the system. For local dynamics, one can divide the volume into several

smaller ones, each of which gives a smaller t_{mix} with almost the same t_{transit} until global dynamics is possible in these element cells. This is feasible because t_{mix} is dependent on the size of the system (equation (2.1)), but t_{transit} is controlled by enzyme concentration (equation (2.3)). In summary, the only interesting regime for new dynamics is

$$t_{\text{turnover}} \gg t_{\text{transit}} \gg t_{\text{mix}}, \quad (2.4)$$

that is, the enzyme turnover cycle is the slowest process in the system, and the diffusion of substrate is the fastest. This criterion may be fulfilled by the compartmentalization of enzymes in a volume of appropriate size, which is common in biological systems.

When equation (2.4) is satisfied, any substrate molecule has traveled throughout the system before finding an active site, and enzyme reaction is very slow compared with these two processes. From an enzyme's point of view, at every moment of its turnover cycle, it can collide with any substrate with the same possibility, and thus the change of the number of substrates affects the dynamics of each enzyme with the same strength. In such a system, the dynamics of a group of enzymes may be synchronized through the diffusion of small molecules, and it is called a molecular network [58]. To enable the synchronization, there should be a mechanism in the system for small molecules to exert their effect on the dynamics of enzymes. Various feedback mechanisms have been proposed, including product inhibition [137], product activation [96, 136, 139], allosteric activation [80], and substrate recycling [138].

It is the interest of this report to concentrate on the substrate recycling scheme in which the product is converted back to substrate by an additional fast pathway in the background (figure 2.10). It is easy to find such a reaction system in the repertoire of biological pathways. A case in point is the cycling between adenosine triphosphate (ATP) and pyrophosphate (PPi) [54] (figure 2.11). In the language of figure 2.10, substrate ATP is converted to product PPi by luciferase, and product is transformed back to substrate by another enzyme (ATP sulfurylase) which has a faster dynamics than luciferase [54]. A more well-known example, namely, enzyme cascade, is also related to this scheme. In this case the substrate is not a small molecule but a protein which can be covalently

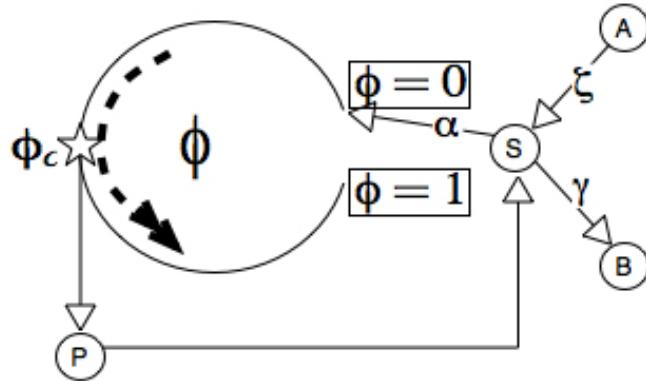


Figure 2.10. A schematic description of an enzyme in a molecular network. Substrate S is supplied and decays at a constant rate ζ and γ , respectively. Enzymatic cycle is characterized by a phase variable ϕ . When $\phi = 0$, the enzyme is available to bind with a substrate at a probability rate of α . Product P is released once $\phi > \phi_c$ and converted back to substrate immediately.

modified, say, by a phosphorylating kinase. The modified protein is then subjected to dephosphorylation spontaneously or catalytically (by a phosphatase). This cycle of phosphorylation level is equivalent to the hydrolysis of ATP. The energy stored in ATP is seemingly wasted, so it is also called a futile cycle. In a futile cycle, although the substrate is not a small molecule, its diffusion constant may assume an appropriate value to make a molecular network possible. There has been a large body of literature on the significance of this futile cycle since 1970s. It is clear now the futile cycle renders ultrasensitivity [17, 28, 29, 123, 135], zero-order sensitivity [46], memory storage [85, 86], noise reduction [158], and stochastic focusing [105]. To the best of my knowledge, all these works, either simulations or experiments, are concerned with steady-state behavior or fluctuation around the steady state. The uniqueness of molecular network is its departure from steady-state simulation to real dynamic exploration. The surprising results show the synchronization of the dynamics of the enzymes in a molecular network [138].

Before showing the numerical result reproduced by me, it is useful to briefly describe how to simulate a system described in figure 2.10 (for more detail, see reference [137]). One first approximates the turnover cycle of an enzyme as a one-dimensional diffusion process on a biased potential surface. Numerical simulation has shown the possibility of this reduction of dimensionality in the context

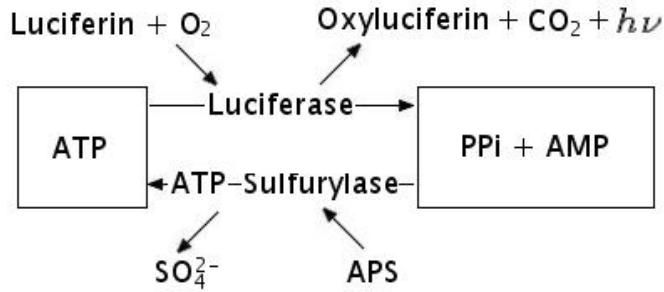


Figure 2.11. A possible realization of the enzymatic cycle in figure 2.10. Substrate ATP is converted to product PPi by luciferase with the consumption of luciferin. PPi is readily used by ATP sulfurylase to regenerate ATP using APS (adenosine phosphosulfate). The activity of luciferase can be monitored by luminescence.

of protein folding [10]. A phase parameter $\phi \in [0, 1]$ is used to characterize the state of enzyme or enzyme–substrate complex in its configuration space [79]. When $\phi = 0$, the enzyme is ready to bind with a substrate with probability rate α . In a molecular network, the possibility of the binding of a substrate molecule with an enzyme in time interval dt is

$$1 - (1 - \alpha dt)^s \approx \alpha s dt,$$

where s is the number of free substrate molecules in the system. After the formation of enzyme–substrate complex, phase parameter ϕ is evolved according to the following Langevin equation,

$$\frac{d\phi}{dt} = v + \eta(t), \quad (2.5)$$

where v is the drift velocity along the potential surface of enzyme–substrate complex, and $\eta(t)$ is a Gaussian noise introduced to simulate the effect of thermal fluctuation. Drift velocity is simply assumed to be the reciprocal of t_{turnover} , which is equivalent to approximating the conformational evolution of enzyme–substrate complex as a “clock.” This clock, however, is not accurate due to the

existence of Gaussian noise $\eta(t)$ with a delta correlation function

$$\langle \eta(t)\eta(t') \rangle = 2\sigma\delta(t - t'),$$

where correlation intensity σ can be related to the dispersion ξ of turnover time as follows [79, 138]

$$\xi \approx (2\sigma\tau_{\text{enzyme}})^{1/2}. \quad (2.6)$$

As ϕ is increased from 0 according to equation (2.5), it will pass a critical point ϕ_c when a product is released and the enzyme begins the relaxation toward the initial configuration which is also the end of an enzymatic turnover cycle. This clock approximation, despite of its roughness, grasps the essence of a typical enzymatic cycle. Indeed, the binding of substrate molecule changes the potential surface of a single enzyme so that the initial configuration is not an energy minimum for the enzyme–substrate complex, and this complex relaxes along the rugged downslope of potential surface [10]. This is consistent with the fact that it is the transition state rather than substrate that is stabilized by an enzyme. After the release of product, the potential energy surface recovers and the enzyme relaxes backward toward initial configuration. The rate of relaxation may be different from that of the previous downslope motion because the substrate is absent. It should be noted that the detail of the reaction is not covered in the turnover cycle simulated by equation (2.5).

Under this simulation scheme, the dynamics of a group of enzymes in a molecular network shows the synchronization of the turnover cycle of individual enzymes. In particular, the enzymes self-organize into group(s) which behaves like a single enzyme. Figure 2.12 depicts the time evolution of the number of substrate molecules in a molecular network. The oscillation is the result of synchronized activity of individual enzymes. Without synchronization, the number of substrates would be distributed randomly around the steady-state value. With synchronization, the enzymes in the system release products, which are converted back to substrates at once, almost simultaneously and thus peaks of substrate generation are observed. The number of free enzymes also oscillates at the same frequency as expected.

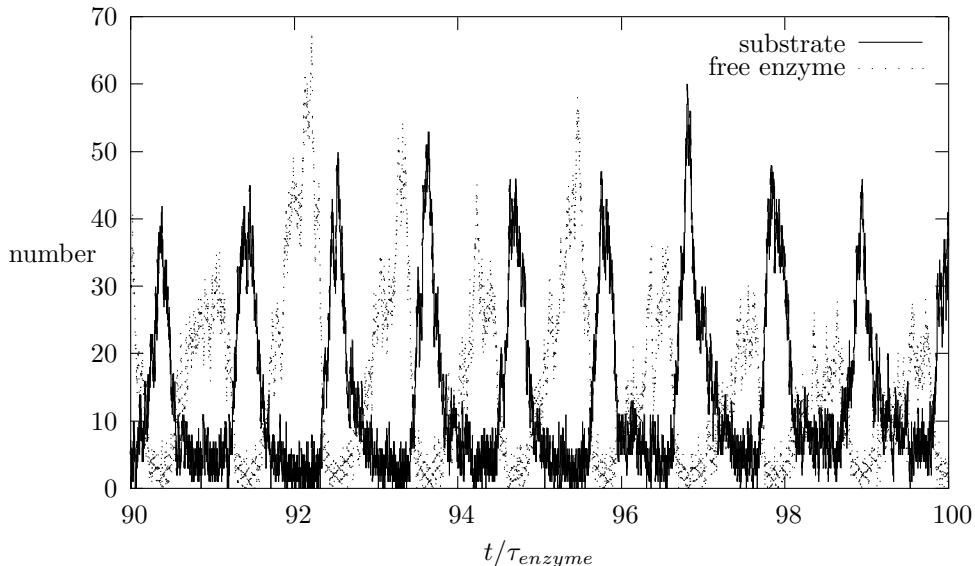


Figure 2.12. The number of substrates (line) and free enzymes (dots) as functions of time in a molecular network consisting $N = 1000$ enzymes. Other parameters were $\phi_c = 0.2$, $\xi = 0.02$, $\gamma = 15/\tau_{\text{enzyme}}$, $\zeta = 200/\tau_{\text{enzyme}}$, and $\alpha = 10/\tau_{\text{enzyme}}$.

The generation of synchronization might be related to the clock nature of every single enzyme. In a highly abstract mathematical theory of Kuramoto, clocks can be synchronized by strong enough coupling to give coherent dynamics [143]. This idea has been verified in a real chemical system [73, 98]. It is interesting to ask if similar behavior is possible in biochemical systems. It seems difficult to find a biological system that can be described by a clock as simple as that in Kuramoto theory. Also, the coupling in real world is more complicated. But the molecular network mentioned in this section is a reasonable approximation to real biological dynamics and possible connection between real biological systems and Kuramoto theory. It would be instructive to understand molecular network within the framework of a mathematical theory on coupled oscillators or clocks.

Oscillation is important to the control of metabolism [25, 75], so people have been interested in its origin. In general, oscillation is typical of nonlinear chemical reactions whose dynamics have inherent periodic solutions. The resulting periodicity is a feature of the nonlinear differential equation based on law of mass action. Researchers have proposed different mechanism to elucidate the oscillations observed in biological systems along this line [45]. In contrast, in the cycling scheme in figure 2.10, traditional kinetic equations do not have nonlinearity [138]. When space constraints are considered,

periodicity emerges in the dynamics. It is interesting to study if this mechanism can be exemplified in a real biologically relevant system *in vitro*. If that is the case, one may further look for its existence *in vivo*.

2.2.3 The Coupling among Molecular Networks

As mentioned above, the emergence of oscillation is related to the clocklike assumption on the dynamics of the enzymes in the molecular network. In the original model [138], the communication among enzymes is analyzed *a priori* in terms of the separation of timescales, so the detail of this communication is not embodied in the model. Other researchers have made simulations of similar systems using more realistic models, for example, network [147] or random walk [168], to describe communication among enzymes and revealed their significance on the synchronization of enzymes. In contrast to these studies, the dynamics of several molecular networks coupled together by the diffusion of small molecules seems interesting if we think about the problem at a higher level. Here, each molecular network rather than each enzyme behaves like a clock and influences other networks through different communication methods. This idea is not new in literature. For example, some researchers were interested in the synchronization of many neuron networks rather than that of neurons in a single network [97]. The interaction among several molecular networks was only reported very recently [14].

In terms of experimental research, it is also useful to study the behavior of several molecular networks for it is easier to observe the dynamics of each network (\sim micrometers) than that of a single enzyme (\sim nanometers). In our lab, protein patterns have been made with a high precision of spatial definition [67]. These protein patches might act as molecular networks and interact among one another by the diffusion of small molecules. Furthermore, in real biological systems, like cell membrane, domains with a high concentration of a certain enzyme are universal, and they might talk with each other rather than work independently [51, 92]. All these can be cast into a model composed of several molecular networks connected by the diffusion of small molecules. Before tackling this problem, it is also useful to start from a much simpler scheme.

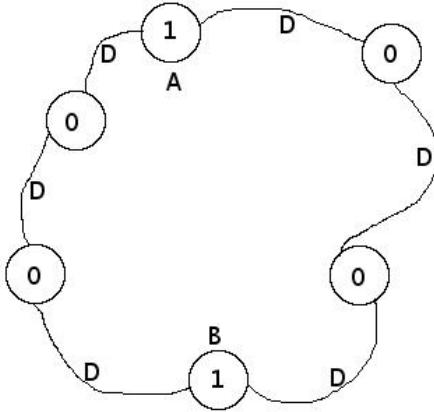


Figure 2.13. Connected molecular networks. The transportation of small molecules among vesicles is approximated by a stochastic hopping process with probability rate D . Each sphere labeled “1” denotes a GUV in which the criterion for molecular network is satisfied and enzymatic cycles in figure 2.10 are also found. Those labeled “0” lack the key enzymes in figure 2.10, but still have the source and the sink pathways of substrates.

2.2.3.1 Hopping Mechanism

The simplest method to model the diffusion among molecular networks is random walk. Actually there have been efforts to set up a network of giant unilamellar vesicles (GUVs) connected by lipid tubules [69, 71]. Normally, these microtubules are barely visible under light microscopy because of their small diameters, but mass transportation through these lipid “bridges” has been confirmed by fluorescence microscopy. By tuning the diameter and length of the bridge, the effective transportation coefficient among connected GUVs may be adjusted. If the reaction in each GUV can be treated as a molecular network, it is possible to study the effect of connectivity on the dynamics of coupled molecular networks by running biochemical reactions in vesicle networks. Thus, it is useful to get some insight of the dynamics of coupled molecular networks by numerical simulation.

Figure 2.13 shows the cyclic connection of six molecular networks which may be vesicles in real experiments. These vesicles are of an appropriate size to satisfy the criterion of molecular networks (equation (2.4)), but they may or may not be loaded with the key enzymes in figure 2.10 and

thus labeled “1” or “0”, respectively. A parameter D is used to describe the probability rate for small molecules to jump to adjacent vesicles. By increasing this parameter, the correlation between adjacent molecular networks is enhanced. As a proof of concept, the numbers of substrates in the two type 1 vesicles in figure 2.13 are simulated and compared in figure 2.14 for two different values of D . The cyclic arrangement in figure 2.13 is equivalent to an infinite array of molecular networks separated by two “empty” compartments lacking enzymes. As D increases, the phase difference between the substrate oscillation in the two molecular networks changes from 180° to 0° . This is consistent with our intuition because D is a measure of coupling strength between molecular networks. In a different context, increasing coupling also leads to the synchronization of electrochemical oscillators [73]. Note that usually it is the cumulative behavior that is the output of a biological unit, so the ability to tune the correlation among constituent molecular networks enables a cell to generate qualitatively different output to other pathways. In particular, the oscillation of a single molecular network is buried in the overall output if there is no correlation among constituent networks, whereas oscillatory output is achieved if every molecular network rocks almost in phase. In a case as simple as figure 2.13, a phase difference of 180° makes the overall output only half of that for two synchronized networks. The ability of generating oscillation and tuning its amplitude is significant in the control of the metabolism and signal transmission, because oscillatory signal can trigger cell response even if average signal intensity is below the threshold [25]. It is also straightforward to test this conclusion in a real GUV network.

2.2.3.2 Effect of Free Diffusion and Distance

Let us returned to address the question put forward at the beginning of this section: the dynamics of several molecular networks coupled by the diffusion of small molecules. Besides the effect of diffusion on a single molecular network, the significance of internetwork distance is also considered in this subsection.

To simulate this system numerically, one has to combine the simulation of diffusion equation and molecular networks. We start this endeavor from the simplest case, that is, one-dimensional

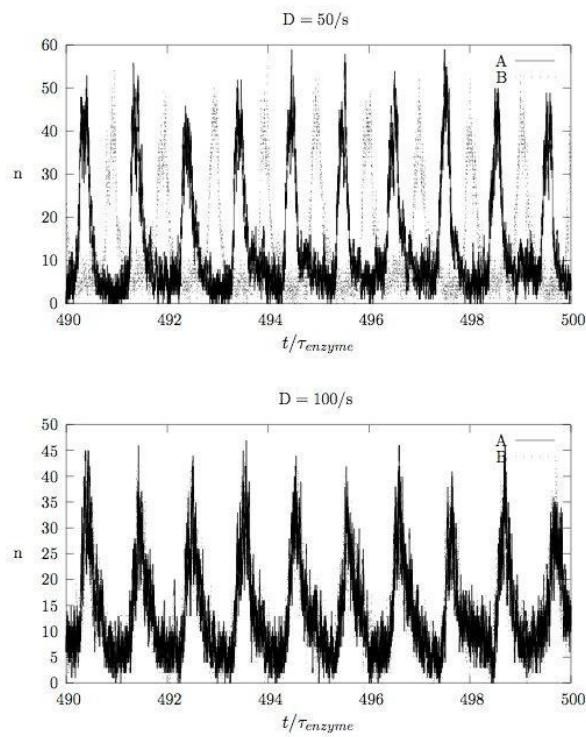


Figure 2.14. The number of substrates (n) in the two molecular networks (labeled “1” in figure 2.13) as a functions of time. Upper panel: Hopping probability rate (D) is $50/\tau_{\text{enzyme}}$ and the dynamics of two networks (A and B in figure 2.13) are not synchronized. Lower panel: D is increased to $100/\tau_{\text{enzyme}}$ and the dynamics of A and B are synchronized.

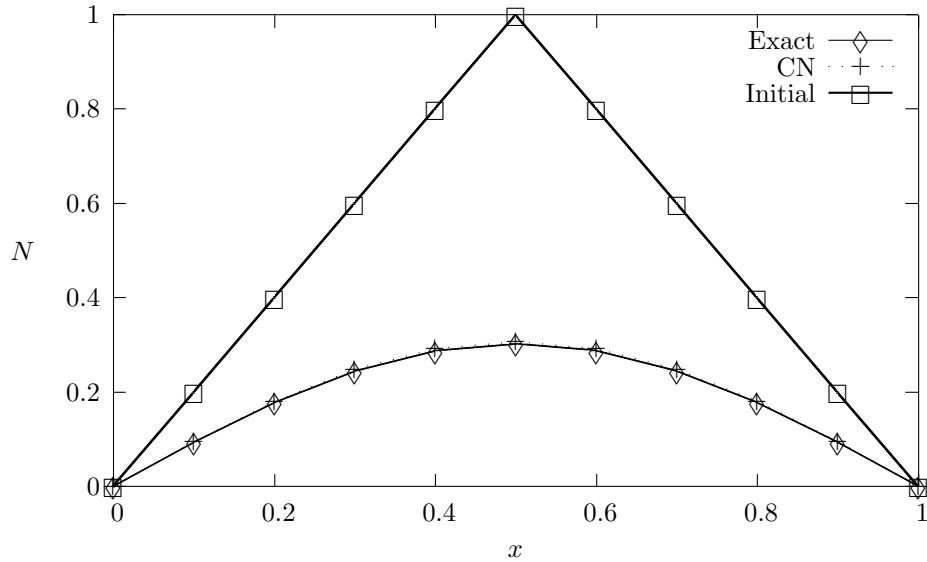


Figure 2.15. The solutions to a test diffusion equation $\frac{\partial N(x,t)}{\partial t} = D \frac{\partial^2 N(x,t)}{\partial^2 x^2}$ with Dirichlet boundary condition and a given initial condition (square) by Crank–Nicholson (cross) and analytical (diamond) method.

diffusion. First of all, one-dimensional space was divided to many cells of dimension Δx and the substrate concentration inside each cell was assumed to be uniform. Diffusion equation

$$\frac{\partial N(x,t)}{\partial t} = D \frac{\partial^2 N(x,t)}{\partial^2 x^2} \quad (2.7)$$

was solved numerically to give the number of substrates (N) at time t and in the cell with coordinate x . Among different discretization schemes to approximate equation (2.7), Crank–Nicholson method was chosen [107]. It gave a fairly good result in a test case (figure 2.15). In mathematical language, the discretization of diffusion equation couples different oscillators by adding to the governing differential equation of each oscillator an additional term dependent linearly on the state of other oscillators. In this sense, although the individual oscillator is not modeled here by a differential equation explicitly, the coupling of molecular networks can still be compared to the coupling of neurons. The difference lies in the fact that the coupling term for a neural network is highly abstract and not necessarily an approximation to Fick diffusion. Still, it would be interesting to compare the coupling of neurons to that of molecular networks in future works.

0	1	0	0	1	0	0	1	0
---	---	---	---	---	---	---	---	---

Figure 2.16. The discretization of one-dimensional space into several cells. The meanings of “1” and “0” are clarified in the text.

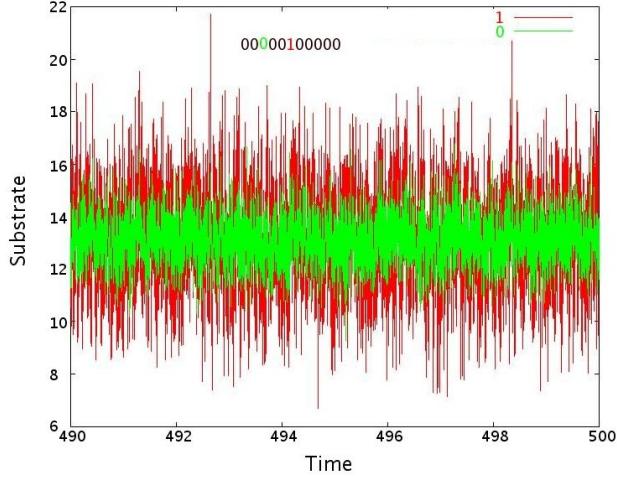


Figure 2.17. The numbers of substrates in compartments with (red) and without (green) enzymes as functions of time (in unit of τ_{enzyme}). One-dimensional space was divided into 11 compartments of either types. Further increase of the number of surrounding type 0 cells did not change the result qualitatively.

In my simulation, the elementary cells of size Δx fell into two kinds labeled by “0” or “1” (figure 2.16) as in section 2.2.3.1. From the property of molecular networks (equation (2.4)) and equation (2.1), it is reasonable to choose $\Delta x \approx \sqrt{D\Delta t}$. The parameters defining a molecular network were the same as those to obtain figure 2.12. In solving the diffusion equation, Dirichlet boundary condition was always assumed. Boundary condition had no effect on the dynamics of type 1 cells as long as the number of type 0 cells is big enough at the two ends.

As the first example, the effect of free diffusion of small molecules into surrounding medium is investigated. In figure 2.17, the numbers of substrates in the molecular network and an enzyme-free compartment are depicted. Due to the diffusion of small molecules into the surrounding medium,

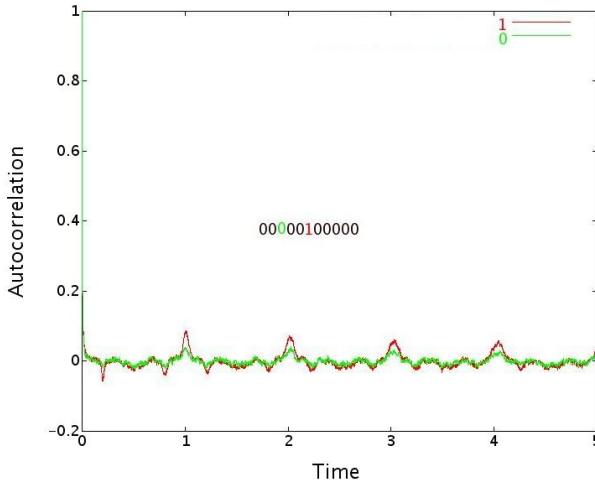


Figure 2.18. The correlation functions of the numbers of substrates in compartment with (red) and without (green) enzymes calculated from figure 2.17. The unit for time is τ_{enzyme} .

one can hardly find any oscillation in the number of substrates.

Another method to show the oscillatory behavior of a function $y(t)$ is to calculate the time correlation function $C(\tau)$,

$$C(\tau) = \frac{\langle (y(t) - \langle y(t) \rangle)(y(t + \tau) - \langle y(t) \rangle) \rangle}{\langle (y(t) - \langle y(t) \rangle)^2 \rangle},$$

where angular brackets denote time average. For a sinusoidal function, its correlation function is still sinusoidal with the same period. Generally the periodicity and amplitude of a correlation function is an indication of the oscillatory component of the original function. For figure 2.17, the corresponding correlation function is shown in figure 2.18 which exhibits the remnant oscillation in the molecular networks. In addition, it is not a surprise that correlation is smaller in the surrounding cells than that in the molecular network.

Furthermore, it is interesting to study the dynamics when more molecular networks are included in the system. In figure 2.19 the correlation functions of substrate numbers in molecular networks are depicted to show the effect of distance between two molecular networks. When they are separated by an appropriate distance, they can reinforce each other to counter the effect of diffusion, as we

can observe that the amplitude of the correlation function reaches maximum when internetwork distance is $4\Delta x$, or, in another work, there are three type 0 cells separating two molecular networks. Here, we are interested in the effect of coupling on the amplitude of autocorrelation functions which indicates the extent to which enzymes in a single molecular network are synchronized (figure 2.19). Additionally, similar to section 2.2.3.1, it is meaningful to study the cross-correlation between two

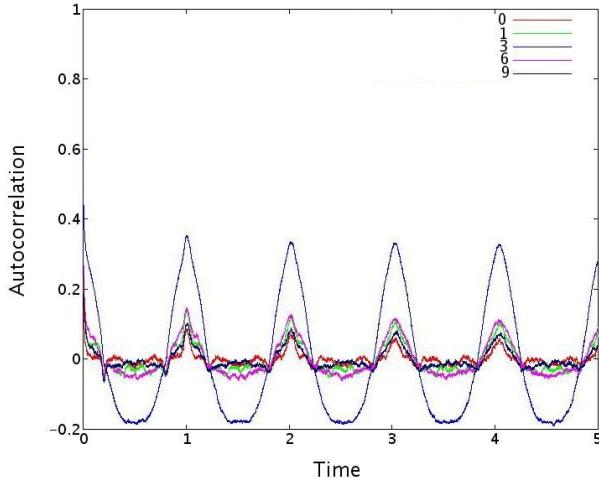


Figure 2.19. The correlation functions of substrate number as functions of distance between two molecular networks. The configuration of the system was similar to figure 2.18 except that there were two molecular networks separated by 1 (green), 3 (blue), 6 (purple), or 9 (black) type 0 cells. Red line is the same as figure 2.18 for comparison. Note that it is the autocorrelation function in one of the two molecular networks in each system that is depicted here. The amplitude of correlation function reaches a maximum when internetwork distance assumes an optimal value indicating the substrate oscillation in the coupled molecular networks can reinforce each other by the diffusion of substrates. The unit for time is τ_{enzyme} .

molecular networks separated by different numbers of type 0 cells. In general, for two time-dependent functions $y_1(t)$ and $y_2(t)$, we may calculate their difference $y(t) = y_1(t) - y_2(t)$ and compare the variance of $y(t)$ ($\sigma_y = \langle (y(t) - \langle y(t) \rangle)^2 \rangle$) with the sum of variances of $y_1(t)$ and $y_2(t)$. Here, the ensemble average used to calculate variance is replaced by average over time. The ratio

$$R = \frac{\sigma_y}{\sigma_{y_1} + \sigma_{y_2}} \quad (2.8)$$

should be one for two independent processes y_1 and y_2 , and it differs from one if y_1 and y_2 are correlated such that

$$\langle(y_1 - \langle y_1 \rangle)(y_2 - \langle y_2 \rangle)\rangle \neq \langle y_1 - \langle y_1 \rangle \rangle \langle y_2 - \langle y_2 \rangle \rangle = 0, \quad (2.9)$$

where brackets denote average over time. Along this line, we plot in the upper panel of figure 2.20 the variances (σ_1 and σ_2) of the number of substrates in the two molecular networks as functions of the number of type 0 cells in between. These variances reach maximum when two molecular networks are separated by three type 0 cells, because oscillation is more pronounced in this case. The variance σ of the difference between two networks is also plotted. Note that σ is not a maximum when distance is $4\Delta x$. According to equation (2.8), R is calculated and depicted in the lower panel of figure 2.20. R is significantly different from one when distance is 3, suggesting significant correlation between the two molecular networks. On the other hand, when R is not far from one and thus equation (2.9) almost becomes an equality, although it is mathematically insufficient to rule out the correlation between y_1 and y_2 , it is still safe to say that there is little correlation between the two molecular networks. In fact, the cross-correlation¹ between the two molecular networks in question is almost zero for large spatial separation where $R \approx 1$ (figure 2.21). Finally, note that the more molecular networks in a system, the bigger the amplitude of correlation functions (figure 2.22).

On the basis of the numerical simulation, one can discuss its biological significance. As mentioned in section 2.2.3.1, the ability of generating and tuning oscillatory signal output is important for biological entities. Using a simple one-dimensional model, we show a possible method to achieve this, namely, simply changing the spatial arrangement of several molecular networks. A similar result was also reported in a different context [126]. This scheme is plausible in that nature does not have to evolve new enzymes to generate qualitatively different signal output, and smart utilization of existing function is usually the trick of nature.

¹The cross-correlation between $y_1(t)$ and $y_2(t)$ is defined as

$$C(\tau) = \frac{\langle(y_1(t) - \langle y_1(t) \rangle)(y_2(t + \tau) - \langle y_2(t) \rangle)\rangle}{\sqrt{\langle(y_1(t) - \langle y_1(t) \rangle)^2\rangle}\sqrt{\langle(y_2(t) - \langle y_2(t) \rangle)^2\rangle}}.$$

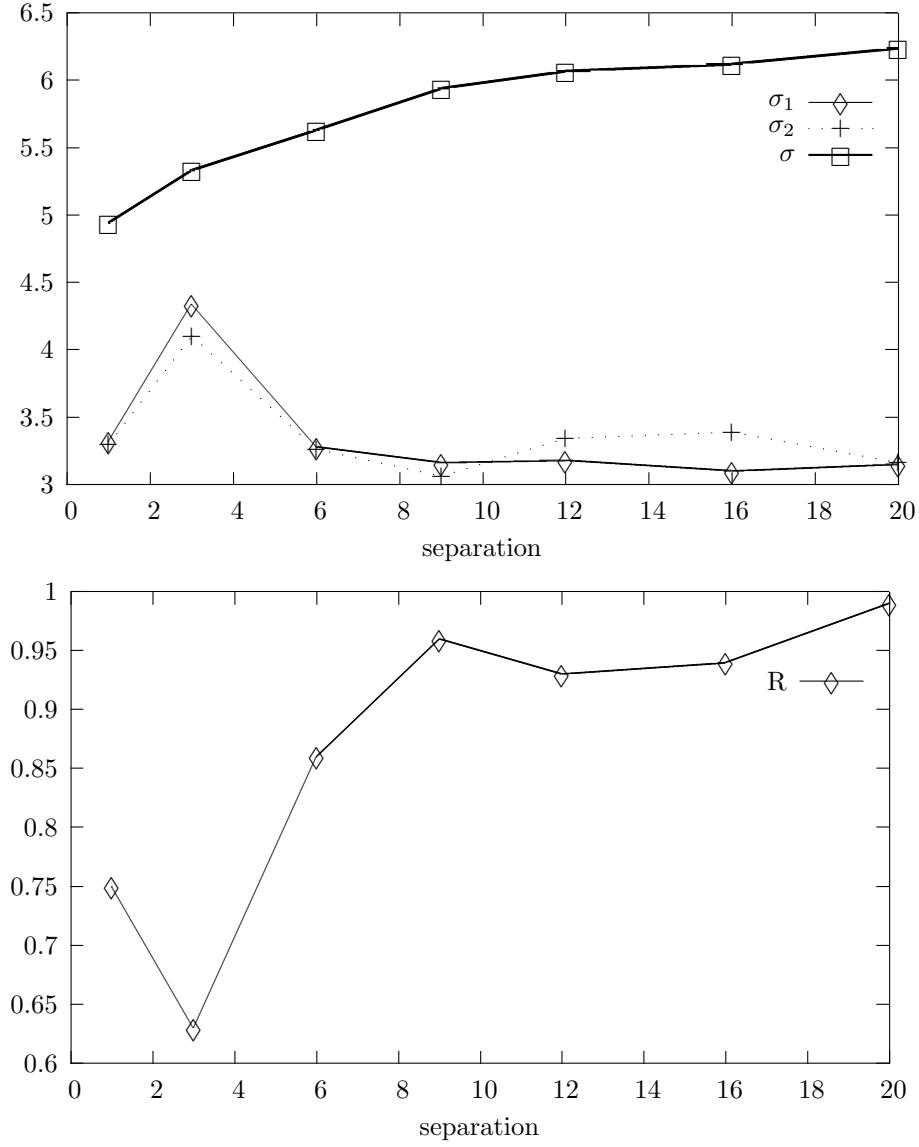


Figure 2.20. Upper panel: the variances of the number of substrates in two molecular networks (diamond, σ_1 ; cross, σ_2) as functions of their separation and the variance of the difference between the numbers of substrates in two molecular networks (square, σ). Lower panel: the ratio R between σ and $\sigma_1 + \sigma_2$.

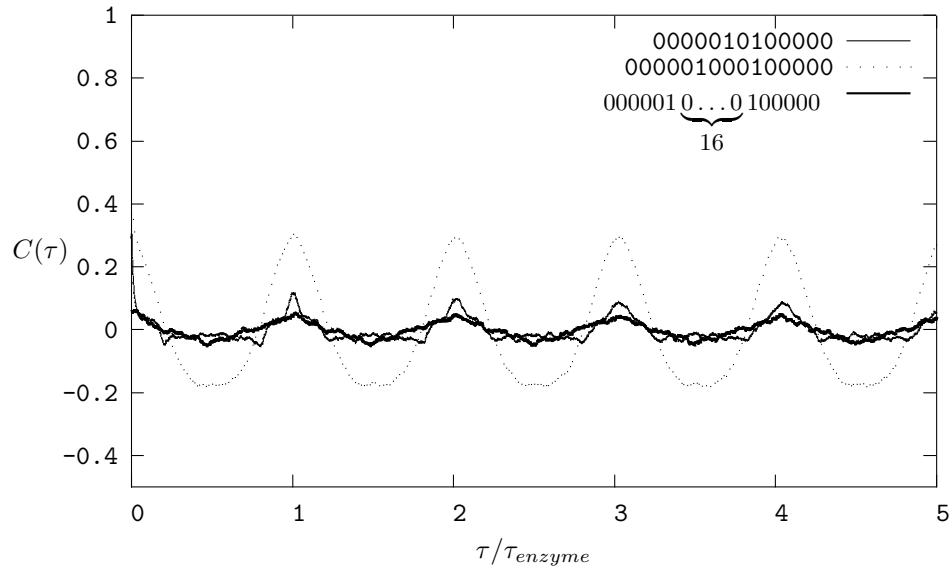


Figure 2.21. The cross-correlation functions (as defined in note 1) between two type 1 molecular networks separated by 1 (line), 3 (dots), or 16 (bold) type 0 cells. When the distance between two networks is too large, there is little correlation between their dynamics.

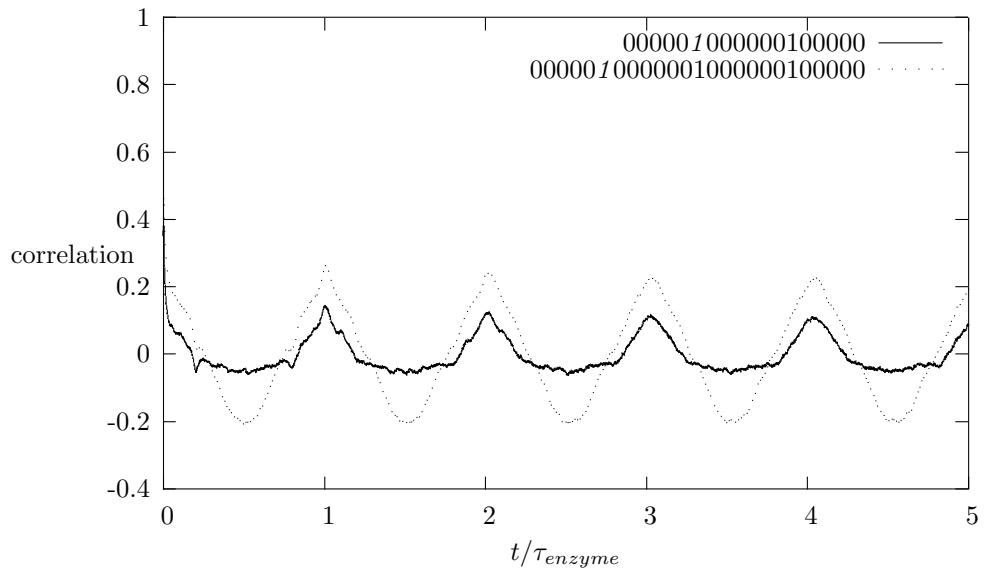


Figure 2.22. The correlation function is dependent on the number (solid, two; dot, three) of molecular networks in the system. The configurations of the systems are also labeled, and the autocorrelation functions of the italicized molecular network are shown.

In summary, this section discusses the coupling of several molecular networks which can also be studied in experiments. Further work will be concentrated on its verification by experiments and generalization to more complex configurations in simulation [14]. These works, however, are constructed on the basis of molecular network. Therefore, it is still important to study other aspects of this model and search for new implications which are significant to biological systems and verifiable experimentally.

2.2.4 Effect of Source Oscillation

Molecular network also makes it possible to control the dynamics of a group of enzymes by external manipulation. Experimentally, many enzymes can be enslaved by pulsed light to act synchronously [49, 50]. Similar to this, in the regime of molecular network, external source of substrate may also serve to control the dynamics of enzymes.

To simulate this effect, the same molecular network as in figure 2.10 is used as a model. In addition to the substrate supply at constant rate ζ , an oscillatory source with rate

$$\zeta_1 = f \left(1 + \cos\left(\frac{2\pi t}{T}\right) \right) \quad (2.10)$$

is used to enslave the enzymes in the network, where f is the adjustable variable equivalent to the average rate of this oscillatory substrate influx. In real simulation, the number of substrates supplied by this source in an interval dt is generated by a Poisson distribution with mean $\zeta_1 dt$. The period T in equation (2.10) is chosen to be close to the average period without ζ_1 .

The results may be analyzed in terms of the number of free enzymes as a function of time. Similar to figure 2.12, the number of free enzymes is still periodic with the existence of oscillatory substrate source. Further analysis reveals two periods in the timecourse of the number of free enzymes (see figure 2.23). This indicates that there are two groups of enzymes in the molecular network. The enzymes in each group behave almost in phase such that they return to the free state (i.e., ready to bind substrate) almost at the same moment. The appearance of a second enzyme group is attributed

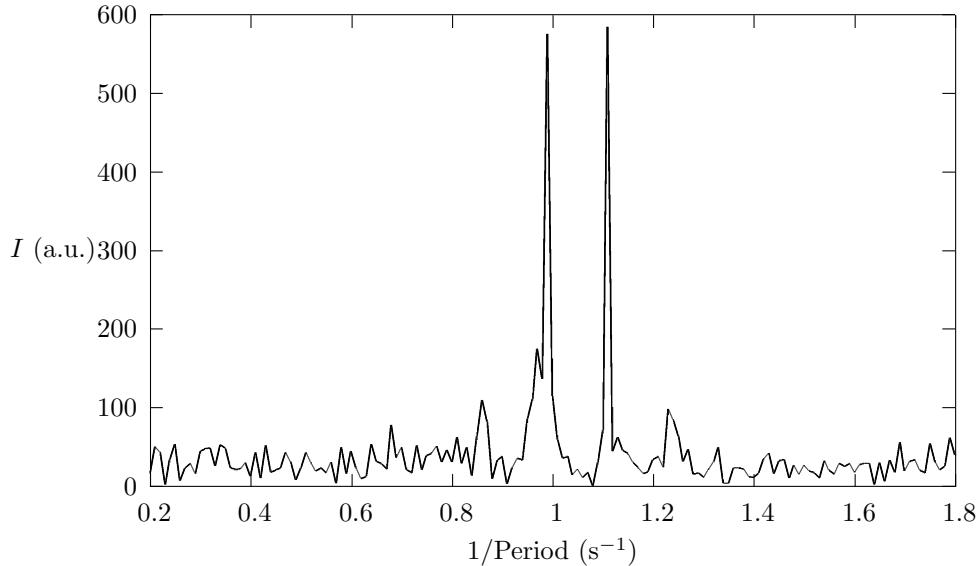


Figure 2.23. The Fourier analysis of the number of free enzymes in a molecular network. Two peaks show up as opposed to one without oscillatory source. In the simulation, $f = 512 \text{ s}^{-1}$, $T = 0.9 \text{ s}$, and other parameters were the same as figure 2.12 and $\tau_{\text{enzyme}} = 1 \text{ s}$.

to the oscillatory source without which only one group of enzyme is formed for the system simulated.

The two peaks in figure 2.23 have different origins. One is from the intrinsic oscillation which is independent of the source oscillation, and the other is the result of an external oscillatory substrate source. This is obvious if we pay attention to the positions of these peaks. In figure 2.24, the periods corresponding to intrinsic and external peaks are plotted as functions of f with external period T being equal to 0.9 s. The external peak has the same period as the oscillatory source as expected, but the period of the intrinsic oscillation approaches asymptotically the period under constant source as f is decreased. Furthermore, we plot the ratio of the heights of the two peaks as a function of f in figure 2.25. With the increase of f , the intrinsic peak is weaker compared with the external one. Note that the height of two peaks is almost equal when $f = 512 \text{ s}^{-1}$, although the strength ζ of the constant source is only 200 s^{-1} . This is an indication of the robustness of the intrinsic oscillation. Figure 2.25 also gives a threshold of $4 < f_c < 8$ for the external peak to be observable.

As noticed by others, some biological systems have responses to the changes in thermal noise [105].

In this context, we will study the change of external and intrinsic peaks when the intensity of

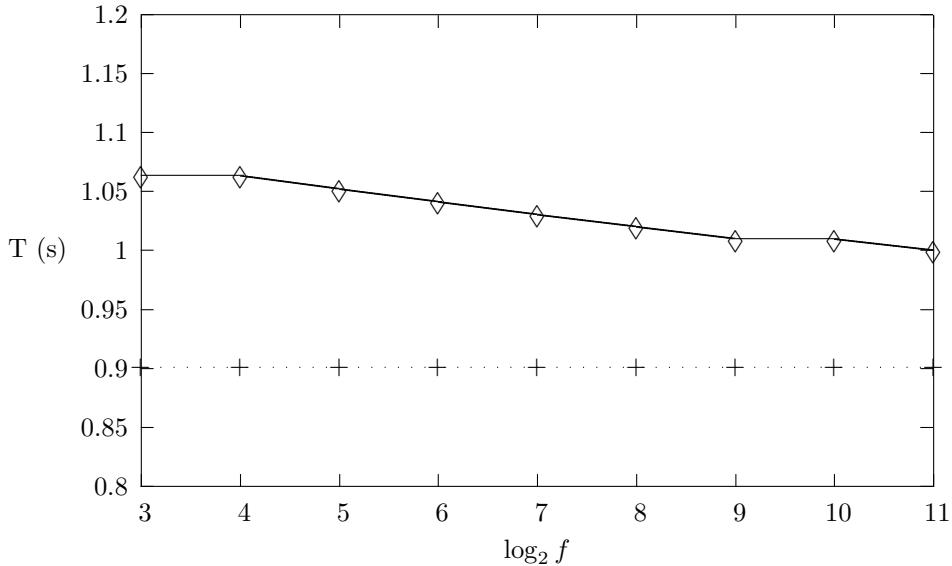


Figure 2.24. The period corresponding to intrinsic (diamond) and external (cross) peak as a function of f . The period of oscillatory source was kept as 0.9 s. As f is decreased, the intrinsic period approaches 1.08 s which is the period when only constant source is used. Different f was used to simulate the dynamics of a molecular network, and the timecourse of the number of free enzymes is Fourier transformed to obtain the positions of peaks. Other parameters were the same as figure 2.23.

noise (equation (2.6)) is varied to see, for example, if there is stochastic focusing [105]. Indeed, as shown in figure 2.26, the height of intrinsic and external peak changes with noise strength, and the ratio between the two can be varied qualitatively by tuning noise strength through, for example, temperature. If we replot figure 2.25 with a different noise strength (figure 2.27), it is obvious that the response of the system to periodic substrate source in terms of the ratio of the two peaks is enhanced by bigger noise intensity. In addition, the maximum in figure 2.26 might also be related to stochastic resonance [32].

The simulation in this section proposes a method to effectively change the collective dynamics of enzymes through the application of an external source of substrate. In addition, it is interesting to ask if nature utilizes this phenomenon to encode the information from a source to some output, because enzyme compartments in the cell may have the property of molecular networks. People have been using the radio wave to transport information for a long while, and the above simulation

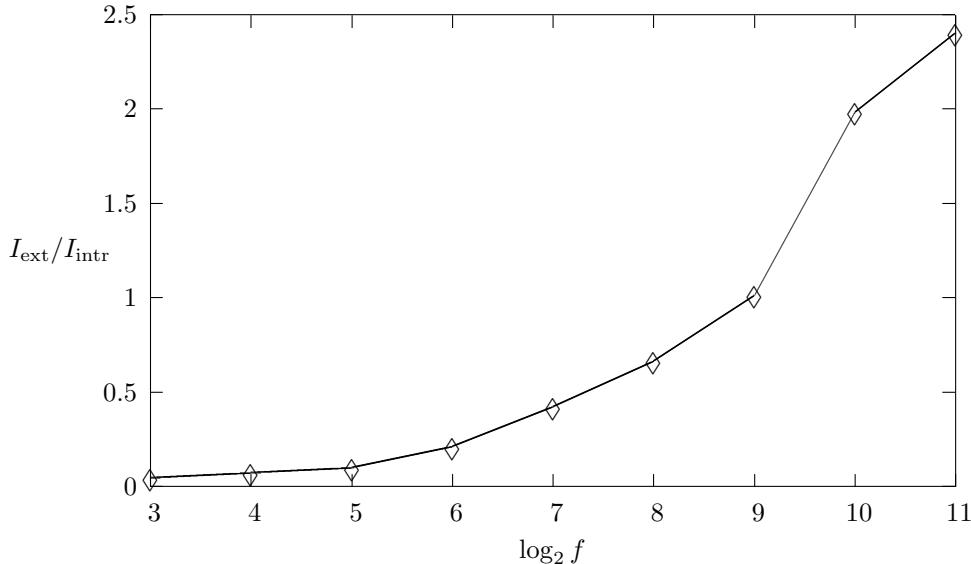


Figure 2.25. The ratio of the height I_{ext} of external peak over that (I_{intr}) of intrinsic peak as a function of f . Different f was used to simulate the dynamics of a molecular network and the timecourse of the number of free enzymes was Fourier transformed to obtain the height of peaks. Other parameters were the same as figure 2.23.

shows the possibility for a molecular network to encode and carry the information to other metabolic pathways. Interestingly enough, there is evidence that specificity is realized by enzyme's differential responses to signals of different frequencies [25], so the ability of preserve frequency information should be indispensable to biological systems. Of course, this simulation can and should be subject to experimental verification in the future.

2.3 Statistical Properties of Molecular Collisions

2.3.1 Analytical Result of Off-time Distribution

An intuitive way to study the effect of diffusion on biochemical reactions is to consider the statistical property of molecular collisions which are a prerequisite of all bimolecular reactions. As a first step, let us consider the collision between a mobile particle A and a static particle B (figure 2.28). According to Smoluchowsky theory, the probability of such an event in a small time interval dt is $k_D dt$, where $k_D = 4\pi r D_A/V$, determined by the distance r at contact between the two particles,

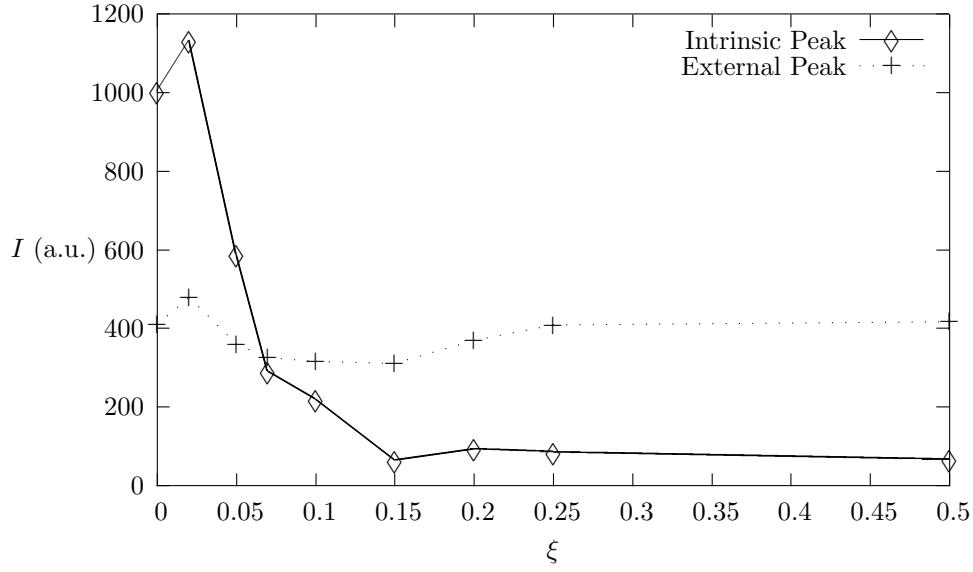


Figure 2.26. The heights (I) of external and internal peaks as functions of noise strength ξ when $f = 128 \text{ s}^{-1}$ and $T = 0.9 \text{ s}$. Other parameters were the same as figure 2.23.

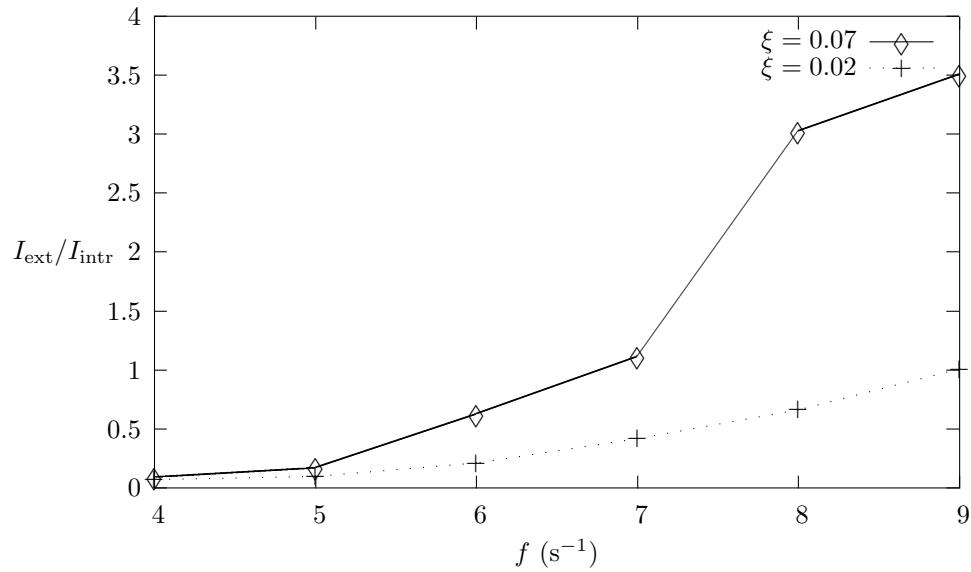


Figure 2.27. The ratio of the height I_{ext} of external peak over that (I_{intr}) of intrinsic peak as a function of f under different noise strength $\xi = 0.02$ or 0.07 . Different f and ξ were used to simulate the dynamics of a molecular network and the timecourses of the number of free enzymes were Fourier transformed to obtain the height of interested peaks. Other parameters were the same as figure 2.23.

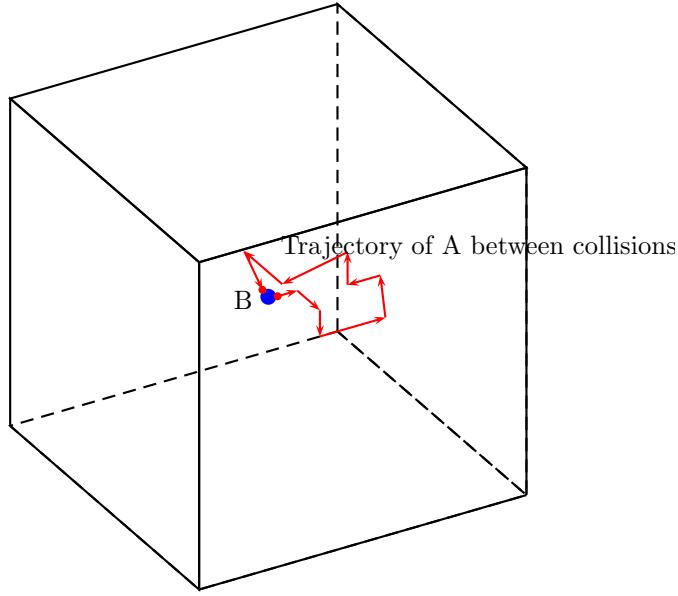


Figure 2.28. The collision between a static particle B (blue dot) and a mobile particle A (red dot) in a cubic reactor of size L .

the diffusion coefficient D_A of particle A, and the system volume V . In other words, the off-time t_{off} between consecutive collision events should follow an exponential distribution, or

$$p(t_{\text{off}}) \propto e^{-k_D t_{\text{off}}}. \quad (2.11)$$

A closer look at the process, however, suggests that after a collision event, particle A should have higher probability to collide with particle B than it would if placed far away from particle B [101]. Equation (2.11), in fact, implicitly assumes that the initial position of A is random, that is, particle A has no memory of its trajectory history. The collision–recollision clustering is of important biological consequence. For example, it may be responsible for the rapid recognition of gene sequence by transcription factors [52]. Quantitatively, the off-time distribution scales as $t_{\text{off}}^{-3/2}$ when $t_{\text{off}} \ll V^{2/3}/D$ [111]. The divergence at $t_{\text{off}} = 0$ is consistent with the intuition that the probability of a collision event immediately after the previous collision approaches certainty. Indeed, a numerical simulation verifies this power-law scaling in figure 2.29.

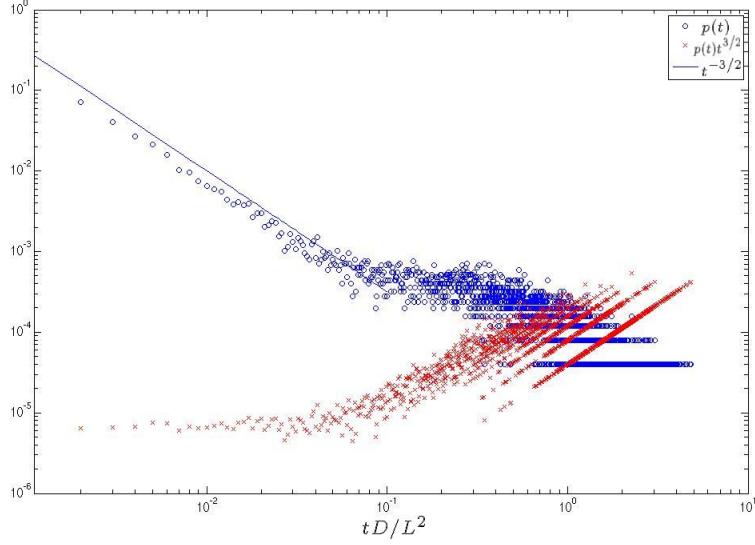


Figure 2.29. The off-time distribution obtained from Brownian motion simulation of the collision between a moving and a static particle in a cubic box of size L (for details, see section 2.3.2). The power-law scaling for small t is indicated by plotting $p(t)t^{3/2}$ as well.

A closely related problem is the collision of a molecule with the boundary of the reactor. Experimentally, one could observe the fluorescence from the transient association of a single Nile Red molecule, confined inside a unilamellar lipid vesicle, with the bilayer wall [33]. The off-time, defined as the dark time between fluorescence events, also follows different distribution in different regimes. To calculate this distribution, we consider the diffusion of a particle inside a spherical reactor with radius R (figure 2.30). Due to the symmetry of this system, the probability distribution c of the particle's position is only a function of its radial position r and time t . The initial condition, as a result, should also exhibit this symmetry as

$$c(r, t = 0) = \delta(r - r_0)/(4\pi r_0^2),$$

where r_0 is the initial position of the particle [111]. Here, $c(r, t)$ is the solution to the well-known diffusion equation,

$$\frac{\partial c}{\partial t} = D \nabla^2 c, \quad (2.12)$$

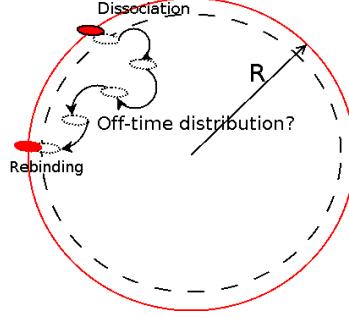


Figure 2.30. The collision between a diffusing particle and the reactor wall.

where D is the diffusion coefficient of particle A. The solution of equation (2.12) can be facilitated by a transformation of variable $u \equiv rc$. The new variable satisfies wave equation, or

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial r^2}.$$

and the following boundary conditions,

$$\begin{aligned} u(r, t=0) &= \frac{\delta(r - r_0)}{4\pi r_0}, \\ u(r=0, t) &= 0, \\ u(r=R, t) &= 0. \end{aligned}$$

By using Laplace transformation, one obtains its solution in Laplace space [111],

$$u(r, s) = \frac{\sinh(\sqrt{s/D}r_-) \sinh(\sqrt{s/D}(R - r_+))}{4\pi r_0 \sqrt{sD} \sinh(\sqrt{s/D}R)}, \quad (2.13)$$

where $r_- \equiv \min(r, r_0)$, and $r_+ \equiv \max(r, r_0)$. The dark time distribution can be identified with the flux out of the absorbing boundary,

$$p(t) = -4\pi R^2 D \left. \frac{\partial c}{\partial r} \right|_{r=R}. \quad (2.14)$$

From equation (2.13), one has

$$c(r, s) = \frac{\sinh(\sqrt{s/D}r_0) \sinh(\sqrt{s/D}(R-r))}{4\pi r r_0 \sinh(\sqrt{s/D}R)}, \quad (2.15)$$

for $r > r_0$. After each collision, the particle starts diffusion from a region very close to the boundary, so $r_0 \lesssim R$. This prompts us to work with a new variable $y \equiv R - r$ which is expected to be much smaller than R . In terms of y , the solution equation (2.15) is written as

$$c(y, s) = \frac{\sinh(\sqrt{s/D}(R-y_0)) \sinh(\sqrt{s/D}y)}{4\pi r_0(R-y) \sinh(\sqrt{s/D}R)}, \quad (2.16)$$

where $y_0 = R - r_0$. The off-time distribution is then evaluated using equation (2.14) and equation (2.16) to be

$$\begin{aligned} p(s) &= \frac{R}{R-y_0} \left[\cosh(\sqrt{s/D}y_0) - \sinh(\sqrt{s/D}y_0) \coth(\sqrt{s/D}R) \right], \\ &= \frac{R}{R-y_0} \frac{\sinh(\sqrt{s/D}(R-y_0))}{\cosh(\sqrt{s/D}R)}. \end{aligned} \quad (2.17)$$

It is instructive to consider two limiting cases of equation (2.17) whose inverse Laplace transforms admit analytical solutions. When $sR^2/D \gg 1$ and $sy_0^2 \ll 1$,

$$\begin{aligned} p(s) &\approx \frac{1}{1 - \frac{y_0}{R}} \left[1 + \frac{sy_0^2}{2D} - \sqrt{s/D}y_0 \right], \\ &\approx \left(1 + \frac{y_0}{R} + o(y_0^2) \right) \left(1 - \sqrt{s/D}y_0 + o(sy_0^2) \right), \\ &\approx 1 - \left(\sqrt{s/D} - 1/R \right) y_0 + o(y_0^2) + o(sy_0^2), \\ &\approx 1 - \sqrt{s/D}y_0. \end{aligned}$$

This implies $p(t) \propto t^{-3/2}$ when $y_0^2/D \ll t \ll R^2/D$. On the other end of the spectrum when

$$sR^2/D \ll 1,$$

$$\begin{aligned} p(s) &\approx \frac{R}{R - y_0} \left[1 + \frac{sy_0^2}{2D} - \frac{y_0}{R} \left(1 + \frac{sR^2}{2D} \right) \right], \\ &\approx 1 - \frac{sy_0 R}{2D}. \end{aligned}$$

This implies that $p(t) \propto e^{-t/(y_0 R/2D)}$ when $t \gg R^2/D$. Interestingly, the scaling behavior of off-time distribution is qualitatively similar to that of the first case we considered. There is some difference, though, when figure 2.29 is compared with figure 2.31 at intermediate t . The algebraic scaling for small t is illustrated in figure 2.31 obtained by numerical inverse of equation (2.17) [165].

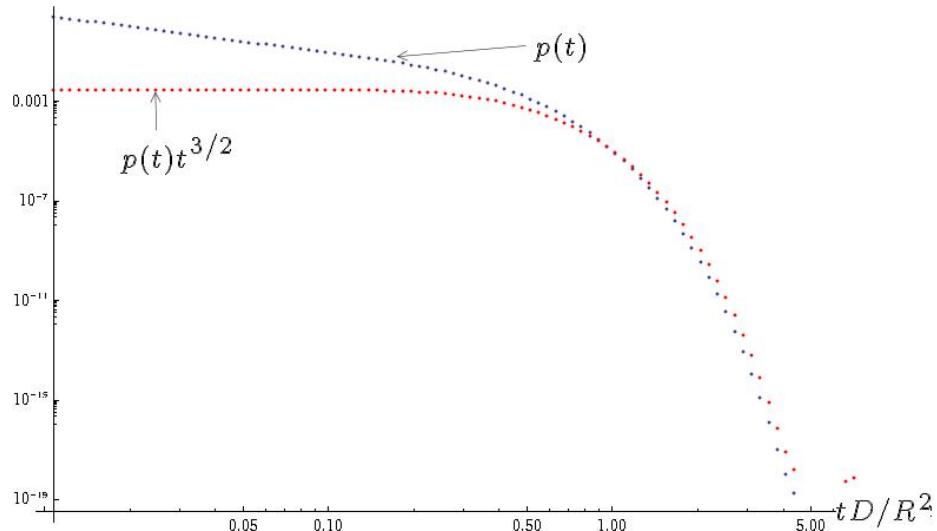


Figure 2.31. The off-time distribution exhibits power-law scaling when $t \ll R^2/D$. Off-time distribution was obtained by numerical inversion of equation (2.17) with $y_0/R = 0.01$.

2.3.2 Numerical Simulation of the Reactant–Reactant and Reactant–Wall Collision

2.3.2.1 Algorithms for Spatial Stochastic Reaction Simulation

Biochemical reactions rarely happen in a well-mixed reactor as assumed throughout section 2.1. Instead, the coupling of reaction, diffusion, and convection gives rise to spatial patterns in the distri-

bution of biomolecules. To resolve the spatial features in reaction network simulation, it is necessary to incorporate mass transport into the theoretical framework. As a first step, diffusion should be accounted for appropriately in the simulation algorithm. There are currently two approaches popular in literature with different emphasis on the discretization of space or time.

Space discretization is the approach to divide the reactor into small compartments where well-mixed condition is assumed to be valid. Diffusion between adjacent compartments is treated as a first-order reaction with rate constant deduced from diffusion coefficient. Although consistent with intuition, it has several practical and theoretical issues to be resolve before gaining wider applicability. First of all, the number of reactions is increased drastically by the inclusion of diffusion steps, although modified stochastic simulation algorithm was proposed to accelerate the process [55]. More importantly, there is some ambiguity in the choice of compartment size. On one hand, if it is too big, diffusion cannot mix reactants thoroughly within each compartments. On the other hand, if it is too small, there would be virtually no chance for bimolecular reactions to take place, because the possibility of both reactants being in the same compartment approaches zero. To solve this problem, people have proposed to modify the reaction rate according to the size of the compartments such that simulation result is less dependent on the fineness of spatial discretization [27].

Temporal discretization essentially simulates the Brownian motion trajectory of each molecule in the system. Different software packages [4, 140] are available to couple the diffusion with biochemical reactions. The spatial resolution of the algorithm is determined by the time step Δt used for simulation as well as the diffusivity of the molecules. Although the diffusion process can be faithfully reproduced, the correct treatment of reaction is not trivial. Recognizing that unimolecular reaction do not have spatial dependence and trimolecular reactions are extremely rare, one only needs to simulate bimolecular reactions appropriately. It is widely accepted that two molecules have to be close enough for reaction to take place and not every close encounter leads to product formation. As such, the interaction radius σ and the probability of reactive collision p (during Δt) are required to simulate bimolecular reactions even in this oversimplified model. Because it is difficult to estimate σ and p from first principle, one has to make the choice by some phenomenological arguments.

The most intuitive one is to stipulate that the effective rate constant calculated according to the

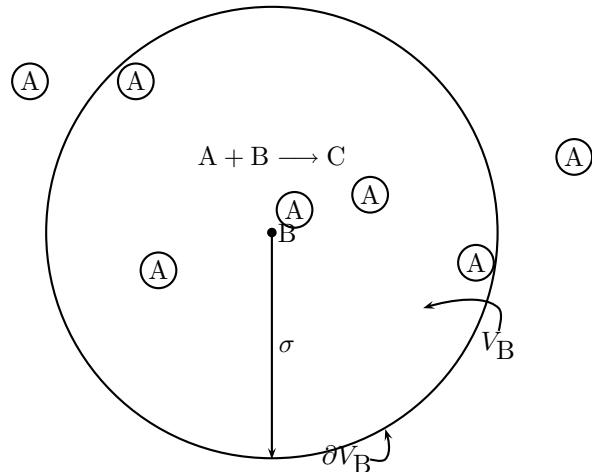
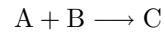


Figure 2.32. The influence sphere V_B of B is centered at B with radius σ . Molecule A may react with B if it is within the sphere.

simulation algorithm is consistent with bulk measurement results, which is normally supplied as an input to the simulation. Mathematically, one has to choose Δt , σ and p such that

$$k = f(\Delta t, \sigma, p),$$

where k is the rate constant for the bimolecular reaction of interest with unit m^3/s . The exact functional form of f will be discussed in the context of heteroreaction



following the discussion in reference [27]. For convenience, we will generally consider the movement of A relative to B. Furthermore, we fix our attention on a single B located at the center of our coordinate system, and A molecules are floating around with an effective diffusion coefficient of $D = D_A + D_B$, where D_A and D_B are the diffusion coefficients for A and B in lab frame, respectively.

When $\Delta t \rightarrow 0$, the simulation approximates the infinitely detailed Brownian dynamics coupled with reaction inside B's influence sphere V_B (figure 2.32). The steady-state concentration of A is

the solution to the following reaction–diffusion equation

$$\begin{aligned} D\nabla^2 c &= 0 \text{ for } r > \sigma, \\ D\nabla^2 c - \lambda c &= 0 \text{ for } r < \sigma, \end{aligned} \quad (2.18)$$

where r is the radial coordinate, and $\lambda = p/\Delta t$ is the reaction rate inside the sphere. The reaction rate due to the removal of A can be expressed in terms of the total flux of A through the surface ∂V_B of the influence sphere V_B as

$$\begin{aligned} \text{Reaction rate} &= \frac{N_B}{V} \int_{V_B} \lambda c dV, \\ &= c_B \int_{V_B} D\nabla^2 c dV, \\ &= c_B D \int_{\partial V_B} \left. \frac{\partial c}{\partial r} \right|_{r=\sigma} dS, \end{aligned}$$

where N_B is the total number of B in a system of volume V and c_B is the concentration of B. The solution to equation (2.18) is straightforward, and the resulting reaction rate divided by the bulk concentration of A and B gives the desired expression for k :

$$k(\Delta t \rightarrow 0) = 4\pi D \left(\sigma - \sqrt{D/\lambda} \tanh \left(\sigma \sqrt{\lambda/D} \right) \right). \quad (2.19)$$

In particular, if all the collisions are reactive, that is, $p = 1$ or $\lambda \rightarrow \infty$, Equation (2.19) reduces to the well-known Smoluchowsky equation $k(\Delta t \rightarrow 0) = 4\pi D\sigma$.

When $\Delta t \rightarrow \infty$, the simulation process can no longer be described by the reaction–diffusion equation. During Δt , the concentration gradient of A is removed by diffusion. Reactions are simulated at the end of the diffusion step by removing A inside the influence sphere according to probability p . In another word, the reaction rate is proportional to the size of V_B :

$$\begin{aligned} \text{Reaction rate} &= p \frac{N_B \frac{4}{3}\pi\sigma^3 N_A}{V \Delta t}, \\ &= \frac{4}{3}\pi\sigma^3 c_A c_B \frac{p}{\Delta t}. \end{aligned}$$

Here, c_A and c_B are the bulk concentration of A and B, respectively, so the rate constant is

$$k(\Delta t \rightarrow \infty) = \frac{\frac{4}{3}\pi\sigma^3 p}{\Delta t}, \quad (2.20)$$

For intermediate values of Δt , no analytical solution is available yet. The following numerical scheme can be used to correlate k with σ , p , and Δt . Let us consider the effect of diffusion during Δt . Reactions not being simulated in this period, diffusion equation can be used to describe the spatial distribution of A around a single B. To render the equation dimensionless, we scale c by the bulk value c_A , r by the interaction radius σ , and t by σ^2/D . The diffusion equation now reads

$$\frac{\partial g(\rho, \tau)}{\partial \tau} = \nabla_\rho^2 g(\rho, \tau), \quad (2.21)$$

where $g \equiv c/c_A$, $\rho \equiv r/\sigma$, and $\tau \equiv tD/\sigma^2$ are dimensionless variables. The boundary conditions are

$$g(\rho \rightarrow \infty) = 1,$$

$$g(\rho = 0) < \infty.$$

Our question is to determine the distribution $g(\rho, \tau + \Delta\tau)$ given an arbitrary $g(\rho, \tau)$. This can be achieved by obtaining the Green's function $G(\rho, \rho'; \Delta\tau)$ which is the solution to equation (2.21) subjected to the initial condition $g(\rho, \tau = 0) = \delta(\rho - \rho')/(4\pi\rho'^2)$. As before, the spherical symmetry allows for a convenient change of variable $u \equiv g\rho$, and the differential equation for u ,

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial \rho^2},$$

with initial condition $u(\rho, \tau = 0) = \delta(\rho - \rho')/(4\pi\rho')$ and boundary condition $u(\rho = 0, \tau) = 0$ can be solved to obtain the Green's function. In the end, one has

$$G(\rho, \rho'; \Delta\tau) = \frac{1}{\rho\rho' \sqrt{4\pi\Delta\tau}} e^{-\frac{(\rho-\rho')^2}{4\Delta\tau}} - \frac{1}{\rho\rho' \sqrt{4\pi\Delta\tau}} e^{-\frac{(\rho+\rho')^2}{4\Delta\tau}}.$$

The evolution of $g(\rho, \tau)$ can be expressed using Green's function

$$g(\rho, \tau + \Delta\tau) = \int_0^\infty G(\rho, \rho'; \Delta\tau)g(\rho')4\pi\rho'^2 d\rho'.$$

The effect of reaction is to remove molecule A within the influence sphere with probability p . It can be realized before or after the diffusion step. After some iteration, $g(\rho)$ will approach a steady state which satisfies the following equation

$$\begin{aligned} g(\rho) &= (1-p) \int_0^1 G(\rho, \rho'; \Delta\tau)g(\rho')4\pi\rho'^2 d\rho' \\ &+ \int_1^\infty G(\rho, \rho'; \Delta\tau)g(\rho')4\pi\rho'^2 d\rho' \quad (\text{reaction first}) \end{aligned} \quad (2.22)$$

$$\begin{aligned} g(\rho) &= \int_0^\infty G(\rho, \rho'; \Delta\tau)g(\rho')4\pi\rho'^2 d\rho' \\ &- pH(1-\rho) \int_0^\infty G(\rho, \rho'; \Delta\tau)g(\rho')4\pi\rho'^2 d\rho' \quad (\text{diffusion first}), \end{aligned} \quad (2.23)$$

where $H(1 - \rho)$ is the Heaviside function. To be self-consistent, the solutions to these two equations should be related to each other by a diffusion step. In fact, equations (2.22) and (2.23) can be rewritten in matrix format if $g(\rho)$ is discretized and then expressed as a column vector $\mathbf{g} = \{g(\rho_i)\}$ where $\rho_i \in (0, \infty)$:

$$\mathbf{g}_1 = \mathbf{DRg}_1 \quad (\text{reaction first}) \quad (2.24)$$

$$\mathbf{g}_2 = \mathbf{RDg}_2 \quad (\text{diffusion first}). \quad (2.25)$$

Here the effect of reaction and diffusion is accounted for by the operation of matrices \mathbf{R} and \mathbf{D} , respectively. Multiplying equation (2.25) with \mathbf{D} from left, one immediately realizes that \mathbf{Dg}_2 is a solution to equation (2.24). Indeed, if the solution to equation (2.23) is further evolved by a diffusion step, the result is the same as the solution to equation (2.22) (figure 2.33).

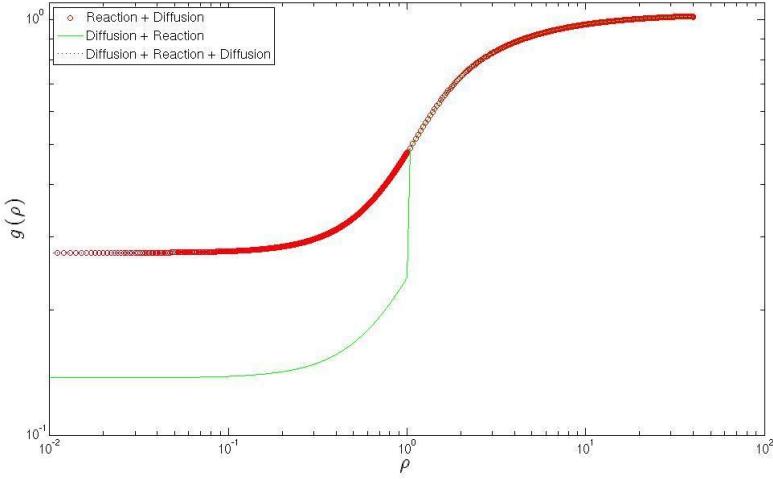


Figure 2.33. The solution to equation (2.22) (red circle, reaction followed by diffusion) and equation (2.23) (green line, diffusion followed by reaction). The latter is further evolved by a diffusion step to obtain the blue dotted line which is equivalent to the solution of equation (2.22). Parameters: $p = 0.5$ and $\sqrt{2\Delta\tau} = 0.5$. Discretization: 1000 points each in $(0, 1]$ and $(1, 40]$ and $g(\rho > 40) \equiv 1$.

The solution to equation (2.22) or equation (2.23) can then be used to compute reaction rate

$$\begin{aligned}\text{Reaction rate} &= \frac{N_B}{V\Delta t} p \int_0^\sigma c_A g(r/\sigma) 4\pi r^2 dr, \\ &= \frac{c_A c_B}{\Delta t} p \int_0^1 g(\rho) 4\pi \rho^2 \sigma^3 d\rho.\end{aligned}$$

The rate constant is thus

$$k = \frac{\sigma^3 p}{\Delta t} \int_0^1 g(\rho) 4\pi \rho^2 d\rho. \quad (2.26)$$

In summary, given p , σ , and Δt , one could solve for $g(\rho)$ from equation (2.22) (or equation (2.23) supplemented by an extra diffusion step) and then k is related to these parameters by equation (2.26).

In particular, when $\Delta t \rightarrow \infty$, $g(\rho)$ approaches unity, and equation (2.26) reduces to equation (2.20) as expected.

It is also possible to derive such a relation on the basis of other diffusion models. Based on

Table 2.4. Classification of spatial stochastic simulation algorithms

Methods	MCell [140]	Smoldyn [3]	reference [27]
Independent parameter	$\sigma, \Delta t$	$p(=1), \Delta t$	Δt , either p or σ
Dependent parameter	p	σ	p or σ

ray-tracing method, MCell [140] uses the following expression

$$\begin{aligned} k &= p\pi\sigma^2 \left(\frac{4D}{\pi\Delta t} \right)^{1/2}, \\ &= \sqrt{2\pi} \frac{\sigma^3 p}{\Delta t} \frac{\sqrt{2D\Delta t}}{\sigma}, \end{aligned} \quad (2.27)$$

to correlate p , σ and Δt with k . The last expression (equation (2.27)) suggests that the reduced reaction rate $\kappa \equiv k\Delta t/\sigma^3$ is proportional to p and $\gamma \equiv \sqrt{2D\Delta t}/\sigma$, the latter of which is the root-mean-square displacement of A (relative to B) scaled by σ . This scaling is consistent with the result of Green's function approach [27] when γ is big.

In all of the schemes discussed above, it is not sufficient to fully determine all of the three parameters required for the simulation using a single constraint on k . In practice, user would normally choose Δt first according to the desired temporal and spatial resolution; a rule of thumb is that the spatial resolution scales as $\sqrt{D\Delta t}$. Then either p or σ can be freely adjusted. Available literatures differ in their choices of user inputs as summarized in Table 2.3.2.1.

2.3.2.2 Numerical Results from Smoldyn Simulation

Smoldyn was used to study the collision between reactants or reactant and wall in a confined space. The source code of Smoldyn was slightly modified to report the position and time of each collision event. In a typical simulation, an enzyme, with diffusion coefficient $7 \times 10^{-11} \text{ m}^2/\text{s}$, and a substrate, with diffusion coefficient $4.4 \times 10^{-10} \text{ m}^2/\text{s}$, were placed randomly in a cubic box. By design, every collision leads to reaction in Smoldyn simulation, so one can conveniently follow the collision between molecules. Figure 2.34(a) shows the positions of the collision between a pair of substrate and enzyme in a 60 nm box. Clearly, some of the collision events cluster at a “hot spot.” To determine if they also cluster in time domain, we also plot the time of the collision events in figure 2.34(b). Indeed, the

30,000th to 70,000th collision happen during a short time window of less than 0.5 μ s. The collision frequency in this short period of time is much higher than the average collision frequency. This was also reported by others [15]. The close proximity of the molecules to the edge or corner of the reactor is expected to be the reason for the clustering of collisions. If we consider the diffusion of substrate molecule relative to the enzyme, after a collision event the asymptotic probability of the substrate returning to the enzyme for another collision is zero. But if they are temporarily confined by the edge or corner of the reaction, repetitive collision is more likely to happen, and this will have important implications on the reaction rate.

Similar results were also obtained when we considered the collision between a diffusing molecule and the reactor wall (figure 2.35). Repetitive collisions take place in clusters, and this may effectively increase the interaction between the reactor wall and the diffusing molecules. When the reactor size is decreased, this interaction may result in the significant nonspecific adsorption of macromolecules to the wall and change the activity of the molecule due to either denaturation or size-exclusion effect. The efficiency of surface passivation is thus required to be optimized further in micro- or nanoreactors. Alternatively, more frequent interaction between diffusing molecules and the reactor wall may significantly modify the apparent reaction kinetics, as shown recently in the oxidation of Amplex Red catalyzed by horseradish peroxidase confined in femtoliter chambers [48].

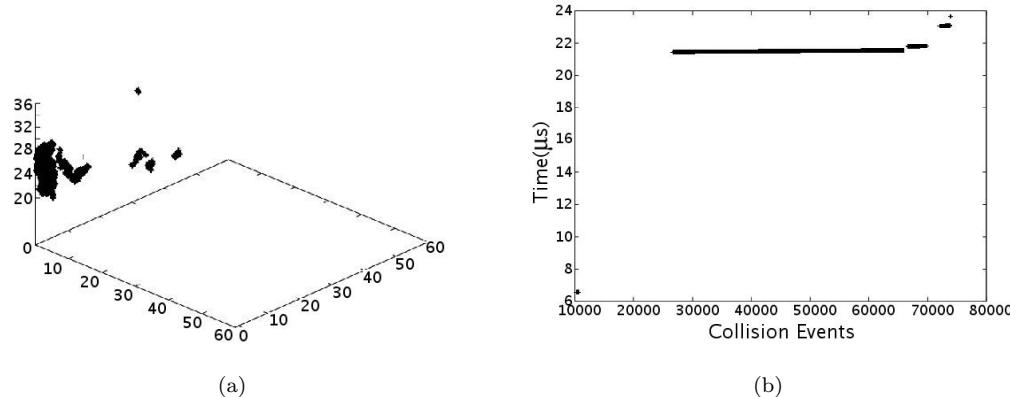


Figure 2.34. The position (a) and time (b) of the collision events between a pair of substrates and enzymes. The box was 60 nm \times 60 nm \times 60 nm. The time step was 1.5 ps. The interaction radius was 3.5 nm.

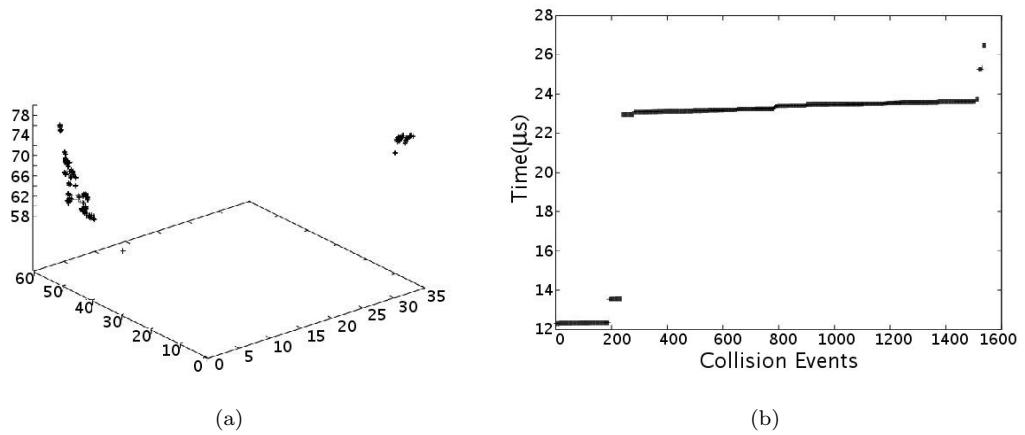


Figure 2.35. The position (a) and time (b) of the collision events between an enzyme and reactor wall. The box was $100 \text{ nm} \times 100 \text{ nm} \times 100 \text{ nm}$. The time step was 1.5 ps.