

UNIVERSITY OF CALIFORNIA
Santa Barbara

Space in Discrete Stochastic Simulation of
Chemical Kinetics

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Computer Science

by

Sotiria Lampoudi

Committee in Charge:

Professor Linda R. Petzold, Chair

Professor Frank L. Brown

Professor John R. Gilbert

December 2008

The Dissertation of
Sotiria Lampoudi is approved:

Professor Frank L. Brown

Professor John R. Gilbert

Professor Linda R. Petzold, Committee Chairperson

September 2008

Space in Discrete Stochastic Simulation of Chemical Kinetics

Copyright © 2008

by

Sotiria Lampoudi

Dedicated to the memory of my two
grandfathers, who both gave me solid advice
before I left the nest.

Acknowledgements

I would like to acknowledge the intellectual contributions and support of Dr. Dan Gillespie, without whom none of the work presented in this thesis would have happened.

Curriculum Vitæ

Sotiria Lampoudi

Education

- 2004 Master of Science in Computer Science, University of California, Santa Barbara.
- 2002 Bachelor of Arts in Computer Science, University of Chicago.

Experience

- 2003 – 2008 Graduate Research Assistant, University of California, Santa Barbara.
- Summer 2007 Associate Lecturer, University of California, Santa Barbara.
- 2002 – 2003 Teaching Assistant, University of California, Santa Barbara.

Selected Publications

S. Lampoudi, D. T. Gillespie, L. R. Petzold, “The multinomial diffusion method for discrete stochastic simulation of reaction-diffusion systems,” in preparation for *J Chem Phys*, August 2008.

S. Lampoudi, D. T. Gillespie, L. R. Petzold, “Effect of reactant size on 2D discrete stochastic chemical kinetics,” submitted to *Phys Rev E*, December 2007.

D. T. Gillespie, **S. Lampoudi**, L. R. Petzold, “Effect of reactant size on discrete stochastic chemical kinetics,” *J Chem Phys* 126(3):034302 (2007).

Abstract

Space in Discrete Stochastic Simulation of Chemical Kinetics

Sotiria Lampoudi

This thesis is concerned with studying what happens when space must be taken into account in discrete stochastic simulation of chemical kinetics. The Stochastic Simulation Algorithm of Gillespie, which is the simulation technique of choice for mesoscopic stochastic chemical kinetics, is rigorously applicable to systems of point molecules which are well-stirred. The former condition means that the molecules must occupy volume negligible compared to the total volume of the system. The latter condition means that the distribution of the molecules' positions in space must be uniformly random.

The SSA has been steadily gaining ground as a method of simulating intracellular kinetics. However, one or both of the conditions which make it rigorously applicable are frequently not met in that biological context. In the first section of this thesis we present our results regarding relaxing the point molecule condition. We find that, even when molecules are allowed to exclude substantial volume, the distribution of inter-collision times is still close to exponential, making the SSA still applicable once appropriate propensity functions are chosen. In the second

section of the thesis we concern ourselves with relaxing the well-stirred condition. The Inhomogeneous SSA is applicable in that situation, but it may not be a realistic choice because of its computational cost. We present the Multinomial Simulation Algorithm for stochastic reaction-diffusion, which interlaces approximate stochastic diffusion with reactions simulated according to the SSA.

Professor Linda R. Petzold
Dissertation Committee Chair

Contents

Acknowledgements	ix
Curriculum Vitæ	xi
Abstract	xiii
List of Figures	xvii
List of Tables	xix
1 Introduction	1
1.1 Motivation, Scope and Outline	1
1.2 Background: the Stochastic Simulation Algorithm	3
2 Relaxing the Point Molecule Assumption	9
2.1 Introduction	9
2.2 The Effect of Excluded Volume in One Dimension	13
2.2.1 The Dilute Case	16
2.2.2 The Dense Case	18
2.2.3 Simulation Results	25
2.2.4 Discussion	30
2.3 The Effect of Excluded Volume in Two Dimensions	33
2.3.1 Hard Sphere Molecular Dynamics	36
2.3.2 Theory for the two-dimensional case	43
2.3.3 Simulation Results	46
2.3.4 Discussion	62

3	Relaxing the Homogeneity Assumption	63
3.1	Introduction	63
3.2	The Inhomogeneous SSA	64
3.3	The Multinomial Method	70
3.3.1	Motivation	70
3.3.2	Diffusion in one dimension	74
3.3.3	Reaction-diffusion in one dimension	99
3.3.4	Reaction-diffusion in two dimensions	107
3.3.5	Discussion	108
	Bibliography	117
	Appendices	121
A	Uniform distribution of hard rods on a fixed interval	123
A.1	The distribution of centers	123
A.2	The gap distribution	126
A.3	Monte Carlo placement of the rods	129

List of Figures

2.1	1D: Simulation vs. analytical distributions for $N = 1000$	28
2.2	1D: Simulation vs. analytical distributions for $N = 4$	29
2.3	1D: Simulation vs. analytical distributions for $N = 2$	29
2.4	2D: Convergence of approximate molecule placement algorithm . .	41
2.5	2D: Diagram of area swept by one molecule relative to another. .	44
2.6	2D: Exponentiality of the τ -distribution.	47
2.7	2D: Piecewise histogram of dilute system.	49
2.8	2D: Piecewise histogram of dense system.	54
2.9	2D: Exponentiality indicator for the τ -distribution.	55
2.10	2D: Approximate results for high population and density.	57
2.11	2D: Effective excluded volume and inverse packing fraction.	60
3.1	Boundary conditions and the resulting graphs	77
3.2	The error for $s = 1$ and systems with $n = 4, 5, 6, 8$	86
3.3	$p_{11}^{(n)}(\kappa\Delta t)$ and $p_{12}^{(n)}(\kappa\Delta t)$ for systems with $n = 4, 5, 6, 8$	87
3.4	The error for $s = 2$ and $p_{13}^{(n)}(\kappa\Delta t)$ for $n = 6, 8, 10$	98
3.5	1D Annihilation: Mean population for ISSA, MSA(1), MSA(2) . .	110
3.6	1D Annihilation: K-distance between ISSA and MSA(1), ISSA and MSA(2)	111
3.7	1D Fisher: The ratio (diffusion)/(reaction)	112
3.8	1D Fisher: Speedup of the MSA(1) and MSA(2) over ISSA. . . .	113
3.9	1D Fisher: Space-averaged K-distance	114
3.10	2D Annihilation: Mean population for ISSA, MSA(1)	115
3.11	2D Annihilation: K-distance between ISSA and MSA(1)	116
A.1	Diagram of hard rods in a finite interval	123

List of Tables

Chapter 1

Introduction

1.1 Motivation, Scope and Outline

The Stochastic Simulation Algorithm (SSA) is the workhorse algorithm of discrete stochastic chemical kinetics. The SSA was first formulated in 1976 [19], and interest in it has surged in the biomolecular simulation community since 1998 [3], when it began to be widely recognized that species present in cells at low copy numbers were playing important regulatory roles.

The SSA has been rigorously derived for *point molecules* in a *well-stirred*, dilute gas setting [22]. The point molecule assumption stipulates that the volume occupied by the reactants must be negligible compared to the volume of the system. The well-stirred assumption requires the positions of the molecules to be uniformly randomly distributed inside the system volume.

The SSA has been increasingly applied to simulating biomolecular kinetic pathways inside cells. In many cases, such application is well justified. In general, however, the intracellular environment is neither dilute, nor is it always well-stirred. The interior of cells is over-stuffed with chemically active species, many of which are large proteins [13]. In addition, some species are purposely localized to specific locations in the cell [5], while some are present at such low copy number that their location simply cannot be considered uniformly random over the timescales of the reactions occurring in the system.

In this thesis we study what happens when we relax the point molecule and well-stirred assumptions. We study the relaxation of the point molecule assumption in a very simple single species system, in one and two dimensions. This allows us to assess the impact of the volume excluded by the reactants on the propensity functions of the SSA.

The Inhomogeneous SSA (ISSA) is the extension of the SSA which can be applied when the well-stirred assumption does not hold, or in the case of systems for which information about the locations of the reactants is desired. Unfortunately, in many cases the ISSA is too time consuming to be applicable to systems of *practical* interest. We present the Multinomial Simulation Algorithm as an alternative to the ISSA for systems in which diffusion is much faster than reactions.

This thesis is organized as follows: the Stochastic Simulation Algorithm is described in the next section (Section 1.2). Relaxation of the point molecule assumption is studied in Chapter 2 (the results on the effect of excluded volume in one dimension are given in Section 2.2, and the results for two dimensions are given in Section 2.3). Chapter 3 deals with algorithms which address inhomogeneity (the ISSA is described in section 3.2, while the Multinomial Simulation Algorithm is presented in Section 3.3).

1.2 Background: the Stochastic Simulation Algorithm

The problem of simulating the time evolution of a system of coupled chemical reactions in the mesoscopic discrete stochastic setting formulated as follows:

Given a vector \mathbf{X} of N species $\{S_1, \dots, S_N\}$ participating in M chemical reactions $\{R_1, \dots, R_M\}$, if at time t_0 the population vector is \mathbf{x}_0 , determine the population vector \mathbf{x}_t at any later time t .

The vector \mathbf{X} is usually called the *state*. The set of chemical reactions is often called a *network*, in recognition of the fact that many of the species involved participate in multiple reactions either as reactants or products.

The effect of each reaction can be concisely denoted by a vector with components representing the amount by which the state vector changes as a result of a

single firing of the reaction. That is, if the k -th reaction has the form:



it will be described by a vector of the form

$$(0, \dots, 0, r_{ki}, 0, \dots, 0, r_{kj}, 0, \dots, 0) \quad (1.2)$$

where the r_{ki} occurs in the i -th position, the r_{kj} occurs in the j -th position and $r_{kl} = 0$ for all l except i and j . The r are often called *stoichiometric coefficients*.

All M stoichiometric vectors, treated as row vectors, can be collected in a matrix $\boldsymbol{\nu}$ called the *stoichiometric matrix*, which describes the entire reaction network:

$$\boldsymbol{\nu} = \begin{pmatrix} r_{11} & \dots & r_{1N} \\ & \vdots & \\ r_{M1} & \dots & r_{MN} \end{pmatrix} \quad (1.3)$$

The formula for advancing the state of the system \boldsymbol{x}_t at time t to $\boldsymbol{x}_{t+\Delta t}$ at time $t + \Delta t$ if one R_k reaction has taken place in $[t, t + \Delta t)$ is simply:

$$\boldsymbol{x}_{t+\Delta t} = \boldsymbol{x}_t + \boldsymbol{\nu}_k \quad (1.4)$$

where ν_k is the k -th row of ν .

Suppose our system, described by \mathbf{x}_0 at time t_0 , is contained in the volume Ω . Suppose further that the volume can be considered to be *well-stirred*, meaning that the location of a given molecule is uniformly random within Ω . We would like to have an expression for the following quantity:

$$\begin{aligned} a_j(\mathbf{x})dt &= \text{the probability that channel } R_j \\ &\text{will fire once in } [t, t + dt), \text{ given} \\ &\text{that } X(t) = \mathbf{x} \end{aligned} \tag{1.5}$$

It turns out that this quantity is well defined. $a_j(\mathbf{x})$ is called a (*reaction*) *propensity function* and has the form:

$$a_j(\mathbf{x})dt = c_j h_j(\mathbf{x}) \tag{1.6}$$

where c_j is the *specific probability rate constant*, defined such that $c_j dt$ gives the probability that a randomly chosen pair of R_j reactant molecules will react in time dt . The function $h_j(\mathbf{x})$ is defined to be the number of distinct combinations of R_j reactant molecules possible in state \mathbf{x} .

By using the standard stochastic process technique of conditioning it is possible to obtain this expression for the time evolution of a chemically reacting system:

$$\begin{aligned} \frac{\partial}{\partial t} P(\mathbf{x}_t, t | \mathbf{x}_0, t_0) &= \sum_{j=1}^M [a_j(\mathbf{x}_t - \boldsymbol{\nu}_j) P(\mathbf{x}_t - \boldsymbol{\nu}_j, t | \mathbf{x}_0, t_0) \\ &\quad - a_j(\mathbf{x}_t) P(\mathbf{x}_t, t | \mathbf{x}_0, t_0)], \end{aligned} \quad (1.7)$$

This formula is known as the *Chemical Master Equation*. For any but the simplest systems its extremely large dimensionality makes it very difficult to solve, analytically or numerically. It is possible, however, to generate samples, specifically sample trajectories, from it.

The Stochastic Simulation Algorithm [19] is a rigorous way to generate sample trajectories of the CME. The heart of the algorithm is the *next reaction density function*:

$$\begin{aligned} p(\tau, j | x, t) d\tau &= \text{the probability that, given } X(t) = \mathbf{x}, \\ &\text{the next reaction in } \Omega \text{ will occur in} \quad (1.8) \\ &[t, t + d\tau) \text{ and will be a } R_j \text{ reaction} \end{aligned}$$

The formula for the next reaction density function turns out to be:

$$p(\tau, j | x, t) = a_j(\mathbf{x}) \exp\left(-\sum_{k=1}^M a_k(\mathbf{x}) \tau\right) \quad (1.9)$$

Generating random pairs (τ, j) from the distribution (1.9) and advancing the state of the system accordingly allows us to sample one possible trajectory of the CME. Collecting the endpoints of multiple such trajectories allows us to sample the CME along a different dimension, this time at a specific t . In simulation terminology, one trajectory, from time t_0 to time t_f is called a *realization* or *replica*. Multiple trajectories comprise an *ensemble*.

In order to generate random pairs according to (τ, j) we produce two uniformly random numbers r_1 and r_2 in $(0, 1]$ and compute τ and j respectively using the formulas:

$$\begin{aligned} \tau &= \frac{1}{a_0(\mathbf{x}_t)} \ln \left(\frac{1}{r_1} \right) & (1.10) \\ j &= \text{smallest integer in } [1, M] \text{ such that } \sum_{j'=1}^j a_{j'}(\mathbf{x}_t) > r_2 a_0(\mathbf{x}_t). \end{aligned}$$

where we use the shorthand $a_0 = \sum_{j=1}^M a_j$.

The procedure for generating (τ, j) above is an implementation of the SSA called the *Direct Method*. There exist two more methods for generating (τ, j) that are equally correct, but have different computational costs associated with them.

The *First Reaction Method*, also proposed by Gillespie, generates a putative reaction time τ_j for each channel. It then chooses the smallest among those, and realizes that reaction. The remaining τ_j are discarded. The method requires the

generation of M random numbers (rather than the two of the Direct Method) and so is usually less efficient than the Direct Method.

The *Next Reaction Method*, proposed by Gibson and Bruck [17], is similar to the First Reaction Method, in that it uses M putative reaction times. It is more efficient than the First Reaction Method because it reuses many of the random numbers generated. The NRM employs two tricks that in principle can be used independently. First, it uses a dependency graph between reactions to keep track of which reaction propensities must be recomputed if a certain reaction has just taken place. Second, it uses an (indexed) priority queue data structure in which reactions are ordered according to their putative firing times, and a scheme for reusing unused putative firing times. The top event in the queue (which has the smallest putative firing time) is realized. The dependency graph is used to update only those propensities and putative firing times which are affected by the event that took place. The newly computed putative firing times are modified in place in the queue, one-by-one, and allowed to bubble up or down to their proper place. Then these steps are repeated.

Chapter 2

Relaxing the Point Molecule Assumption

2.1 Introduction

The Stochastic Simulation Algorithm (SSA) [19] is the workhorse algorithm for discrete stochastic simulation of networks of coupled chemical reactions. The physical system, in this case, is a collection of molecules of various chemical species that move around inside a fixed volume, and are subject to a set of chemical reactions in which the molecules may be reactants or products or both. The chemical reactions are all assumed to be “elementary” in the sense that they occur essentially instantaneously. Elementary reactions will invariably be either unimolecular or bimolecular; all other types of reactions (trimolecular, reversible, etc.) will consist of a series of two or more elementary reactions. If the system is well-stirred, we can define its state simply by giving the vector \mathbf{x} of the molecular populations of

the various chemical species. In this circumstance, we recall from Chapter 1 that the dynamics of each reaction channel R_j is described by a "propensity function" $a_j(\mathbf{x})$, defined so that if the system is in state \mathbf{x} , then $a_j(\mathbf{x})dt$ gives the probability that the reaction will occur somewhere inside the system in the next infinitesimal time interval dt . The magnitude of $a_j(\mathbf{x})$ thus measures the "propensity" of reaction R_j to occur in the immediate future.

The propensity function is very close to, and sometimes numerically equal to, what in deterministic chemical kinetics is called the "reaction rate". But the propensity function does not make the assumption that reactions occur continuously and deterministically, and its product with dt is mathematically treated as a probability. In the thermodynamic limit (infinite populations and infinite system volume with finite concentrations), the CME (Eq. 1.7) and SSA almost always reduce to the ordinary differential equations of deterministic chemical kinetics.

The SSA generates times τ between successive reactions as samples of an exponential distribution whose mean is equal to the inverse of the sum of the propensity functions. The most commonly used propensity functions are of a mass action form, according to which the rate of a reaction is proportional to the combinatorial product of the reactants' populations.

Mass action propensity functions for elementary reactions have been rigorously derived in a well-stirred, dilute hard sphere setting [22]. In this setting molecules

are represented by hard spheres moving ballistically in a vacuum. We refer to them as point molecules, because, although they must have non-zero diameter l in order to collide, the volume of all the molecules combined is negligible compared to the volume of their container. Our work addresses the question: if the point molecule assumption is relaxed, to what extent does the volume occupied by the reactant molecules themselves affect the rates of the reactions in which they participate?

We will be studying the effect of reactant-excluded volume in a simple but computationally tractable physical model. Specifically, we will attempt to answer the following two questions. First, is the time between successive reactions in a well-stirred, non-point molecule system exponentially distributed, as it must be for the stochastic process theory which underlies the SSA to hold? Second, if the reaction times are exponentially distributed, what is the mathematical form of the propensity functions in this setting?

Since bimolecular reactions are always initiated by a collision, the probability of a reaction between two molecules can be broken down into a) the probability that the two molecules will collide, times b) the probability that they will react given that they have collided. Throughout this work we make the simplifying assumption that (b) is unity, and thus use the terms *collision* and *reaction* probability (and inter-collision and inter-reaction time) interchangeably.

In Section 2.2 we show how, for a one dimensional system, the mass action propensity functions need to be modified when the volume of the reactant molecules is comparable to the total system volume [23]. We have analytically derived an exact formula for the reaction probability, in the next infinitesimal time dt , of the reaction $A + A \rightarrow \text{products}$ in a one-dimensional system of N non-overlapping hard rods of length l moving ballistically in a volume of length L . We then confirmed our analytical derivation using an extensive series of Hard Sphere Molecular Dynamics simulations.

An analytical treatment of the two-dimensional hard disc system has proved to be challenging. The difficulty arises when trying to find an analytical intermolecular distance distribution function for non-overlapping, non-zero sized hard disks in a finite area. To the best of our knowledge, a two or three-dimensional exact result has not been reported in the literature, and we have not been able to derive it ourselves.

In Section 2.3 we computationally investigate the effect of molecule size on the propensity for the $A + A \rightarrow \text{products}$ using the Hard Sphere Molecular Dynamics methodology. We considered a system of N hard disks, each of diameter l , initially distributed uniformly randomly with no overlap inside a circular container with hard reflective walls and diameter L .

In both one and two dimensions, the choice of hard instead of periodic boundaries was made after careful consideration. We believe that hard boundaries bring our simple system closer to being realistic. Periodic boundaries would introduce the unphysical appearance of molecules from nowhere, as they cross the boundary. Also, for molecules that have nonzero diameter, periodic boundaries make choices regarding initial random placement and inter-molecular collision detection awkward, if not arbitrary.

2.2 The Effect of Excluded Volume in One Dimension

In this section we try to answer the following question: What modifications to the mass action law are required to capture the kinetics of a bimolecular reaction in a well-stirred gas-phase system in which volume is excluded only by the reactant molecules? We should note that this scenario ignores several important features of cellular biochemistry, for instance, molecular kinetics in solution, and crowding by large non-reacting molecules. However, it has the advantage of being mathematically tractable, at least in the idealized case of a one dimensional, single species system. Therefore, our aim here will be to derive, from first principles, the propensity function for a reaction of the form $A + A \rightarrow \text{products}$ in a well-stirred,

one-dimensional system that contains only molecules of species, these molecules having a finite size.

It is known that for the reactions $A + A \rightarrow A$ and $A + A \rightarrow \emptyset$ in one dimension, neither ballistic (gas-phase) nor diffusive molecular motion can sustain a well-stirred condition over many successive reactions [2]. For such systems, maintaining the well-stirred condition would require either external stirring or a strongly coupled heat bath. However, our goal here is not to explore how the well-stirred condition might be maintained in one dimension; rather, it is to take the first step towards understanding the effects of reactant size in three dimensions, where the well-stirred condition can often be sustained over many successive reactions in both the ballistic and diffusive motion cases.

Since bimolecular reactions are always initiated by a collision between two reactant molecules, our main task here will be to compute the probability $p_{col}(dt)$ that a collision will occur, in the next infinitesimal time dt , in a system of N identical hard-rod molecules, each of length l , which are confined to the real interval $[0, L]$ with $L > Nl$. Our analysis will make two important assumptions: First, the positions of the rods are randomly uniform in $[0, L]$, subject to the constraint that no two rods overlap. Second, the system is at some constant absolute temperature T , so that the velocities of the molecules can be regarded

as independent normal (Gaussian) random variables with mean zero and variance

$$\sigma^2 \equiv k_B T / m \tag{2.1}$$

Here, k_B is Boltzmann's constant, and m is the mass of a molecule. Equation (2.1) can be seen most easily from the equipartition theorem, which says that in thermal equilibrium the average kinetic energy per molecule is $\langle \frac{1}{2} m v^2 \rangle = \frac{1}{2} k_B T$. The normality of the velocity distribution follows from the Maxwell-Boltzmann formula. Since each rod velocity v is an independent normal random variable with mean 0 and variance $2\sigma^2$, it follows that the difference in the velocities of any two rods, i.e. their relative velocity v_{rel} , is a normal random variable with mean 0 and variance σ^2 . It is not hard to show that the average of the absolute value of v_{rel} , i.e., the average relative speed of a randomly chosen pair of rods, is

$$s_{rel} \equiv \langle |v_{rel}| \rangle = \sqrt{\frac{4}{\pi}} \sigma \tag{2.2}$$

Note that, although the initial configuration of the system is chosen randomly, so as to mimic a well-stirred, thermally equilibrated state, the subsequent evolution of the system up until the first collision is entirely deterministic: the molecules move on the interval with their assigned velocities as though the system were isolated. However, in our subsequent analysis we do not continue the evolution of the

system beyond the first collision (if we did, we would notice the system “cooling” as fast-moving molecules are depleted, and correlations between the velocities of the molecules would eventually start to form). We make this choice because the focus of this work is not isolated chemical systems, but rather systems that are in thermal contact with their surroundings, as, for instance, a typical cellular system would be.

In the next section we review the standard derivation of $p_{col}(dt)$ in the dilute ($l = 0$) case. After that we give our derivation for the dense ($l > 0$) case, using the mathematical properties of a randomly uniform, non-overlapping distribution of rods in a fixed interval as derived in Appendix A. In section 2.2.3 we present some molecular dynamics simulation results that confirm our theoretical results, making use of the Hard Sphere Molecular Dynamics Algorithm. We conclude by discussing some possible implications of these results.

2.2.1 The Dilute Case

To compute $p_{col}(dt)$ in the dilute gas limit, where each hard-rod molecule is essentially a point, we begin by choosing from the system a pair of molecules at random, without regard for their positions. In an infinitesimal time dt , one of these chosen molecules will sweep out relative to the other an average distance $s_{rel}dt$. If the second molecule happens to lie inside the length swept out by the

first molecule, the two molecules will collide in the next dt . Since the molecules are distributed uniformly inside the interval $[0, L]$, the probability that they will collide in the next dt is simply the ratio of the size of the interval the second molecule needs to be in for a collision to occur to the size of the interval the second molecule is known to be in: s_{rel}/L . We invoke here the fact that the space occupied by the point molecules themselves is zero, or at least negligible compared to L . Since there are $N(N - 1)/2$ distinct ways in which the pair of molecules could have originally been chosen, it follows by the addition law of probability that the probability that some one of those pairs will collide in the next dt is

$$\frac{1}{2}N(N - 1) \times \left(\frac{s_{rel}dt}{L} \right)$$

Thus we conclude that, in the dilute limit,

$$p_{col}(dt) = \frac{N(N - 1)s_{rel}}{2L}dt \tag{2.3}$$

In the foregoing argument, the vanishing smallness of dt ensured that the probability for two or more pairs of molecules to collide in the next dt is negligibly small compared to the probability for only one pair to collide in dt . That in turn implied that pair collisions in time dt can effectively be regarded as mutually exclusive events, and this is what allowed us to invoke the addition law of probability

theory to obtain Eq. (2.3). Equation (2.3) is thus valid only for infinitesimally small dt .

Some may feel that the foregoing derivation is flawed because it appears to accord the same collision probability to all pairs of molecules, regardless of how far apart the members of the pair might be. But this criticism is unjustified. The single pair collision probability $s_{rel}dt/L$ is expressly conditioned on the premise that the relative locations of the two chosen molecules inside $[0, L]$ are not known – that is what was meant by stipulating that the two molecules be chosen “at random”. In a sense, $s_{rel}dt/L$ is the single-pair collision probability averaged over all possible relative positions of the pair members. As has been shown elsewhere [20], a derivation that explicitly computes the collision probabilities for adjacent point molecules, which are the only ones that can physically collide, yields the same result (2.3).

2.2.2 The Dense Case

If each of the N molecules in the system just considered had a non-zero extent l , how would the result (2.3) be altered? Naively, we might suppose that since a portion Nl of the interval $[0, L]$ is occupied by the rods themselves, then the probability that the center of one randomly chosen rod will lie inside the infinitesimal collision length $s_{rel}dt$ of the other randomly chosen rod would be the ratio of the

collision length to the free interval length $(L - Nl)$. This logic would result in simply replacing L in (2.3) with $(L - Nl)$. But further reflection might suggest that things should not be that simple: In a situation where the gap between two adjacent rods is less than l , that gap is not available for occupation by another rod, so perhaps it, too, should somehow be counted as part of the interval that is excluded by the rods themselves. We shall show in this section that this last concern is unfounded, and that the seemingly naive replacement of L in Eq. (2.3) with $(L - Nl)$ does in fact yield the correct generalization of Eq. (2.3) to the case $l > 0$.

We start by choosing at random a pair of adjacent rods, noting that only adjacent rods are capable of colliding with each other. Our selection of this adjacent pair is “random” in the sense that it is made irrespective of the positions and the velocities of the two rods that comprise the pair. All we know about these two rods is that the gap ξ between them is randomly distributed in a way that is consistent with a randomly uniform, non-overlapping distribution of N rods inside the interval $[0, L]$, and further that their relative velocity v_{rel} is randomly distributed according to the normal (Gaussian) density function $f_G(v_{rel})$ with mean 0 and variance $2\sigma^2$.

The probability density function $P_{\text{gap}}(\xi)$ for the inter-rod gap distance is derived in Appendix A, and the result is shown in Eq. A.10. Using that function,

we assert that the probability that our randomly chosen adjacent pair will collide in the next nearly infinitesimal time interval δt is, to first order in δt ,

$$p_{\text{col}}^{\text{s.p.}}(\delta t) = \int_{\xi=0}^{L-Nl} P_{\text{gap}}(\xi) d\xi \times \left(\int_{v_{\text{rel}}=-\infty}^{-\xi/\delta t} f_G(v_{\text{rel}}) dv_{\text{rel}} \right) + o(\delta t) \quad (2.4)$$

The superscript “s.p.” signifies that this is the collision probability for a single pair of adjacent rods, and we note that there are $N - 1$ such pairs in the system. The integral in parentheses in Eq. (2.4) is the probability that the relative velocity of the two rods will have a sufficiently large negative value that the rods will collide in the next δt if their present gap separation is ξ . That probability is multiplied by the probability that the gap separation will be infinitesimally close to ξ , and the result is then summed over all possible values of ξ to obtain the desired collision probability. The term $o(\delta t)$, which goes to zero with δt faster than δt but is otherwise unspecified, accounts for errors in this reasoning caused by interfering collisions of either rod with its other neighbor (or a boundary). The basic assumption being made here is that the probability for two or more collisions to occur in time δt can be made negligibly small compared to the probability of only one collision simply by taking δt close enough to zero. In the end, we will be interested only in the case in which δt is taken to be infinitesimally small.

Changing the limits in the v_{rel} -integral in Eq. (2.4) from $(-\infty, -\xi/\delta t)$ to $(\xi/\delta t, \infty)$, which is permissible because $f_G(v_{\text{rel}})$ is symmetric in gives

$$p_{\text{col}}^{\text{s.p.}}(\delta t) = \int_{\xi=0}^{L-Nl} d\xi \int_{v_{\text{rel}}=\xi/\delta t}^{\infty} dv_{\text{rel}} [P_{\text{gap}}(\xi) f(v_{\text{rel}})] + o(\delta t)$$

Next we change the order of integration over ξ and v_{rel} :

$$\begin{aligned} \int_{\xi=0}^{L-Nl} d\xi \int_{v_{\text{rel}}=\xi/\delta t}^{\infty} dv_{\text{rel}} &= \int_{v_{\text{rel}}=0}^{(L-Nl)/\delta t} dv_{\text{rel}} \int_{\xi=0}^{v_{\text{rel}}\delta t} d\xi + \int_{v_{\text{rel}}=(L-Nl)/\delta t}^{\infty} dv_{\text{rel}} \int_{\xi=0}^{L-Nl} d\xi \\ &= \int_{v_{\text{rel}}=0}^{\infty} dv_{\text{rel}} \int_{\xi=0}^{v_{\text{rel}}\delta t} d\xi + o(\delta t) \end{aligned}$$

The first step here can be seen by graphing the integration region in (ξ, v_{rel}) -space. The second step follows by taking δt so small that the limits $(L - Nl)/\delta t$ in the two v_{rel} -integrations can be effectively replaced by ∞ . Our assumption that doing this introduces errors of higher order in δt than the first is justified by the fact that the Gaussian function $f_G(v_{\text{rel}})$ falls off extremely rapidly with increasing v_{rel} , and in practice there will always be some finite upper bound on that variable. So now we have

$$p_{\text{col}}^{\text{s.p.}}(\delta t) = \int_0^{\infty} dv_{\text{rel}} \left(f_G(v_{\text{rel}}) \int_0^{v_{\text{rel}}\delta t} d\xi [P_{\text{gap}}(\xi)] \right) + o(\delta t)$$

Since $P_{\text{gap}}(0) > 0$, the inner integral over ξ can be easily evaluated to first order in δt :

$$\int_0^{v_{\text{rel}}\delta t} d\xi [P_{\text{gap}}(\xi)] = P_{\text{gap}}(0) (v_{\text{rel}}\delta t) + o(\delta t)$$

Therefore,

$$\begin{aligned} p_{\text{col}}^{\text{s.p.}}(\delta t) &= \int_0^\infty dv_{\text{rel}} (f_{\text{G}}(v_{\text{rel}}) P_{\text{gap}}(0) v_{\text{rel}}\delta t) + o(\delta t) \\ &= P_{\text{gap}}(0) \delta t \left(\frac{1}{2} \int_{-\infty}^\infty |v_{\text{rel}}| f_{\text{G}}(v_{\text{rel}}) dv_{\text{rel}} \right) + o(\delta t) \\ p_{\text{col}}^{\text{s.p.}}(\delta t) &= \left(\frac{N}{L - Nl} \right) \delta t \left(\frac{1}{2} s_{\text{rel}} \right) + o(\delta t) \quad (2.5) \end{aligned}$$

In the second step we have invoked the evenness of the function $f_{\text{G}}(v_{\text{rel}})$. In the last step we have invoked the value of the gap density function $P_{\text{gap}}(\xi)$ at $\xi = 0$, as prescribed by Eq. (A.10), and also the definition (2.2) of s_{rel} .

Recalling now that there were $N - 1$ possible ways in which this adjacent pair of rods could have been chosen, we apply the addition law of probability once more to infer that the probability that some one of those adjacent pairs will collide in time δt is

$$p_{\text{col}}(\delta t) = (N - 1) \times p_{\text{col}}^{\text{s.p.}}(\delta t) + o(\delta t) \quad (2.6)$$

Again, a term $o(\delta t)$ accounts for corrections to the addition law caused by the small possibility that more than one pair of rods might collide in time δt . Substituting

Eq. (2.5) into Eq. (2.6), and finally taking the limit $\delta t \rightarrow dt$ so that the $o(\delta t)$ term can finally be dropped, we obtain the principal result of this chapter:

$$p_{\text{col}}(dt) = \frac{N(N-1)s_{\text{rel}}}{2(L-Nl)}dt \quad (2.7)$$

We observe that putting $l = 0$ in Eq. (2.7) indeed gives the expected dilute gas result (2.3). More remarkable, though, is the implication that the dilute result (2.3) can be generalized to the dense case simply by replacing the interval length L with the free interval length, $L - Nl$. This goes against the supposition that, in estimating the available system volume, allowance must somehow be made for gaps between molecules that are smaller than a molecular diameter, since such gaps are not available for occupancy by a molecule. We suggest that the intuitive basis for that concern stems from trying to solve the related but different problem of uniformly scattering the rods inside the interval $[0, L]$, where the non-overlap condition would indeed preclude placing a rod between two already placed rods that are separated by a distance less than l . But, as is shown in Appendix A, the most efficient algorithm for randomly placing the rods inside $[0, L]$ in a uniform, non-overlapping way deals out the rods from left to right in that interval, and thus never tries to place a rod between two already placed rods.

The result (2.7) can be viewed from two different perspectives:

$$p_{\text{col}}(dt) = \begin{cases} \frac{1}{2}N(N-1) \times \frac{s_{\text{rel}}dt}{(L-Nl)} \\ (N-1) \times \frac{Ns_{\text{rel}}dt}{2(L-Nl)} \end{cases} \quad (2.8)$$

The first form in Eq. (2.8) is the probability that a randomly chosen pair of rods will collide in the next dt , multiplied by the number of such pairs. The second form is the probability that a randomly chosen adjacent pair of rods will collide in the next dt , multiplied by the number of such pairs. This dual perspective on $p_{\text{col}}(dt)$ was pointed out in Ref. [20] for the dilute case $l = 0$, with the aim of demonstrating that both of these lines of reasoning lead to the same correct formula for the dt -collision probability. Our present result therefore extends that earlier finding to the case $l > 0$.

If every collision between two A molecules produced a chemical reaction, then the propensity function for that reaction would simply be the coefficient of dt in Eq. (2.7). More generally, the propensity function should contain as a factor the probability $p_{\text{reac}|\text{coll}}$ that a colliding pair of molecules will actually undergo the reaction. For simplicity, however, in the following discussion we shall omit this factor and assume that every collision leads to a reaction.

2.2.3 Simulation Results

The logic used in discrete stochastic chemical kinetics to deduce the bimolecular propensity function – a logic which forms the foundation for both the chemical master equation and the stochastic simulation algorithm [19, 21, 22] – interprets the result (2.7) as implying that the time to the next collision/reaction is an exponential random variable whose mean is the reciprocal of the coefficient of dt :

$$\langle \tau_{next} \rangle = \frac{2(L - Nl)}{N(N - 1)s_{rel}} \quad (2.9)$$

But this result requires Eq. (2.7) to hold at the beginning of every infinitesimal time interval prior to the first collision¹, and hence that at all those times the system be well-stirred. Of course, the system is well-stirred, by assumption, at the beginning of the first infinitesimal interval $[0, dt)$. But can the same be said for the succeeding intervals $[dt, 2dt)$, $[2dt, 3dt)$, etc.? A simple heuristic argument giving an affirmative answer to this question goes as follows: So long as no collision has occurred before a given dt interval, we will not have learned anything at all about the positions of the rods beyond what we knew at time $t = 0$; thus, for any such dt interval we may regard the system as still being well-stirred. To put it in another

¹Suppose αdt is the probability of a collision in *any* infinitesimal time interval dt . The probability of no collision in any finite interval $[0, t)$ can be computed by subdividing $[0, t)$ into n subintervals of size t/n , and then computing the probability of no collision in all of those subintervals. If n is sufficiently large, the latter probability is $(1 - \alpha(t/n))^n = e^{-\alpha t}$, and this signifies an “exponentially distributed” collision time.

way, we are suggesting that the Maxwellian ballistic motion of the rods, at least up until the moment of the first collision, does not undo the randomly uniform spatial distribution of those rods.

To see if this expectation is true, and also to check the validity of formula (2.9), we have performed the following Hard Sphere Molecular Dynamics experiment:

- Using the generating formula A.14, deal out a configuration of the positions of the N rods in $[0, L]$. (This distributes the molecules randomly and uniformly in $[0, L]$.)
- Assign each rod a velocity in the form of an independent sample of the normal random variable with mean 0 and standard deviation $\sigma = \sqrt{\pi/4}s_{rel}$. (This puts the molecules in thermal equilibrium at temperature T [see Eqs. 2.1 and 2.2].)
- Allow the rods to move accordingly, with rods 1 and N reflecting elastically off the boundaries, until the first collision between a pair of rods occurs, and record the time τ of that collision.
- Repeat the above steps to obtain many samples of the time to the first collision/reaction, and then make normalized frequency histograms of those times.

We simulated ensembles for a range of values for N and the fraction of the interval occupied by the molecules themselves, κN , where κ is the dimensionless parameter defined by

$$\kappa \equiv \frac{l}{L} \tag{2.10}$$

We obtained for each case 20,000 samples of the time to the first collision/reaction. The estimated probability density function (pdf) of the time to the first collision is shown as the solid histograms in Figs. 2.1, 2.2 and 2.3. The dashed curves in these figures are the pdfs for the exponential distribution with the dense-case mean (2.9), and the dotted curves are the pdfs for the exponential distribution for the dilute-case mean with $l = 0$.

These results show that the distribution of t_{next} is, as conjectured, accurately described by an exponential distribution with the dense-case mean (2.9). For a very low fraction of volume occupied ($\kappa N = 1\%$), the difference between the three curves is practically impossible to detect. As the fraction of volume occupied by the molecules increases ($\kappa N = 50\%$), the difference between the dilute and dense exponential distributions becomes pronounced, and the simulation results in every case match the dense distributions. This happens consistently over reactant populations ranging from $N = 1000$ (Fig. 2.1) to $N = 4$ (Fig. 2.2). For $N = 3$ (not shown) and $N = 2$ (Fig. 2.3), a slightly larger difference between the analytical

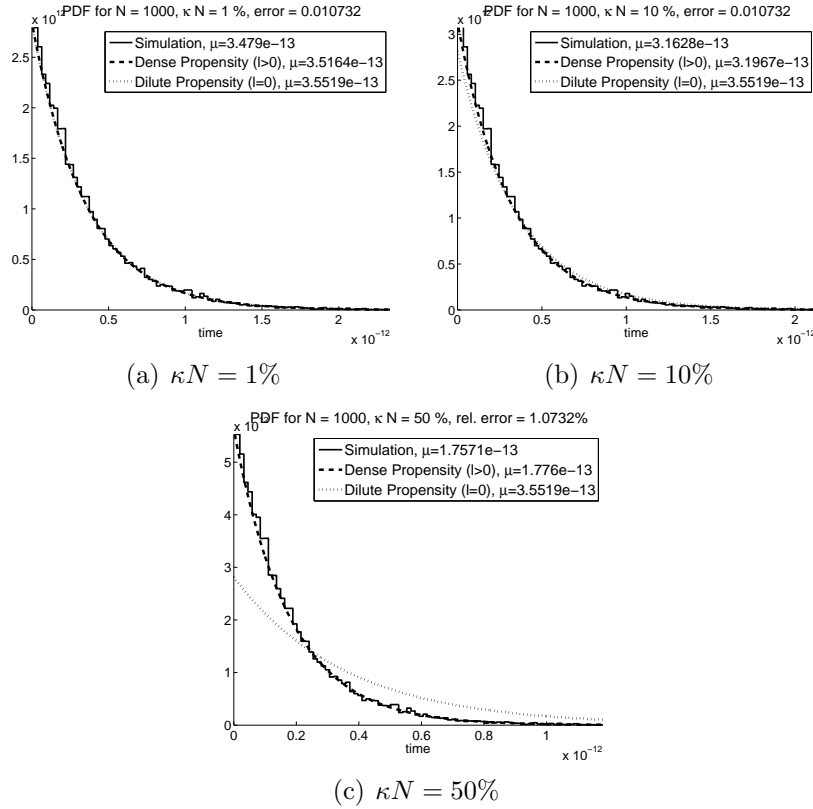


Figure 2.1: Histograms of simulation vs. analytical pdfds (dense and dilute exponentials) for population $N = 1000$ molecules. Here $\kappa = l/L$, where L is the “volume” of the system, and l is the “size” of one molecule, so κN is the fraction of the volume of the system occupied by the molecules themselves. Each simulation ensemble contains 20,000 samples. The histograms are computed using 140 equal-sized bins.

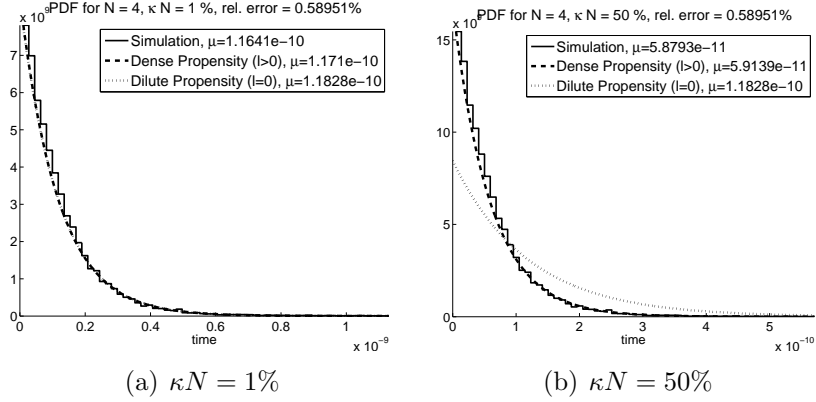


Figure 2.2: Histogram of simulation vs. analytical pdfs for $N = 4$. All other parameters are the same as for the previous figure.

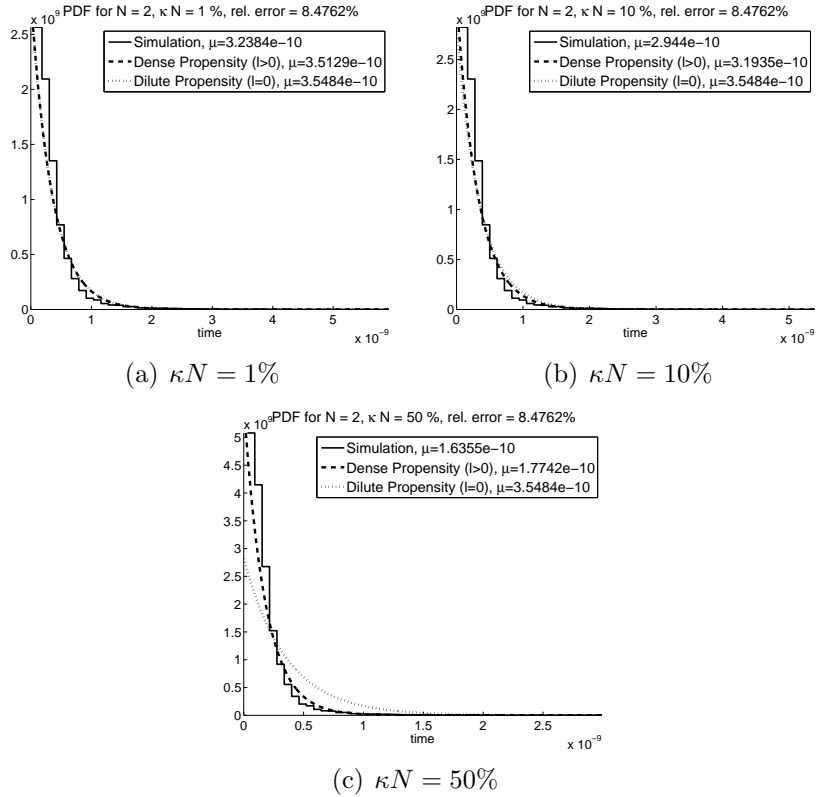


Figure 2.3: Histogram of simulation vs. analytical pdfs for $N = 2$. All other parameters are the same as for the previous figure.

and simulation curves is observed. We attribute this discrepancy to the fact that collisions with the hard boundaries are frequent enough as to be non-negligible when only two or three molecules are present in the system, and it vanishes already at $N = 4$. Nevertheless, even for $N = 2$ and $\kappa N = 50\%$, the dense exponential curve provides a reasonably good approximation to the simulation result.

2.2.4 Discussion

We have derived a formula for the propensity function of the well-stirred reaction $A + A \rightarrow \text{products}$, assuming that the N molecules of species are hard rods of length l that move ballistically with Maxwellian velocities in the one-dimensional interval $[0, L]$. Our formula generalizes the known $l = 0$ formula by replacing the system volume L with the “free volume” ($L - Nl$); perhaps surprisingly, it shows no effect from spaces between adjacent rods that are smaller than l . We also reported results of molecular dynamics experiments that confirm not only the accuracy of our derived formula, but also the exponentially distributed nature of the next-collision time that is implicit in both the Chemical Master Equation and the Stochastic Simulation Algorithm.

In the molecular dynamics experiments described in the previous section, the initial state of the system was chosen randomly – according to a uniform distribution of molecular positions and a Maxwell-Boltzmann distribution of molecular

velocities – but thereafter, until the instant of the first collision/reaction, the system evolved as though it were mechanically isolated. Most real systems of interest are not isolated. A cellular system, for instance, is typically immersed in some form of heat bath with which it interacts, in an essentially random way, to keep its temperature (but not its energy) constant. Since it is not clear how such heat bath interactions should be modeled (see for example Kopelman [24] for three alternatives), the following simple but admittedly artificial approach is implicit in our treatment here: During the time between successive collisions the system is treated as though it were isolated. The effects of the random interactions with the heat bath are then assimilated all at once at the moment of each collision, thereby establishing a new spatially homogeneous and thermally equilibrated initial state for the next collision. In the absence of a heat bath, the system would typically “cool” as a result of the removal of usually faster molecules by the reactions. But that thermally isolated case is not of interest to us here, and that is why we did not continue our molecular dynamics simulations beyond the time of the first collision.

It is worth noting the well-behaved nature of our propensity function implied by Eq. 2.7. As the size of the molecules goes to zero, our dense propensity function limits to the well-known point molecule result. Further, the divergence of our propensity function as the size of an A molecule becomes so large that N of

those molecules occupy the entire interval L seems physically quite reasonable, since in that limit the probability of a collision in the next dt should approach something close to 1. More generally, it seems reasonable to expect that, even under diffusing conditions, the N -dependence of the propensity function should change as the ratio of the size of an A molecule to the size of the system volume (the ratio κ in our analysis) changes.

Another interesting property of Eq. (2.7) is revealed by expanding that formula in a power series, using the dimensionless parameter defined in Eq. (2.10):

$$a(N) = \frac{s_{rel}}{L} \frac{\frac{1}{2}N(N-1)}{1 - \kappa N} \quad (2.11)$$

$$\frac{s_{rel}}{L} \frac{1}{2}N(N-1)(1 + \kappa N + \kappa^2 N^2 + \dots) \quad (2.12)$$

Form (2.11) shows how the usual “combinatorial” dependence of the propensity function on the number of molecules gets altered by molecular crowding. For a given system volume size L , as the size l of an A molecule goes from zero, where the propensity function behaves quadratically, to its maximum value L/N , where the propensity diverges, the propensity passes successively through stages in which its N -dependence can be roughly approximated by a power law, $a(N) \approx kN^X$, where $X = 2, 3, 4, \dots$. As κ is taken larger and larger, the dominant term in parentheses

on the right of Eq. (2.11) will be one with an increasingly higher exponent on N . An amusing instance of this power-law-like fit occurs for $l = 0.65$ and $L = 1000$: for this value of κ our propensity function mimics rather closely the $dN/dt \sim N^3$ behavior that is expected in a one-dimensional isolated, diffusion-limited system [28, 1, 26].

2.3 The Effect of Excluded Volume in Two Dimensions

We have previously shown how, for a one dimensional system, the mass action propensity functions need to be modified when the volume of the reactant molecules is comparable to the total system volume [23]. We analytically derived the following exact formula for the reaction probability, in the next infinitesimal time dt , of the reaction $A + A \rightarrow products$ in a one-dimensional system of N non-overlapping hard rods of length l moving ballistically in a volume of length L :

$$p_{\text{col}}(dt) = \frac{N(N-1)s_{\text{rel}}}{2(L-Nl)}dt \quad (2.13)$$

In the limit of $l \rightarrow 0$ this is equal to the usual dilute gas reaction rate law. The propensity function for the reaction is, by definition, this probability divided by

dt. The correctness of this formula was then confirmed through an extensive series of exact hard rod molecular dynamics simulations.

We do not see a way to extend this analytical result into two dimensions. Therefore, we use the hard spheres molecular dynamics simulation methodology to computationally investigate the effect of molecule size on the propensity for the $A + A \rightarrow \text{products}$ reaction in the *two dimensional* version of the system. We consider a system of N hard disks, each of diameter l , initially distributed uniformly randomly with no overlap inside a circular container with hard reflective walls and diameter L .

We find that the distribution of inter-collision times τ in this system is approximately exponential, but with noticeable deviations in certain circumstances. We study how the τ distribution varies with the parameters l and L , which, for a fixed N , determine the area density of the system (defined as the ratio of the area of the molecule disks to the total area of the system).

For small numbers of molecules, it appears that three types of τ distribution are present: at intermediate values of the area density, the distribution is indistinguishable from an exponential; as the system tends to the point molecule limit (low area density), long inter-molecular collision times are overrepresented; as the area density of the system becomes high, short inter-molecular collision times are overrepresented.

It is known that the choice of container shape affects the degree of ergodicity of the molecules' trajectories, with some container shapes encouraging trajectories that sample only small parts of the container's area. In the low population and small molecule size limit, we find that the small number of molecules, combined with a choice of non-ergodic container shape (e.g. circular, as opposed to “stadium”), gives rise to over-represented long times, while the τ distribution for short times is just as predicted theoretically. Either increasing the number of molecules, or improving the shape of the boundary, ameliorates this effect.

As the area density of the system becomes high (i.e. for large l/L), and for low population, we find that the excluded area inferred from the simulation measurements is larger than what one might expect from taking into account the area of the molecule disks, or their close-packed area, or even several less dense packings [35]. However, as the number of molecules is increased, the excluded area appears to approach the close-packed area, as one might have expected.

These results suggest that excluded area in two and three dimensional, finite, dense systems has a somewhat predictable impact for systems with a large number of molecules, but may have greater impact than one might have initially supposed for systems with low population.

In Section 2.3.1 we present the physical system under consideration and the computational algorithms used to simulate its kinetics. In Section 2.3.2 we give a

brief derivation of the probability of collision from first principles. In Section 2.3.3 we present the results of our simulations. Section 2.3.4 summarizes and explains our findings.

2.3.1 Hard Sphere Molecular Dynamics

The algorithm

The hard sphere molecular dynamics simulation algorithm is a simple billiard balls simulator. Given randomly uniform initial positions and random Maxwellian velocities for the molecules, it returns the time to the first inter-molecular collision.

The velocities of the molecules are Maxwell-Boltzmann distributed, meaning that each Cartesian component of the velocity vector is a statistically independent normal random variable with mean 0 and temperature-determined variance $\sigma^2 \equiv k_B T/m$, where T is the system temperature, m the mass of the molecules, and k_B is Boltzmann's constant.

Statistics for the distribution of times τ to the first collision are obtained by running an ensemble of 100,000 realizations of the hard spheres molecular dynamics algorithm (for a given N , l , and L), each with different random seeds.

Algorithm 1 Hard Sphere Molecular Dynamics

```
t = 0
Generate uniform random non-overlapping positions for N molecules
  (See Generation of Initial Positions)
for molecule i = 1, ..., N
  generate velocity distributed according to the Maxwell-Boltzmann distribution
endfor
loop until exit
  for molecule i = 1, ..., N
    compute the next putative collision time with the boundary
  endfor
  for each pair of molecules
    compute the next putative inter-molecular collision time (if it exists)
  endfor
  choose tnext, the smallest of all molecule-boundary and inter-molecular collision times
  t = t + tnext
  if the collision was inter-molecular
    report t and exit
  else
    reflect the molecule that collided with the wall
    advance the positions of the other molecules
  endif
endloop
```

Generating exact initial positions

For the one-dimensional problem it was possible to derive a rejection-free Monte Carlo algorithm for generating samples of the molecules' initial positions. An analogous procedure for two dimensions has proven elusive (this problem is equivalent to finding the analytical formula for the intermolecular distance distribution function), so a rejection based Monte Carlo algorithm was used to generate uniform random non-overlapping positions for the molecules in two dimensions.

Algorithm 2 Rejection-based generation of exact positions in two dimensions

```
for molecule index  $i = 1, \dots, N$ 
  Generate putative coordinates  $(x_i, y_i)$  for the center of the molecule,
  uniformly and randomly inside the container
  for molecule index  $j = 1, \dots, (i - 1)$ 
    if the placement of molecule  $i$  will overlap with molecule  $j$ 
      discard the entire set of placed molecules  $(1, \dots, i)$  and restart
    endif
  endfor
endfor
```

The only non-obvious step of the placement algorithm is that, if an overlap is discovered, we discard the entire set of already placed molecules, as opposed to only the very last one that produced the overlap. The latter procedure will result in a biased distribution, as can be demonstrated both analytically and by numerical simulation in a one-dimensional setting.

The computational cost of this rejection-based method increases with the total number of molecules, and with the area density of the system. The simulations

we were able to perform using this placement method reached up to area densities of 40% for small values of N (e.g. $N = 6$).

Generating approximate initial positions

An approximate alternative to generating the initial positions of the molecules by the rejection-based Monte Carlo method is to use a pre-stirring procedure. In this scheme we initialize the positions of the molecules on a regular grid and then allow the molecules to bounce around until their positions are randomized. The success of such an initialization depends on choosing a good stopping condition. This can be: a) a time by which an ensemble of positions is guaranteed to have become well-stirred, b) a target value for a function (e.g. the radial distribution function), which when reached would denote a well-stirred state, or c) noticing that a function (again, possibly the radial distribution function) appears to have converged, without explicitly having a target value for it. We were unable to locate either (a) or (b) in the literature, and our experience with checking the radial distribution function for convergence suggested that it was too noisy for systems with small numbers of molecules.

There is a subtle aspect to choosing the stopping criterion when the system is simulated not by traditional Molecular Dynamics (i.e. integration of the laws of motion), but by our exact hard sphere algorithm, which steps from collision to

collision. If we stop pre-stirring the system immediately after an intermolecular collision has occurred, the two molecules which have just collided will be touching. Thus, choosing a stopping criterion that is based on the idea that an exact number of collisions must occur before stopping will yield ensembles of initial positions which systematically contain two molecules touching. One must remember to remove such a bias by evolving the system for some time past the last collision, or avoid introducing such a bias in the first place, by choosing a stopping condition expressed as a duration of time rather than number of collisions. We do the latter.

In our experiments exploring different amounts of pre-stirring we found that unexpectedly long amounts of pre-stirring time were necessary for other indicators, e.g. the mean of τ , which is the focus of this chapter, to reach the values obtained by the rejection-based method. For a given system (i.e. given N, l, L) we quantify the duration of pre-stirring as $FN\tilde{\mu}_\tau$ units of time, where F is a scaling factor, N is the number of molecules, and $\tilde{\mu}_\tau$ is the expected time to the first collision for this system, under the assumption that $V_e = V_d = Na$ (see the next section). Figure 2.4 shows, for a system with $N = 9, L = 10, l = 2$, how the mean time to the first collision computed by simulations initialized with pre-stirring (circles) approaches that computed by a simulation initialized by rejection-based Monte Carlo (solid line with dashed confidence interval), with increasing F .

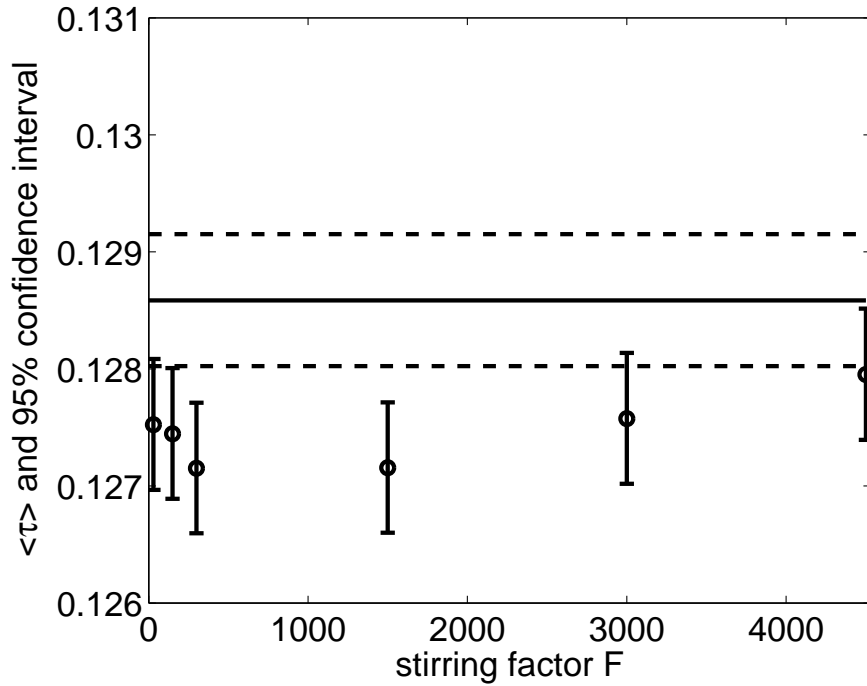


Figure 2.4: The mean of the τ -distribution, as calculated from simulations initialized with rejection-based Monte Carlo, is given by the solid lines. The dashed lines give the 95% confidence interval in the estimate of the mean. The circles and error bars give the means and 95% confidence intervals of the means of the τ -distribution, as calculated from simulations initialized with pre-stirring with F given by the respective x -coordinates. The system is $N = 9$, $L = 10$, $l = 2$.

For a factor $F = 300$, and the system of Figure 2.4 ($N = 9$), we see on average 4113 intermolecular collisions, i.e. each molecule experiences on average 457 collisions with another molecule. We also see on average 2401 collisions with the boundary, i.e. on average 267 collisions with the boundary per molecule. So F is a conservative estimate of the number of intermolecular collisions each molecule experiences during the pre-stirring simulation.

While the pre-stirring method of obtaining initial positions makes it possible to work with dense, high-population systems which we would otherwise not have been able to examine, it is nevertheless very time consuming. As can be seen from Figure 2.4, $F = 300$ is suboptimal in terms of achieving good stirring (if the mean of the τ -distribution is used as a convergence criterion). However it would take 10 times as long to simulate with $F = 3000$, in order to achieve the dubious improvement of having the 95% confidence interval of the pre-stirring method overlap that of the rejection-based Monte Carlo. Simulating ensembles of size 100,000 with a stirring factor of $F = 300$ already takes days on a workstation, so any approximate results we show are based on simulations with $F = 300$, rather than a higher value.

Due to the lack of rigor of our pre-stirring method, we do not consider it to be on equal footing with the rejection-based Monte Carlo. Nearly all our results were obtained using simulations initialized by rejection-based Monte Carlo. The

single exception is Figure 2.10, which would have been impossible to obtain by rejection-based Monte Carlo, and which we therefore consider speculative.

2.3.2 Theory for the two-dimensional case

Here we derive an expression for the probability $p_{col}(dt)$ of an inter-molecular collision occurring in the next infinitesimal time interval dt . This probability is given by the product of (the number of ways one can choose a random pair of molecules) \times (the probability that a randomly chosen pair of molecules will collide in the next dt). Because of the randomly uniform spatial distribution, the second factor can be further decomposed into the ratio: (the area one molecule will sweep out relative to the other molecule in the next dt) over (the total area inside the container accessible to the other molecule).

The initial velocities of the molecules are assumed to follow the equilibrium Maxwell-Boltzmann distribution, i.e. their Cartesian components are normal random variables with means 0 and variances $\sigma^2 \equiv k_B T/m$, where T is the absolute temperature of the system, m is the mass of the molecules and k_B is Boltzmann's constant. The mean relative speed of two randomly chosen molecules in such a distribution can be shown to be

$$s_{rel} = \langle |v_{rel}| \rangle = \sqrt{\pi} \sigma \tag{2.14}$$

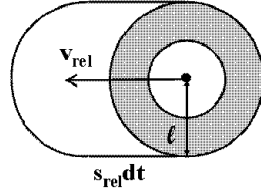


Figure 2.5: To calculate the probability of collision of two randomly chosen molecules of diameter l , moving with relative velocity v_{rel} , in the next infinitesimal time dt we compute the area that one molecule sweeps relative to the other in that amount of time: $2ls_{rel}dt$.

Suppose we randomly choose a pair of molecules. We can do this in $\frac{1}{2}N(N-1)$ ways. Now we change our frame of reference so that we are standing on one of the chosen molecules. Then the area swept out by the other molecule in the next infinitesimal time increment dt relative to the center of the one we are standing on is (see Fig. 2.5) $2ls_{rel}dt$. To get the probability of our randomly chosen pair of molecules colliding in the next dt , we must divide this area by the total area available to the molecules. If we assume that the molecules have no extent ($l = 0$), the area A available to the molecules is the total area of the system. If the shape of the container is circular, and the diameter is L , the area is given by $A_c = \frac{1}{4}\pi L^2$. For a “stadium” container (formally known as a Bunimovich stadium [27]), in which semi-circles of diameter L are separated by a square of side L , the area is $A_s = \frac{1}{4}\pi L^2 + L^2$.

If we relax the point molecule assumption (i.e. $l > 0$), the area available to the molecules will be less than the total area A of the system by at least the

area excluded by the disks of the molecules themselves ($V_d = Na = N\frac{1}{4}\pi l^2$). The question of exactly how large is the excluded area has prompted this research. We will call the excluded area as estimated from the simulation data the effective excluded area V_e , and we will show in the results section that it is quite a bit larger than V_d .

Finally, combining all of the above, we find that the probability of an inter-molecular collision in the next dt is given by:

$$p_{col}(dt) = \frac{N(N-1)}{2} \frac{2ls_{rel}}{A - V_e} dt \quad (2.15)$$

with $V_e = 0$ at the limit $l = 0$.

If we assume that the same probability $p_{col}(dt)$ holds for each successive infinitesimal time increment dt , from the initialization of the system until the first inter-molecular collision occurs, then the times τ will be samples of the exponential distribution whose mean is the inverse of the coefficient of dt in Eq. (2.15). If we further assume that every collision results in a reaction, that coefficient will be the propensity function of the Stochastic Simulation Algorithm and the Master Equation for that reaction.

2.3.3 Simulation Results

Collision time distribution methodology

To fully characterize the distributions of times to the first collision, τ , we asked the following questions: what is the empirical distribution of τ at different area densities, and what is the relationship between this empirical distribution and the analytical exponential in Equation (2.15) with the most conservative estimate of excluded volume ($V_e = V_d = Na$)?

Upon casual visual inspection, all the distributions certainly appear exponential. To quantitatively test for exponentiality, we used a common two-sample comparison test, the Kolmogorov-Smirnoff test [34]. To test a τ distribution (100,000 values) for exponentiality, we generated the same number of random samples from an exponential with the same mean as the τ distribution. We then used Matlab's `kstest` routine at the default significance level $\alpha = 0.05$. The routine rejects the null hypothesis that the samples came from the same distribution (exponential) if the so-called p-value returned by the test is < 0.05 ; it fails to reject the null hypothesis that the samples came from the same distribution, if the p-value is > 0.05 .

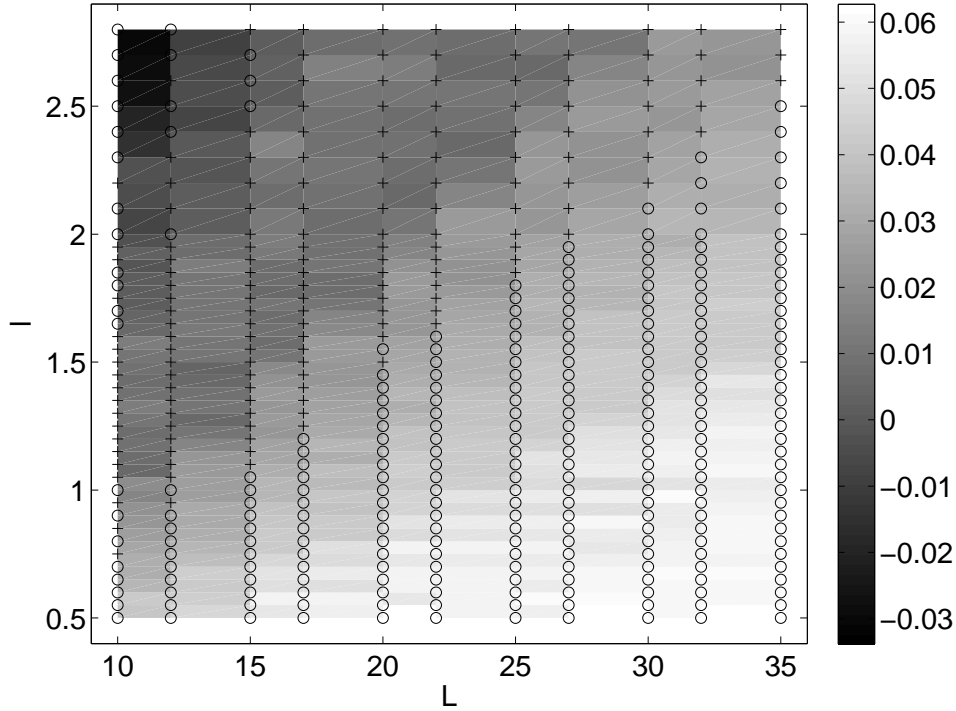


Figure 2.6: Results of K-S test for exponentiality of the τ -distribution at significance level $\alpha = 0.05$, superimposed on a grayscale plot of the quantity $I_\tau \equiv (\sigma_\tau)/(\mu_\tau) - 1$. The population is $N = 6$ for all samples; the standard deviation of the Cartesian velocity components is $\sigma = 1$ for all samples; the vertical axis is l , the diameter of one molecule, and the horizontal axis is L , the diameter of the system. "+" denotes a τ -distribution which is indistinguishable from exponential (p-value > 0.05), while "o" denotes a τ -distribution which deviates from the exponential (p-value < 0.05). Positive I_τ values imply a distribution with more spread than an exponential, while negative values imply a distribution with less spread than an exponential.

Collision time distribution for $N=6$

We began by considering the low molecule count case $N = 6$, with $\sigma = 1$ in a circular container of diameter L . Figure 2.6 shows the regions of (L, l) phase space where the distribution of τ is or is not statistically indistinguishable from an exponential according to the K-S test. At intermediate values of area density, for ensembles denoted with “+”, the distribution is indistinguishable from the exponential. At low and high area density values, for ensembles denoted with “o”, the distribution deviates from the exponential.

However, the p-value of the K-S test does not reveal *how* the empirical τ distribution differs from an exponential distribution. To gain some insight into this issue, we noted that the standard deviation of an exponential distribution is equal to its mean. We therefore computed the mean μ_τ and the standard deviation σ_τ of the empirical τ distribution, and then examined the quantity $I_\tau \equiv (\sigma_\tau/\mu_\tau) - 1$. If τ were exponentially distributed, we would have $I_\tau = 0$. If $I_\tau < 0$ the empirical τ distribution would be more peaked around its mean than an exponential distribution would be, and if $I_\tau > 0$ the empirical τ distribution would be more spread out than an exponential distribution.

The grayscale portion of Figure 2.6 gives a plot of I_τ , again for $N = 6$ and $\sigma = 1$, in the (L, l) -plane. The plot shows where the distribution of τ is more spread out or less spread out than an exponential. One possible interpretation

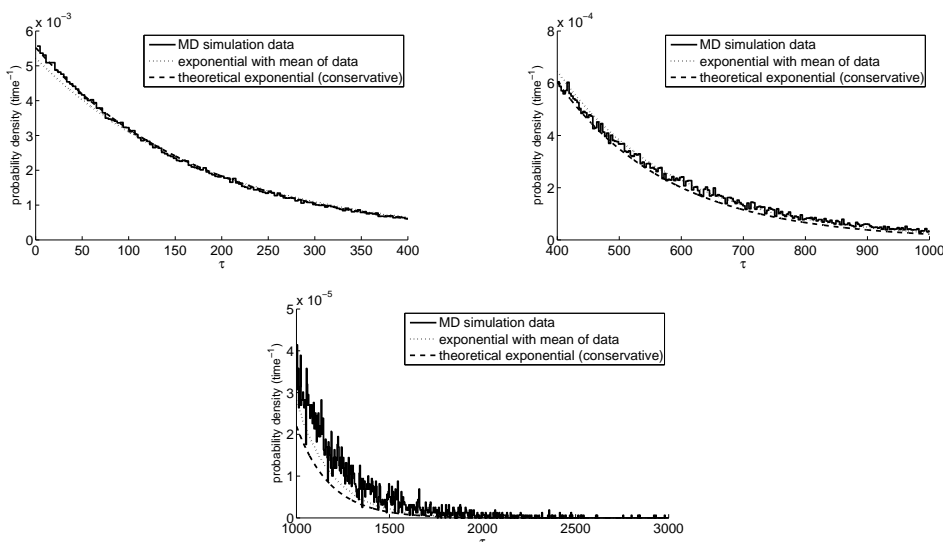


Figure 2.7: Piecewise histogram of the τ distribution for $\sigma = 1$, $N = 6$, $L = 35$, $l = 0.1$ vs analytical pdfs of two exponentials. The dotted exponential has the same mean as the τ distribution (191). The dashed exponential is the most conservative theoretical prediction (with $V_e = V_d = Na$, where a is the area of one molecule).

of this plot is the following: a large portion of the distribution of the τ sample is exponential, but superimposed on that is a component that ruins the overall exponentiality. In areas with $I_\tau < 0$ the values in the added component are close to the mean, i.e. small. In areas with $I_\tau > 0$ the values in the added component are larger than the mean. The latter component, in fact, contains some very large outliers.

This can be seen by comparing the hypothesized analytical exponential, whose mean is given by the inverse of the coefficient of dt in equation (2.15), to the histogram of the τ data. An instance of a system for which $I_\tau > 0$ is the *small*

molecule case $\sigma = 1$, $N = 6$, $L = 35$, $l = 0.1$. The mean for the data is 191 and the standard deviation is 204. Figure 2.7 breaks down the pdf of the τ distribution into three segments for clarity (note the different scales on the vertical axes). For shorter times (first plot), the empirical τ distribution (solid line) follows our analytical prediction (dashed line) very well. Approximately two standard deviations to the right of the mean (600, middle plot), the data becomes heavier than that model prediction, and is better described by an exponential with the empirical mean (dotted line). Finally, at long times (last plot), the data is heavier than both exponentials. So it appears that, in the $I_\tau > 0$ regime, the data follows the analytical exponential at shorter τ values, but has heavier than exponential tails for long τ values.

To better understand this behavior, we looked at some of the realizations that contributed the outlier τ samples, i.e. very long times. The trajectories of the molecules in those systems were notable for their non-ergodicity; more specifically, there would frequently be one or more molecules that moved along the circle boundary (a “whispering gallery” mode), or that crossed very close to the center (and therefore bounced back and forth along the diameter). These trajectories are non-ergodic in the sense that the molecules are not sampling the entire container volume. This has the effect of reducing the number of potential pairings of molecules, which is small to begin with, thereby reducing the probability that

some pair will collide. In this way, the shape of the container boundary combined with the small number of molecules contributes to the increased number of long-time outliers in the τ distribution.

One is also led to suspect the boundary's role from the observation that the number of boundary collisions preceding each inter-molecular collision increases dramatically as one approaches the point molecule limit – something that does not happen in a one-dimensional system. To test the hypothesis that the boundary contributes to the over-represented long times, we tried several modifications to the hard circular boundary that we initially studied.

One class of modifications left the shape of the boundary intact, but changed how the molecules were reflected after they struck it. Since our hypothesis is that specular reflections from the circular boundary tend to orchestrate the molecules' trajectories in such a way that they did not always sample the entire space, we modified the reflection formulas in several ways, in an attempt to increase the randomness in the trajectory of the molecules as they depart the boundary. First, we completely randomized the departing velocity of the molecule after a collision, while keeping the speed constant. Surprisingly, this had the effect of further lengthening the mean collision time. We next tried adding a small random angle to the departure angle after a reflection. This had the effect of very slightly shifting the mean collision time towards the model mean. Both of these were

ways that intuitively seemed to us as though they would increase the ergodicity of the trajectories in the system; yet, they yielded very inconclusive results.

The other class of modification that we tried retained specular reflection but changed the shape of the boundary to one which is supposed to discourage non-ergodic trajectories. More specifically, we moved from a circular boundary to a *Bunimovich stadium* [27]. This is simply the interior of a boundary made by joining two semi-circles of diameter L with a square of side L in between. This boundary shape is supposed to create fewer non-ergodic trajectories than either a square or a circle boundary alone. Indeed, we observed that it affected the τ distribution by moving its mean slightly, but noticeably, closer to the analytically predicted mean, and also the standard deviation closer to the mean (as it should be for an exponential). For instance when comparing two 400K run ensembles with $\sigma = 1$, $N=6$, $l=0.1$, one in a circular container of diameter $L=30$, $A=706.9$ and the other in a stadium with $L=19.9$, $A=707.0$, we found a mean τ of 140.9 for the circle and 138.9 for the Bunimovich stadium, along with a standard deviation of 149.5 for the circle and 145.9 for the stadium. The analytical mean and standard deviation of τ for a volume of this area are 132.9. We confirmed that the effect is general by noticing that the p-values for the K-S exponentiality test imply that the test is much closer to accepting the distributions as exponential for Bunimovich stadium volumes than for circular volumes of the same area.

It seems to us that whether a boundary’s shape and reflection characteristics contribute to non-ergodic trajectories is not a simple yes-or-no question, but rather one of degree. Some boundaries will cause molecules’ trajectories to sample the container’s area in a shorter amount of time than others. Our simulations would be mostly affected by the degree of ergodicity within a set finite amount of time (the expected time to a collision), not *eventually*. Thus, we suspect that there is no such thing as an “ideal” boundary that would completely eliminate the effects of non-ergodic trajectories. If this is so, then it is not surprising that the Bunimovich stadium did not completely abolish the long-time tails; instead, it is satisfying that it produced a measurable change in the expected direction.

At the other extreme of the area density, we see a different picture. An example system for which $I_\tau < 0$ is given by the *large* molecule case $\sigma = 1, N = 6, L = 10, l = 2.8$. The mean for the data is 0.1511 and the standard deviation is 0.1460. Figure 2.8 gives a breakdown of the pdf of the τ distribution for those parameters. The most noticeable feature of the pdf is the complete mismatch between the model exponential curve (dashed line) and the data (solid line). This is due to the fact that the model exponential (dashed line) is computed using the very conservative estimate that $V_e = Na$. But we see that the exponential curve with the empirical mean (dotted curve) follows the data (solid line) rather well. The major deviation in that regard is that at about two standard deviations to

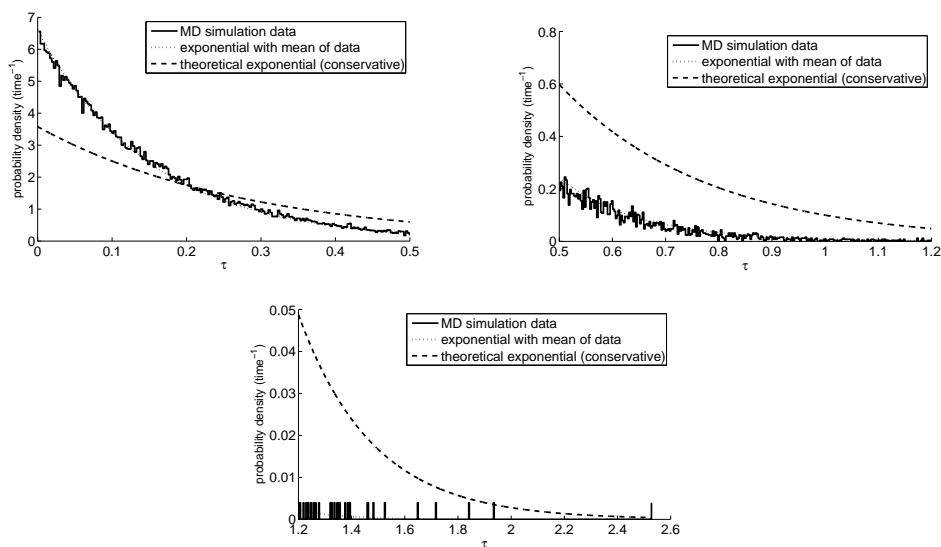


Figure 2.8: Piecewise histogram of the τ distribution for $\sigma = 1, N = 6, L = 10, l = 2.8$ vs analytical pdfs of two exponentials. The dotted exponential has the same mean as the τ distribution (0.1511). The dashed exponential is the most conservative theoretical prediction (with $V_e = V_d = Na$, where a is the area of one molecule).

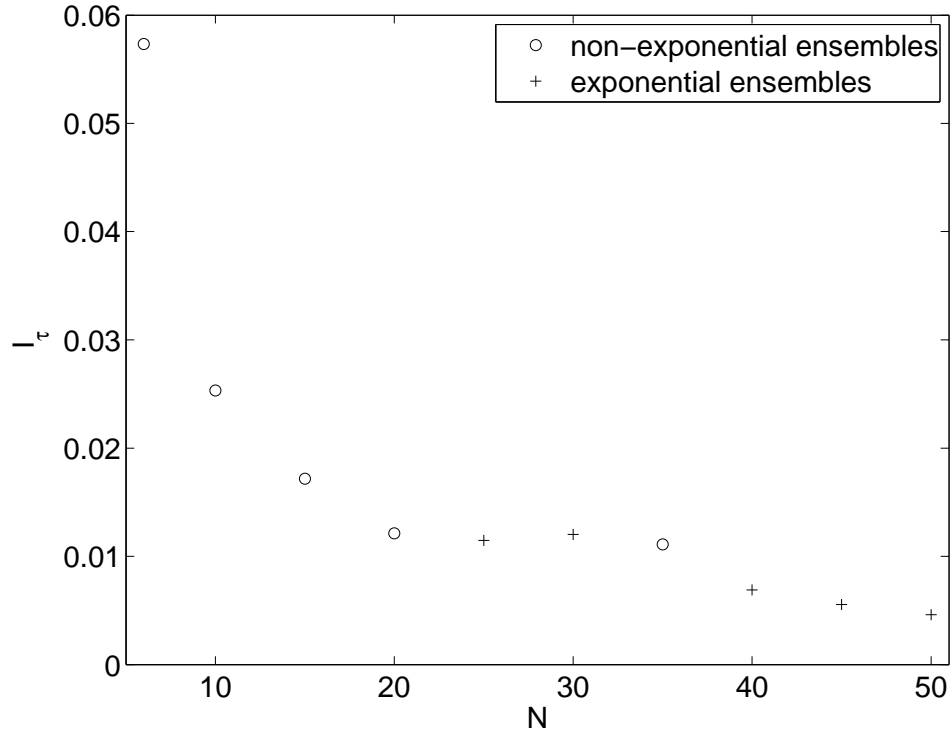


Figure 2.9: Plot of the indicator $I_\tau \equiv (\sigma_\tau/\mu_\tau) - 1$ for fixed low area density 0.5%, $L=10$, and varying N . ”+” denote τ -distributions which are indistinguishable from exponential according to the K-S test at significance $\alpha = 0.05$, while ”o” denote τ -distributions which are non-exponential.

the right of the mean (~ 0.5), the data slightly undershoots the dotted exponential curve, which has the effect of biasing the mass of the distribution closer to the mean.

Collision time distribution for large population

We have shown that in the low molecule population and low area density case the small number of molecules conspires with the non-ergodic boundary to

introduce long-time outliers in the τ -distribution. Figure 2.9 shows that increasing the number of molecules while maintaining the low area density abolishes this effect, restoring exponentiality to the distribution.

But is it also the case that the low population but high area density non-exponentiality, which we described in the previous section, can be abolished by increasing the number of molecules? A definitive answer to this question could be given if our exact simulation methodology were tractable on dense, high N systems. However, the rejection-based Monte Carlo initialization of the positions of the molecules takes prohibitively long to complete for such systems. A tentative answer, which is hopefully a hint in the right direction, can be obtained using the pre-stirring based molecule position initialization routine. Figure 2.10(a) shows that increasing the number of molecules while maintaining a high area density restores exponentiality to the τ -distribution. (We will discuss Figures 2.10(b) and 2.10(c) in the next section.)

Excluded volume

Our initial goal in this effort was to investigate the effect of reactant-excluded volume on the kinetics of the $A + A \rightarrow \text{products}$ reaction in our simulation experiments. The most straightforward way to estimate the excluded volume felt by the molecules in these experiments is to proceed as follows: Assume the distribution

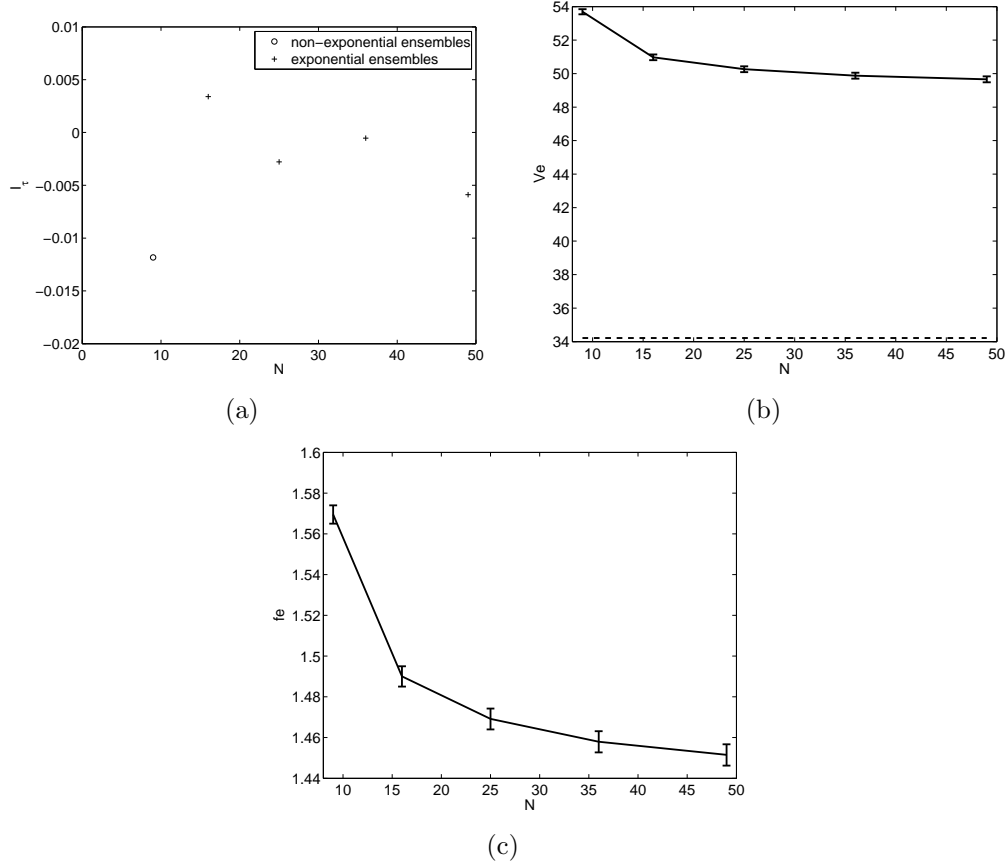


Figure 2.10: (a) Plot of the exponentiality according to the K-S test, and the indicator $I_\tau \equiv (\sigma_\tau/\mu_\tau) - 1$ for fixed high area density 42%, $L=10$, and varying N . ”+” denote τ -distributions which are indistinguishable from exponential, according to the K-S test at significance 0.05, while ”o” denote τ -distributions which are not exponential by the K-S test. The y-coordinate of the ”+” and ”o” gives the I_τ value for each distribution. Plot of the effective excluded volume V_e (b), and inverse packing fraction f_e (c), as a function of N , for the same fixed high area density 42%, and $L=10$. The dashed line in (b) gives $V_d = Na$.

of τ to be exponential; then estimate the propensity P as the inverse of the mean μ_τ of the empirical τ -distribution; finally, compute the effective excluded volume by solving the following equation for V_e :

$$\frac{1}{\mu_\tau} = P = \frac{N(N-1)}{2} \frac{2ls_{rel}}{A - V_e} \quad (2.16)$$

The effective excluded volume computed this way can then be compared to several theoretically plausible excluded volume formulas. The simplest such formula takes into account only the area excluded by the disk molecules themselves: $V_d = Na$.

Another possibility would be to consider volumes of the form $V_i = Na/\varphi_i$, where φ_i is a packing fraction. Some packing fractions that have been theoretically studied by others are: $\varphi_{cp} = \pi/2\sqrt{3} \approx 0.90$, the “close packing” fraction; $\varphi_f = 0.69$, the “freezing” packing fraction; and $\varphi_c = 0.82$, the “random close” packing fraction [35]. It should be noted that in one dimension, all three of these packing fractions are equal to 1; therefore, if we were to find that in two dimensions any one of these fractions is the desired factor, the theory would limit nicely to the lower dimensional result.

Since all the proposed excluded volumes above are of the form $V = fNa$, with f the inverse of a packing fraction, a reasonable quantity to visualize is the

effective inverse packing fraction, $f_e = V_e/Na$. Higher f_e is associated with looser packings, i.e. higher per-molecule excluded volume. The inverse “close packing” fraction is $f_{cp} \approx 1/0.90 \approx 1.1$; the inverse “freezing” packing fraction is $f_f \approx 1.45$; and the inverse “random close” packing fraction is $f_c \approx 1.22$.

The three parameters of the simulation are N , the number of molecules, L , the diameter of the circular container, and l , the diameter of the molecules. In each of the plots in Figure 2.11, we keep two of these parameters fixed and vary the other. In the left plots we show V_e (solid line), as estimated from solving Eq. (2.16), along with $V_d = Na$ for comparison (dashed line). In the right plots we show f_e . Note that f_e is the ratio of the solid and dashed lines from the left plots.

The first two left plots (in which we vary N and l) show that the effective excluded volume is higher than just the disk volume, as expected. But does the effective excluded volume correspond to some packing fraction? The first two plots on the right address just that question: it seems that no constant packing fraction can account for the effective excluded volume we observe, across the whole range of area densities and populations. The effective inverse packing fraction, f_e , decreases with increasing area density and with increasing population. This is also the case in our speculative results based on the pre-stirring initialization of positions (see Figures 2.10(b) and 2.10(c)).

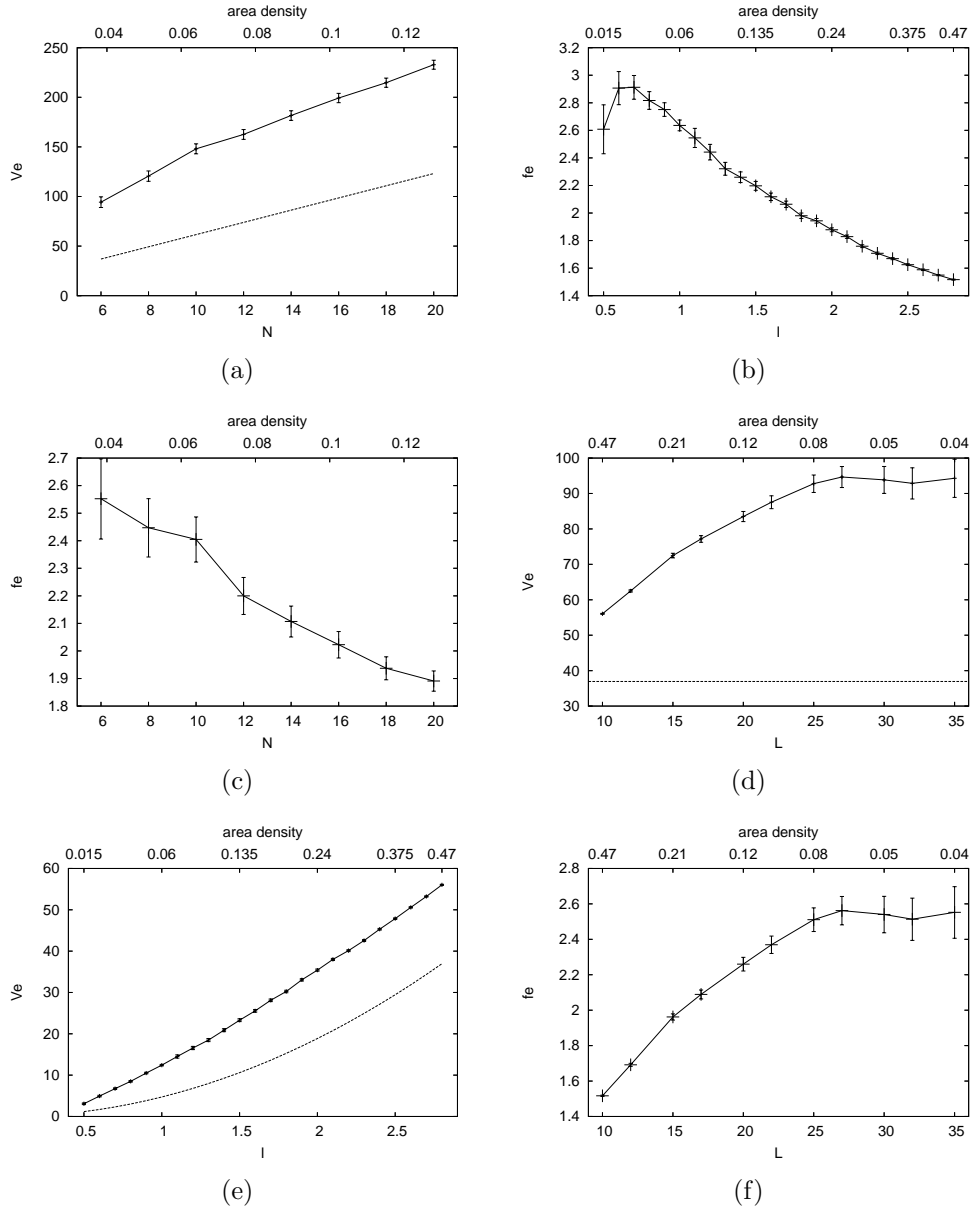


Figure 2.11: Plots of the effective excluded volume V_e (left) and inverse packing fraction f_e (right), as a function of the three parameters L , l , and N , keeping two parameters fixed, and varying the other. For all plots we have $\sigma = 1$, ensembles of size 100,000, and 95% confidence intervals. The dashed lines give $V_d = Na$. In the top plots we vary N , the number of disks, while maintaining $L = 35$, $l = 2.8$; in the middle plots we vary l , the diameter of the molecule disks, while maintaining $N = 6$, $L = 10$; in the bottom plots we vary L , the diameter of the container circle, while maintaining $N = 6$, $l = 2.8$.

This implies that the excluded volume situation for finite, reflective boundary containers is not as simple in two dimensions as it is in one dimension. Since the effective excluded volume is a function of the mean μ_τ of the empirical τ -distribution (Eq. 2.16), the reason why the packing appears more compact as the number of molecules increases must relate back to the τ -distribution. But we have already shown that in situations with low population, the τ -distribution contains artifacts introduced by the reflective boundary, which ruin its exponentiality, and which obviously impact its mean. So it is reasonable that in situations where the τ -distribution is not exponential, the packing estimated from the empirical is surprising, in this case for its looseness.

At high population and intermediate density, e.g. $N = 20$ and area density 12%, at the right end of Figure 2.11(c), the packing reaches $f_e = 1.8$. At high population and high area density (but simulated with pre-stirring), e.g. $N=50$ and area density 42%, at the right end of Figure 2.10(c), the packing reaches $f_e = 1.45$, the freezing packing fraction. So, it appears that for densities at which the excluded area is a significant portion of the area of the system, at high population, the effective excluded area is reasonably close to close packed.

2.3.4 Discussion

We have used computer simulations of hard disk dynamics to study the effect of reactant size on the rates of intermolecular collisions in the ballistic setting. We have found that the distribution of collision rates is close to exponential. It can be thought of as mostly exponential, except at low population, where we observe an additional mode which depends on the area density of the system.

At low population and low area density, non-ergodic trajectories contribute to longer than expected intermolecular collision times. At low population and high area density shorter intermolecular collision times are over-represented. We conclude, on the basis of exact simulations, that increasing the population abolishes the low area density effect; on the basis of approximate simulations, we speculate that the same is true at high area density.

At intermediate and high area densities, the volume excluded by the reactants ranges from about 2.5 times the area of the molecules' disks, at low population or low area density, to near the close packed volume, at high population or high area density.

Chapter 3

Relaxing the Homogeneity Assumption

3.1 Introduction

The idea of treating a spatially inhomogeneous chemically reacting system as a collection of smaller interacting subsystems is not new. It has appeared in the literature since the 1970s under a few names, notably the “reaction-diffusion master equation” [16] and the “multivariate master equation” [29]. The theory behind it has been explored by Nicolis and Prigogine in numerous publications (e.g. [29, 4]). A more recent paper [4] gives a comparison of the reaction-diffusion master equation approach with results from Direct Simulation Monte Carlo [7] in dilute gases, and attempts to experimentally verify the range of validity of the method. Widespread awareness of the multivariate master equation approach was

probably achieved through its inclusion in the classic texts of Gardiner [15] and Van Kampen [25].

This chapter is divided into two parts: in Section 3.2 we present the ISSA, while Section 3.3 presents the Multinomial Simulation Algorithm, which is our major contribution to the field of Inhomogeneous Stochastic Simulation.

3.2 The Inhomogeneous SSA

The inhomogeneous stochastic simulation problem is stated as:

Given \mathbf{X} a $Q \times N$ matrix representing the population of N species $\{S_1, \dots, S_N\}$ distributed among Q subvolumes $\{V_1, \dots, V_Q\}$ participating in M chemical reactions $\{R_1, \dots, R_M\}$, if at time t_0 the population matrix is \mathbf{x}_0 , find \mathbf{x}_t at any later time t .

We assume that the species are distributed among equally sized cubic subvolumes of side length l (and volume l^3). Each subvolume is treated as a homogeneous system, and the SSA applies to each subvolume independently. We also allow for the transfer of molecules between subvolumes which are nearest neighbors. These transfers can be thought of as mathematically equivalent to unimolecular reactions in which a species in the origin subvolume is a reactant, and therefore is consumed, and the same species in the destination subvolume is a product.

The stoichiometric matrix is the same for the inhomogeneous formulation as it is for the homogeneous. The update formula for reactions is also the same (1.4),

except for the fact that it is applied to one subvolume at a time. Specifically, if the state at time t is \mathbf{x}_t , and an R_k reaction happens in subvolume V_l in the time interval $[t, t + \Delta t)$, then the state $\mathbf{x}_{t+\Delta t}$ at time $t + \Delta t$ will be the same as \mathbf{x}_t for all rows except row l , which will be:

$$\mathbf{x}_{t+\Delta t}(l) = \mathbf{x}_t(l) + \boldsymbol{\nu}_k \quad (3.1)$$

A diffusive transfer of S_k from subvolume V_i to subvolume V_j has an update vector of the form:

$$(0, \dots, 0, -1, 0, \dots, 0, +1, 0, \dots, 0) \quad (3.2)$$

where the -1 occurs in the i -th position and the 1 occurs in the j -th position.

These updates are applied to the columns of the state matrix \mathbf{X} . Thus, if at time t the state is \mathbf{x}_t and in the time interval $[t, t + \Delta t)$ a diffusive transfer of S_k happens from subvolume V_i to subvolume V_j , the update formula for the k -th column of the state matrix looks like:

$$\mathbf{x}_{t+\Delta t}(k) = \mathbf{x}_t(k) + (0, \dots, 0, -1, 0, \dots, 0, +1, 0, \dots, 0) \quad (3.3)$$

It is possible to try to collect the diffusive transfer update vectors in a matrix, but there is neither a natural ordering for the matrix, nor any other good reason

to do so. Instead, we will just describe the procedure and structures necessary to realize diffusive transfers.

The most general solution to the problem of keeping track of nearest neighbors is a *nearest-neighbor matrix*. This matrix can take several different forms. Following Mathematica's nomenclature, it can be either an *adjacency matrix* or an *adjacency list* (which for a regular lattice is, in fact, a matrix).

An adjacency matrix for Q subvolumes is simply a $Q \times Q$ matrix with a 1 at position (i, j) if V_i and V_j are nearest neighbors, and a 0 if they are not.

An adjacency list is a collection of Q lists; the i -th list contains the indices of all the nearest-neighbors of V_i , in no particular order. For a cubic lattice, where there are 6 neighbors for interior subvolumes and fewer for boundary subvolumes, it makes sense to set up an adjacency list as a $Q \times 6$ matrix, with -1 (or some other special value) representing empty fields for boundary subvolumes. If the system is two dimensional the adjacency list need only be $Q \times 4$, and for a one dimensional system it need only be $Q \times 2$. Besides consuming less memory, an adjacency list turns out to have advantages for the kind of access we require, so from now on we will restrict ourselves to this implementation of the nearest-neighbor matrix. We call the nearest-neighbor matrix C .

Armed with the nearest-neighbor matrix, which specifies the spatial topology, and consequently the possible diffusive transfers, and with the stoichiometric matrix ν , and given the initial state \mathbf{x}_0 we have a full description of the “system”.

A master equation for the coupled diffusion and reaction process will have components that will conceptually look like:

$$\frac{\partial}{\partial t}P(\mathbf{x}_t, t|\mathbf{x}_0, t_0) = \frac{\partial}{\partial t}P_{\text{diffusion}} + \frac{\partial}{\partial t}P_{\text{reaction}} \quad (3.4)$$

So we set out to find a master equation for the diffusion process.

To describe the diffusive transfers we define the *diffusive transfer propensity function* to be $d_{i,v,v'}(\mathbf{x})$ such that

$$\begin{aligned} d_{i,v,v'}(\mathbf{x})dt &= \text{the probability that one molecule of species } S_i \\ &\text{will diffuse from volume } v \text{ to volume } v' \\ &\text{in } [t, t + dt) \text{ given } \mathbf{X}(t) = \mathbf{x}. \end{aligned} \quad (3.5)$$

At this point we are forced to become a little sloppy with our notation due to the fact that diffusion updates columns of the state matrix and reaction updates rows. To solve the problem of these coexisting in the master equation we are building up, we will use superscripts for row indexing and subscripts for column indexing of the state.

We will use the indexing \mathbf{x}_i to mean the i -th column of the state, corresponding to the population of S_i in all subvolumes. Then the master equation for the diffusion takes the form:

$$\frac{\partial}{\partial t} P(\mathbf{x}_t, t | \mathbf{x}_0, t_0) = \sum_{\text{nn}} \sum_{i=1}^N d_{vv'} [(\mathbf{x}_i - \text{change}) P(\mathbf{x}_i - \text{change}, t) - \mathbf{x}_i P(\mathbf{x}_i, t)] \quad (3.6)$$

where “nn” means the sum over nearest neighbor pairs vv' and “change” is of the form (3.2). Note that both probabilities on the right should be conditioned on \mathbf{x}_0, t_0 , but we are omitting the conditioning to eliminate clutter.

The CME (1.7) will similarly need to be adapted to use over multiple subvolumes. In the case of the CME however we need to index the state matrix by row, meaning by subvolume. We mandate that the indexing \mathbf{x}^v will mean the v -th row of the state, corresponding to the local population of all species in subvolume V_v .

The CME for multiple subvolumes becomes:

$$\begin{aligned} \frac{\partial}{\partial t} P(\mathbf{x}_t, t | \mathbf{x}_0, t_0) = & \sum_{v=1}^Q \sum_{j=1}^M [a_j(\mathbf{x}^v_t - \nu_j) P(\mathbf{x}^v_t - \nu_j, t | \mathbf{x}^v_0, t_0) \\ & - a_j(\mathbf{x}^v_t) P(\mathbf{x}^v_t, t | \mathbf{x}^v_0, t_0)], \end{aligned} \quad (3.7)$$

As (3.4) clearly states, the *Multivariate Master Equation* is the sum of (3.6) and (3.7), which we will not write again.

The diffusive transfer propensity function $d_{i,v,v'}(\mathbf{x})$, for equal-size cubic subvolumes and time-constant, isotropic diffusion, is:

$$d_{i,v,v'}(\mathbf{x}) = \frac{D_i}{l^2} \mathbf{x}^v \quad (3.8)$$

where D_i is the diffusion coefficient of species S_i at the correct temperature, and l is the length of one side of our cubic subvolumes [15]. This choice of $d_{i,v,v'}(\mathbf{x})$ allows this description of diffusion to lead to the deterministic diffusion equation ($\frac{\partial c}{\partial t} = D_i \nabla^2 c$, where c is concentration) in the limit of $l \rightarrow 0$.

The next reaction density function, which is sampled in order to obtain a trajectory of the homogeneous system, can be generalized to a *next event density function*. This function would be the joint distribution of time for the next event and type of the next event. The *type of the next event* can be reaction R_j in subvolume V_v or diffusion of species S_i between V_v and $V_{v'}$.

We note that reaction propensities now need to be labeled by subvolume, in addition to reaction. We institute the notation a_{jv} to denote the propensity of reaction R_j in subvolume V_v . The diffusion propensities are already labeled appropriately, so there is nothing to be done about them. Next we redefine a_0 to be the sum of all propensities over all subvolumes and all reaction channels: $a_0 = \sum_{j=1}^M \sum_{v=1}^Q a_{jv}$. We similarly define d_0 to be the sum of all diffu-

sive transfer propensities over all species and for all pairs of nearest neighbors:

$$d_0 = \sum_{i=1}^N \sum_{\text{nn}} d_{iiv'}$$

Now the formulas required for generating the waiting time τ analogously to (1.11) is:

$$\tau = \frac{1}{(a_0(\mathbf{x}_t) + d_0(\mathbf{x}_t))} \ln \left(\frac{1}{r_1} \right) \quad (3.9)$$

To generate the event type we proceed as in (1.11). First we decide on an ordering of all possible events. Then we start summing their propensities one-by-one according to the predetermined order, until the sum first exceeds r_2 . The event realized is the last event whose propensity we added to the sum.

3.3 The Multinomial Method

3.3.1 Motivation

The ISSA algorithm is prohibitively slow for some systems, and in particular in the presence of fast diffusion.

In this Section we present the Multinomial Simulation Algorithm (MSA), which is intended to be faster than the ISSA in just this type of scenario: when diffusive transfers greatly outnumber reaction events. The MSA is a stochastic-stochastic hybrid method which is based on separating chemical reactions, which

are treated in the usual SSA way, from diffusive transfers, which are treated by an approximate stochastic process. The MSA computes the *net* diffusive transfer from each subvolume to its neighbors in a given time-step. In this sense it is similar to the τ -leaping method, but with some important differences. The MSA has the important property that it conserves the total number of molecules across subvolumes, which is not the case, for example, for τ -leaping.

A number of authors [6, 32, 14] have proposed deterministic-stochastic hybrid methods in which diffusion is treated deterministically. These methods are applicable when the diffusing species are present everywhere in large population, but often this is not the case. The MSA is capable of obtaining spatial resolution even in the low population case.

The MSA is different from the Gillespie Multi-Particle (GMP) method of Rodriguez *et al.* [31], another stochastic-stochastic hybrid method, in two ways. First, although the MSA also relies on a type of operator splitting to separate reactions and diffusive transfers, it interleaves reactions and diffusions differently from the GMP method. We feel that our way is better justified theoretically, and possibly more accurate. Second, the GMP method uses Chopard’s multi-particle method [10] to simulate diffusion. According to this method, molecules from one subvolume are uniformly randomly distributed among the immediately neighboring subvolumes at each diffusion step, and the macroscopic diffusion equation is

recovered in the limit of $\lambda \rightarrow 0$, where λ is the subvolume's side length. In the MSA, molecules from one subvolume are also distributed among the neighboring subvolumes, but the probabilities used for that are multinomial. Rather than being chosen to limit to the diffusion equation as $\lambda \rightarrow 0$, the multinomial way of treating diffusion is intended to limit to the ISSA as the diffusion time-step goes to zero.

Rossinelli *et al.* [32] presented two methods: the $S\tau$ -Leaping is a stochastic algorithm which employs a unified step for both the reaction and diffusion processes, while the hybrid $H\tau$ -Leaping method combines deterministic diffusion with τ leaping for reactions. As in homogeneous τ -Leaping [8], the difficulty in spatial τ -Leaping is choosing a time-step that simultaneously satisfies the leap condition, i.e. that the propensities do not change substantially during the leap (an accuracy condition), but also has a low likelihood of causing the population to become negative. The choice of diffusion time-step for the MSA is also limited by an accuracy condition, but the way in which the jump probabilities are conditioned eliminates the problem of negative population.

Finally, the same stochastic process theory which forms the early steps of the derivation of the MSA appears in a non-spatial context in Rathinam and El Samad's paper on the Reversible-equivalent-monomolecular τ (REMM- τ) method [30]. REMM- τ is an explicit τ -leaping method, which approximates bimolecular

reversible reactions by suitable unimolecular reversible reactions, and considers them as operating in isolation during the time-step τ . The MSA and REMM- τ apply to distinctly different physical systems, but they share a common mathematical foundation, namely an exact, time-dependent stochastic solution for the reversible isomerization reaction set $S_1 \rightleftharpoons S_2$. In the present work we generalize that solution to the reaction set $S_1 \rightleftharpoons S_2 \rightleftharpoons \dots \rightleftharpoons S_n$, for $n > 2$, and we also develop approximations to make the calculations practical. The $n = 2$ solution expresses the instantaneous populations of the species as linear combinations of statistically independent *binomial* random variables. Our $n > 2$ generalization takes the form of linear combinations of statistically independent *multinomial* random variables – hence the name of the MSA.

follows: In Subection 3.3.2 we develop multinomial diffusion for one species in one dimension in the absence of any reactions. In Subection 3.3.3 we extend this to an arbitrary number of species, and add reactions to obtain the Multinomial Simulation Algorithm; we then present simulation results and evaluate the algorithm's performance in one dimension. In Subection 3.3.4 we describe the algorithm for two dimensions and present some simulation results. We conclude with a discussion of how the algorithm can be used as part of a larger adaptive simulation strategy.

3.3.2 Diffusion in one dimension

Theoretical foundations

In this subsection we derive the foundations of the Multinomial Simulation Algorithm. For simplicity we do this for a one dimensional system.

Suppose we have a one dimensional system of length L which contains only one chemical species. Consider n subvolumes of equal size, $l = L/n$, which we index from left to right $1, 2, \dots, n$. Initially, subvolume i contains k_i molecules of a given chemical species, distributed randomly and uniformly. Now suppose that κ is defined as follows:

$$\kappa dt \equiv \text{the probability that a molecule will jump to an adjacent cell} \quad (3.10)$$

in the next infinitesimal dt .

This parameter is usually taken to be $\kappa = D/l^2$, where D is the usual diffusion coefficient of the chemical species [11].

Define the probabilities

$$p_{ij}^{(n)}(t) \equiv \text{the probability that a randomly chosen molecule in cell } i \text{ at time } 0 \text{ will be} \\ \text{found in cell } j \text{ at time } t > 0, (i, j = 1, 2, \dots, n). \quad (3.11)$$

Since these n^2 probabilities satisfy the n relations:

$$p_{i,1}^{(n)} + p_{i,2}^{(n)} + \dots + p_{i,n}^{(n)} = 1, \quad (i = 1, 2, \dots, n), \quad (3.12)$$

only $n(n - 1)$ of them will be independent.

To find these probabilities, note that in an infinitesimal time dt there will be effectively zero probability of more than one molecule jumping between adjacent cells. If the boundaries of our system are reflective (i.e. diffusive jumps between subvolumes 1 and n are not allowed), then the addition and multiplication laws of probability yield

$$\begin{aligned} p_{i1}(t + dt) &= p_{i1}(t) \times [1 - \kappa dt] + p_{i2}(t) \times \kappa dt, \\ p_{ij}(t + dt) &= p_{i(j-1)}(t) \times \kappa dt + p_{ij}(t) \times [1 - 2\kappa dt] + p_{i(j+1)}(t) \times \kappa dt, \\ & \hspace{25em} (j = 2, \dots, n - 1) \\ p_{in}(t + dt) &= p_{i(n-1)}(t) \times \kappa dt + p_{in}(t) \times [1 - \kappa dt] \end{aligned} \quad (3.13)$$

The first equation in (3.13) parses thusly: {The probability that a molecule will be in subvolume 1 at time $t + dt$ given that it was in state i at time 0} is equal to the *sum* of {the probability that the molecule was in subvolume 1 at time t , given that it was in subvolume i at time 0 *and* it did not jump away from subvolume

1 in the next dt } plus {the probability that the molecule was in subvolume 2 at time t , given that it was in subvolume i at time 0 and it jumped from subvolume 1 to subvolume 2 in the next dt }. All other routes to subvolume 1 at time $t + dt$ from a subvolume other than 1 or 2 at time t will be second order in dt (and will thus make no contribution when (3.13) is later converted to an ODE).

If the boundaries of our system are periodic (i.e. subvolumes 1 and n communicate), then we have

$$\begin{aligned}
 p_{i1}(t + dt) &= p_{in}(t) \times \kappa dt + p_{i1}(t) \times [1 - 2\kappa dt] + p_{i2}(t) \times \kappa dt, \\
 p_{ij}(t + dt) &= p_{i(j-1)}(t) \times \kappa dt + p_{ij}(t) \times [1 - 2\kappa dt] + p_{i(j+1)}(t) \times \kappa dt, \\
 & \hspace{20em} (j = 2, \dots, n - 1) \\
 p_{in}(t + dt) &= p_{i(n-1)}(t) \times \kappa dt + p_{in}(t) \times [1 - 2\kappa dt] + p_{i1}(t) \times \kappa dt \quad (3.14)
 \end{aligned}$$

The discussion which follows can be made independent of boundary condition by using the concept of the Laplacian matrix of a graph. There is an isomorphism between the discretization of our system into subvolumes, and a directed graph (a collection of vertices and directed edges). Each subvolume of our system can be represented by a vertex. We can then connect with a directed edge those vertices which correspond to allowable diffusive transfers. The resulting graph G is just another representation of our original system, with vertices denoting

the possible locations of molecules, and edges denoting the possible transitions (diffusive jumps) between those locations. When our system has n subvolumes and periodic boundary conditions, then the resulting graph is R_n , the so-called ring graph with n vertices; a system with n subvolumes and reflecting boundary conditions yields L_n , the *line* graph (see Figure 3.1).

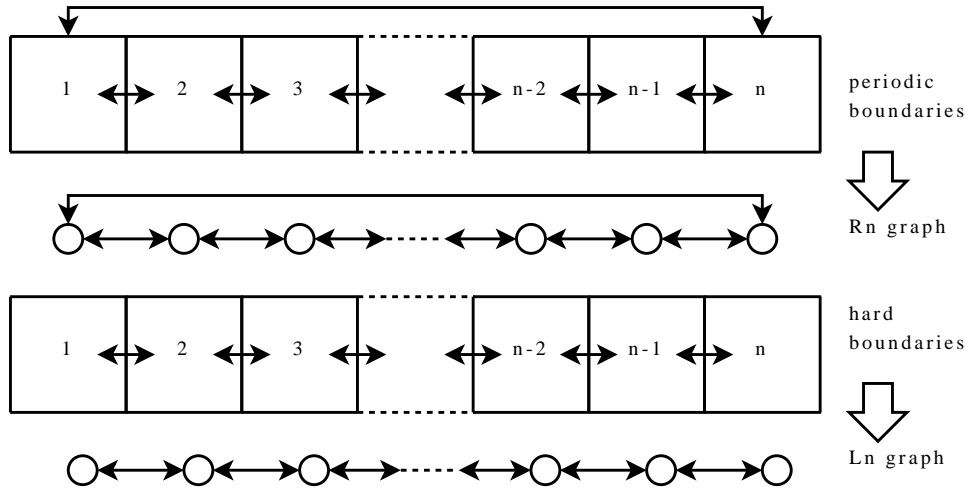


Figure 3.1: Boundary conditions and the resulting graphs. The solid arrows show the allowed diffusive jumps.

Equations (3.13) and (3.14) both lead to the general set of differential equations

$$\frac{d\mathbf{p}(t)}{dt} = -\kappa \mathcal{L}_{GP}(t), \quad (3.15)$$

where

$$\mathbf{p}(t) = \begin{pmatrix} p_{i1}(t) \\ p_{i2}(t) \\ \dots \\ p_{in}(t) \end{pmatrix} \quad (3.16)$$

and \mathcal{L}_G is the so-called Laplacian matrix of the graph $G \in \{L_n, R_n\}$, with entries:

$$\mathcal{L}_G(i, j) = \begin{cases} \text{number of neighbors of } i, & \text{if } i = j \\ -1, & \text{if } i \neq j, \text{ and } i \text{ is adjacent to } j \\ 0, & \text{if } i \neq j, \text{ and } i \text{ is not adjacent to } j \end{cases} \quad (3.17)$$

For the initial condition $\mathbf{p}(0)$, the solution to (3.15) is

$$\mathbf{p}(t) = \mathbf{V} \cdot e^{-\lambda \kappa t} \cdot \mathbf{V}^{-1} \cdot \mathbf{p}(0) \quad (3.18)$$

where \mathbf{V} is the matrix of eigenvectors of \mathcal{L}_G , and

$$e^{-\lambda \kappa t} \equiv \begin{pmatrix} e^{-\lambda_1 \kappa t} & 0 & \dots & 0 \\ 0 & e^{-\lambda_2 \kappa t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{-\lambda_n \kappa t} \end{pmatrix} \quad (3.19)$$

where λ_i ($i = 1, \dots, n$) is the i^{th} eigenvalue of $\mathcal{L}_{\mathcal{G}}$.

Now we introduce the random variables

$$M_{ij}^{(n)}(k_i, t) \equiv \text{the number of the } k_i \text{ molecules in subvolume } i \text{ at time } 0 \text{ that will} \\ \text{be in subvolume } j \text{ at time } t, (i, j = 1, 2, \dots, n). \quad (3.20)$$

These n^2 random variables satisfy the relations

$$M_{i1}^{(n)}(k_i, t) + M_{i2}^{(n)}(k_i, t) + \dots + M_{in}^{(n)}(k_i, t) = k_i \quad (i = 1, 2, \dots, n) \quad (3.21)$$

so only $n(n-1)$ of them will be independent. We choose the independent variables to be $M_{ij}^{(n)}(k_i, t)$ for $i \neq j$. There will be n such statistically independent sets.

Consider first the $(n-1)$ variables $M_{1j}^{(n)}(k_1, t)$ for $j = 2, \dots, n$. These are statistically independent of the $(n-1)^2$ variables $M_{ij}^{(n)}(k_i, t)$ for $i, j = 2, \dots, n$, because individual molecules move independently of each other, but the $M_{ij}^{(n)}(k_i, t)$ are *not* statistically independent of each other.

Denote the joint probability density function of the $(n-1)$ subvolume 1 random variables by

$$P_{2,\dots,n}^{(1;n)}(m_{12}, m_{13}, \dots, m_{1n}; k_1, t) \equiv \text{Prob}\{M_{1j}^{(n)}(k_1, t) = m_{1j} \text{ for } j = 2, \dots, n\}. \quad (3.22)$$

From the addition and multiplication laws of probability we have:

$$\begin{aligned} P_{2,\dots,n}^{(1;n)}(m_{12}, m_{13}, \dots, m_{1n}; k_1, t) &= \frac{k_1!}{m_{12}!m_{13}!\dots m_{1n}!(k_1 - m_{12} - m_{13} - \dots - m_{1n})} \\ &\times \left[\left(p_{12}^{(n)}(t)\right)^{m_{12}} \left(p_{13}^{(n)}(t)\right)^{m_{13}} \dots \left(p_{1n}^{(n)}(t)\right)^{m_{1n}} \right. \\ &\left. \left(1 - p_{12}^{(n)}(t) - p_{13}^{(n)}(t) - \dots - p_{1n}^{(n)}(t)\right)^{k_1 - m_{12} - m_{13} - \dots - m_{1n}} \right]. \quad (3.23) \end{aligned}$$

The second factor on the right hand side is the probability that, of the k_1 molecules in subvolume 1 at time 0, a particular set of m_{12} of them will wind up in subvolume 2 at time t , and a particular set of m_{13} of them will wind up in subvolume 3 at time t , and so on, with the remaining $k_1 - m_{12} - m_{13} - \dots - m_{1n}$ molecules remaining in subvolume 1 at time t . The first factor on the right hand side of (3.23) is the number of ways of choosing groups of $m_{12}, m_{13}, \dots, m_{1n}$ molecules from k_1 molecules. The joint probability function (3.23) implies that the random variables $M_{1i}^{(n)}$ for $i = 2, \dots, n$, are multinomially distributed. We now observe

that (3.23) is algebraically identical to:

$$\begin{aligned}
 P_{2,\dots,n}^{(1;n)}(m_{12}, m_{13}, \dots, m_{1n}; k_1, t) &= \frac{k_1!}{m_{12}!(k_1 - m_{12})!} \left(p_{12}^{(n)}(t)\right)^{m_{12}} \left(1 - p_{12}^{(n)}(t)\right)^{k_1 - m_{12}} \\
 &\times \frac{(k_1 - m_{12})!}{m_{13}!(k_1 - m_{12} - m_{13})!} \left(\frac{p_{13}^{(n)}(t)}{1 - p_{12}^{(n)}(t)}\right)^{m_{13}} \left(1 - \frac{p_{13}^{(n)}(t)}{1 - p_{12}^{(n)}(t)}\right)^{k_1 - m_{12} - m_{13}} \\
 &\times \dots \times \frac{(k_1 - m_{12} - \dots - m_{1(n-1)})!}{m_{1n}!(k_1 - m_{12} - \dots - m_{1n})!} \left(\frac{p_{1n}^{(n)}(t)}{1 - p_{12}^{(n)}(t) - \dots - p_{1(n-1)}^{(n)}(t)}\right)^{m_{1n}} \\
 &\times \left(1 - \frac{p_{1n}^{(n)}(t)}{1 - p_{12}^{(n)}(t) - \dots - p_{1(n-1)}^{(n)}(t)}\right)^{k_1 - m_{12} - \dots - m_{1n}} \quad (3.24)
 \end{aligned}$$

The significance of (3.24) is that it immediately implies the conditioning

$$\begin{aligned}
 P_{2,\dots,n}^{(1;n)}(m_{12}, m_{13}, \dots, m_{1n}; k_1, t) &= P_2^{(1;n)}(m_{12}; k_1, t) \times P_{3|2}^{(1;n)}(m_{13}|m_{12}; k_1, t) \\
 &\times \dots \times P_{n|2,\dots,(n-1)}^{(1;n)}(m_{1n}|m_{12}, \dots, m_{1(n-1)}; k_1, t) \quad (3.25)
 \end{aligned}$$

where

$$P_2^{(1;n)}(m_{12}; k_1, t) = \mathbb{P}_{\mathcal{B}} \left(m_{12}; p_{12}^{(n)}(t), k_1 \right), \quad (3.26)$$

$$P_{3|2}^{(1;n)}(m_{13}|m_{12}; k_1, t) = \mathbb{P}_{\mathcal{B}} \left(m_{13}; \frac{p_{13}^{(n)}(t)}{1 - p_{12}^{(n)}(t)}, k_1 - m_{12} \right), \quad (3.27)$$

...

$$P_{n|2,\dots,(n-1)}^{(1;n)}(m_{1n}|m_{12}, \dots, m_{1(n-1)}; k_1, t) = \mathbb{P}_{\mathcal{B}} \left(m_{1n}; \frac{p_{1n}^{(n)}(t)}{1 - p_{12}^{(n)}(t) - \dots - p_{1(n-1)}^{(n)}(t)}, k_1 - m_{12} - \dots - m_{1(n-1)} \right) \quad (3.28)$$

with $\mathbb{P}_{\mathcal{B}}$ the binomial pdf

$$\mathbb{P}_{\mathcal{B}}(m; p, n) = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}$$

The physical interpretation of this result is as follows: the number m_{12} of the k_1 molecules in subvolume 1 at time 0 that will be found in subvolume 2 at time t , irrespective of the fates of the other $(k_1 - m_{12})$ molecules, can be chosen by sampling the binomial distribution with parameters $p_{12}^{(n)}(t)$ and k_1 . Once the number m_{12} has been selected in this way, the number m_{13} of the remaining $(k_1 - m_{12})$ molecules that will be found in subvolume 3 at time t can be chosen by sampling the binomial distribution with parameters $p_{13}^{(n)}(t)/(1 - p_{12}^{(n)})$ and $(k_1 -$

m_{12}). This procedure can be repeated to generate the remaining m_{1i} for $i = 4, \dots, n$ as samples of the binomial distribution with parameters given by (3.28).

Equations (3.26-3.28) show how to generate the time t fates of the molecules that are in subvolume 1 at time 0. The time t fates of the k_i molecules in subvolume i at time 0, for $i = 2, \dots, n$, are independent of those in any other subvolume, and the procedure for determining them is analogous.

Some additional approximations

At this point it may seem that we have specified an algorithm for generating the number m_{ij} of molecules moving from subvolume i to subvolume j in time t for all $i \neq j$. However this algorithm has a serious drawback, which renders it practically unusable: it requires $O(n^2)$ samples of the binomial distribution per time-step. Generating $O(n^2)$ binomial samples is likely to be a prohibitive computational burden, even for modest n . Furthermore, each $(n-1)$ of the samples are dependent, limiting any speedup that may be obtainable by parallelizing the binomial sample generation.

In this section we take three steps to obtain an algorithm which does not have this quadratic complexity disadvantage. First, to obtain linear complexity, we limit the distance any molecule can diffuse in a single time-step. Second, to maintain accuracy in spite of this approximation, we impose an upper limit on

the time-step. Third, to scale the algorithm to large system sizes, we approximate the diffusion probabilities of systems of arbitrary size n by those of a small, finite system of size \hat{n} .

Step 1: Ideally, rather than $O(n^2)$, we would prefer to generate only $O(n)$ binomial samples per time-step. This can be achieved if we restrict where molecules can go: if a molecule, rather than having n choices of destination subvolume, instead only has a constant number of choices, then only $O(n)$ binomial samples per time-step will be required. By neglecting subvolumes outside a radius s of the subvolume of origin, we reduce the number of binomial samples required from $(n - 1)^2$ (with each $(n - 1)$ dependent) to $2sn$ (with each $2s$ dependent).

Step 2: For an algorithm based on a limited diffusion radius to be accurate, the size of the time-step must be restricted. The *time-step restriction* should satisfy the following condition: the probability of a molecule jumping from subvolume i to any subvolume beyond a radius of s subvolumes away from i in time Δt , should be less than or equal to a given ε . This probability per molecule per time-step represents the *error* from “corralling” the molecules within a radius of s subvolumes from their subvolume of origin in any given time-step. We will denote this error $e_i^{(n)}$, where the subscript i refers to the subvolume of origin, and the superscript (n) refers to the total number of subvolumes in the system. In the case of a periodic system, the subscript i can be dropped (as it will be in Figures 3.2

and 3.3), since the diffusion probabilities, and therefore the error, are identical for all origin subvolumes. Generally (for both periodic and reflective boundaries) the probability that a molecule will, in time Δt , diffuse more than s subvolumes away from its original subvolume i is

$$e_i^{(n)}(s, \Delta t) = 1 - \sum_{j \in J(i,s)} p_{ij}^{(n)}(\Delta t) \quad (3.29)$$

where $J(i, s)$ is the set of subvolumes within a radius of s subvolumes from i (including i).

Thus, if we are willing to incur $e_i^{(n)}(s, \Delta t) \leq \varepsilon$ error in probability per molecule per time-step, we can restrict the distance a molecule can travel from subvolume i in time Δt , to s subvolumes from i in either direction by taking the time-step Δt to be less than or equal to Δt_{max} , where Δt_{max} is given by the solution to $e_i^{(n)}(s, \Delta t_{max}) = \varepsilon$.

The elements $p_{ij}(t)$ of $\mathbf{p}(t)$ (Eq. 3.18) are probabilities which are always functions of the product κt , where κ depends on the diffusion coefficient of the molecular species. Thus, in practice, the maximum time-step Δt_{max} will always be a function of κ .

Step 3: We have shown how to reduce the complexity of the algorithm by limiting the diffusion radius to s , and how to ensure that a level of accuracy ε

is satisfied by limiting the time-step Δt . But up to this point our analysis has depended on the system size n . We will next show how the dependence on the system size n can be dropped, allowing the algorithm to be applied to systems of arbitrary size.

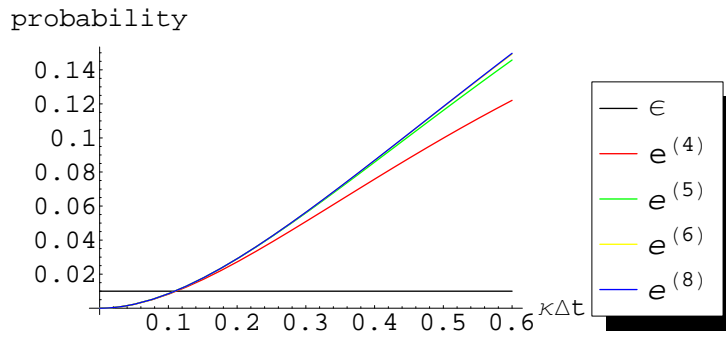


Figure 3.2: A plot of $e_i^{(n)}(s = 1, t)$ versus $\kappa\Delta t$, for a system with periodic boundaries, and different values of n . The horizontal black line indicates error per molecule per time-step $\varepsilon = 1\%$. Note that increasing the system size does not appreciably change the error.

For $t \leq \Delta t_{max}$ it is possible to find a system size \hat{n} , such that the probabilities $p_{ij}^{(n)}(t)$, for $j \in J(i, s)$, are nearly indistinguishable for all $n > \hat{n}$. The error, being a function of these probabilities (see (3.29)), will *also* be indistinguishable for all $n > \hat{n}$. To illustrate this, consider four systems with periodic boundary conditions and $n = 4, 5, 6$ and 8 subvolumes. Figure 3.2 shows the probability of going *past* a radius $s = 1$, i.e. the error $e_i^{(n)}(s = 1, t)$, for these systems. These probabilities were obtained analytically using Mathematica to solve (3.18). A series expansion (again, performed using Mathematica) reveals that for all $n =$

4, 5, 6, 8, $e_i^{(n)}(s = 1, t) = (\kappa\Delta t)^2 + O((\kappa\Delta t)^3)$. Thus for small $\kappa\Delta t$ we do not expect these probabilities to have significantly different values. Indeed, for probability $\leq 1\%$ (horizontal black line), there is almost no visible difference in the error if we compare these systems; for n past $\hat{n} = 4$, the error does not appreciably increase with increasing n .

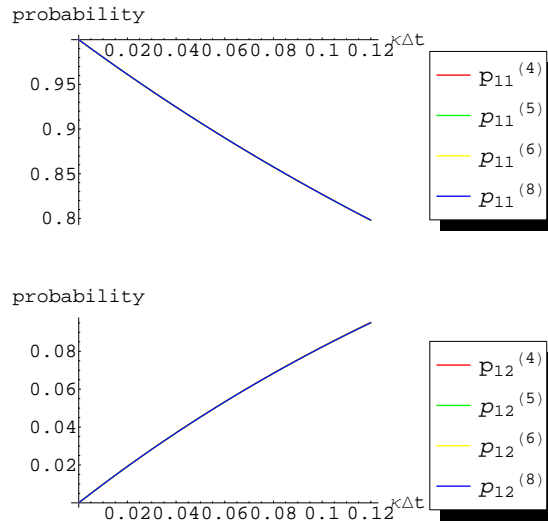


Figure 3.3: Plots of $p_{11}^{(n)}(\kappa\Delta t)$ and $p_{12}^{(n)}(\kappa\Delta t)$ versus $\kappa\Delta t$, for a system with periodic boundaries, and for different values of n .

Figure 3.3 shows that the same pattern holds for the probabilities $p_{11}^{(n)}(t)$ and $p_{12}^{(n)}(t)$, individually. As $t \rightarrow \infty$, $p_{ij}^{(n)}(t) \rightarrow \frac{1}{n}$, i.e. the probabilities tend to a uniformly random distribution. When we perform a series expansion, we see that, for all n , the $p_{11}^{(n)}(t)$ share a leading term which is $O(1)$, the $p_{12}^{(n)}(t)$ share a leading term which is $O(\kappa\Delta t)$, and so on. Thus, for *small* $\kappa\Delta t$, consistent with

$\varepsilon = 1\%$, these probabilities, which we will use directly in the algorithm, are also indistinguishable for $n > \hat{n} = 4$.

This observation suggests a way to scale the algorithm to arbitrary system sizes, given a desired per molecule per time-step error of $\varepsilon = 1\%$ (horizontal black line): since for all $n > 4$, $e_i(n)(1, t) \approx e_i^{(4)}(1, t)$ for $\kappa\Delta t$ consistent with ε , then the probabilities with superscript $\hat{n} = 4$, corresponding to a system with 4 subvolumes, can be used in place of the probabilities of any larger system.

In addition, the observation that $e_i^{(4)}(1, t) \approx (\kappa\Delta t)^2$ suggests that for $s = 1$ we can choose a conservative maximum time-step consistent with a level of error less than or equal to ε by satisfying

$$\Delta t \leq \sqrt{\varepsilon}/\kappa \tag{3.30}$$

This gives a formula for choosing the time-step.

To summarize, the steps that must be followed in order to obtain a practical algorithm from the theory of the previous section are as follows:

1. Choose a diffusion radius s , and a given level of error ε ;
2. Choose Δt_{max} to satisfy $e_i^{(n)}(s, \Delta t_{max}) = \varepsilon$, as a function of κ (for $s = 1$ use (3.30); for $s > 1$ similar formulas exist);
3. Find \hat{n} which satisfies $e_i^{(n)}(s, \Delta t) \approx e_i^{(\hat{n})}(s, \Delta t)$, $\forall n > \hat{n}$ and $\Delta t \leq \Delta t_{max}$.

Implementation of the algorithm

There is one practical consideration in the implementation of the algorithm which we have not yet addressed. Because the sum total of the probabilities of the events which can occur in the simulation must be unity, the probability $e_i^{(n)}(s, \Delta t)$ of going beyond the diffusion radius must be *reassigned* to an event which *can* occur during the simulation. Where should we reassign this probability?

According to our tests, two different strategies work best in two distinct cases. If the subvolume of origin i is an interior subvolume, the best accuracy is achieved by adding $\frac{1}{2}e_i^{(n)}(s, \Delta t)$ to the two probabilities of going as far away as possible from i in either direction. In a periodic system, all subvolumes fall in this category.

For reflective boundary systems, we have found that if the subvolume of origin i has a boundary close to it, the best accuracy is achieved by adding $e_i^{(n)}(s, \Delta t)$ to the probability of staying in subvolume i .

This distinction makes it clear that we need a shorthand notation for the probabilities we will use in the *implementation* of the algorithm. Thus we define

$$\hat{p}_{ij}(t; s, \Delta t) \equiv \text{the system-size independent probability that a single molecule which was in subvolume } i \text{ at time } 0, \text{ will be in subvolume } j \text{ at time } t, \text{ as it will be used in the algorithm with } s \text{ and } \Delta t \quad (3.31)$$

For example, for an interior subvolume i and diffusion radius $s = 1$, the formulas are given by

$$\hat{p}_{i,(i+1)}(t; 1, \Delta t) = \hat{p}_{i,(i-1)}(t; 1, \Delta t) \equiv p_{i,(i+1)}^{(\hat{n})}(t) + \frac{1}{2}e_i^{(\hat{n})}(1, \Delta t) \quad (3.32)$$

For a subvolume directly abutting a reflective boundary on one side, we modify the probability of staying in that subvolume, yielding the formula

$$\hat{p}_{ii}(t; 1, \Delta t) \equiv p_{ii}^{(\hat{n})}(t) + e_i^{(\hat{n})}(1, \Delta t) \quad i = 1, n \quad (3.33)$$

We are now ready to give the procedure for approximate multinomial diffusion for a system with n subvolumes, each of length l , and a single species \mathbf{X} with diffusion coefficient D . The algorithm first computes the $2sn$ values of the variables

ΔX_{ij} , for $i = 1, \dots, n$ and $j = i \pm 1, \dots, i \pm s$, giving the number of molecules which will move from subvolume i to a subvolume j , to the right ($j = i + 1, \dots, i + s$) or to the left ($j = i - 1, \dots, i - s$) of i . A second loop then applies these population changes to the state X_i , $i = 1, \dots, n$, and finally the time is incremented. The function $\mathcal{B}(p, n)$ generates random numbers distributed according to the binomial distribution with parameters p and n (Eq. 3.29). For the sake of simplicity, we will present the algorithm for $s = 1$.

Algorithm 3 Diffusion in one dimension with diffusion radius $s = 1$

```

Choose  $s$  and  $\varepsilon$ 
Calculate  $\Delta t_{max}$  as a function of  $\varepsilon$ ,  $s$ , and  $\kappa = D/l^2$ 
Choose  $\Delta t \leq \Delta t_{max}$ 
while  $t \leq t_{final}$  do
    for  $i = 1$  to  $n$  do
         $\Delta X_{i(i+1)} = \mathcal{B}(\hat{p}_{i(i+1)}(\Delta t), X_i)$ 
         $\Delta X_{i(i-1)} = \mathcal{B}\left(\frac{\hat{p}_{i(i-1)}(\Delta t)}{1 - \hat{p}_{i(i+1)}(\Delta t)}, X_i - \Delta X_{i(i+1)}\right)$ 
    end for
    for  $i = 1$  to  $n$  do
         $X_i = X_i - \Delta X_{i(i+1)} - \Delta X_{i(i-1)} + \Delta X_{(i+1)i} + \Delta X_{(i-1)i}$ 
    end for
     $t = t + \Delta t$ 
end while

```

Error analysis for $s = 1$

Of the three steps outlined in Subsection 3.3.2, steps 1 and 3 represent approximations, and each one introduces some error to our simulation. We can gain some intuition about the *relative* magnitude of the two errors by revisiting Fig-

ure 3.2. The error from the restriction of the diffusion radius to s (step 1) is given by the $e^{(\hat{n})}$ curve. The error from the approximation of the probabilities of an arbitrary-sized system, by the probabilities of a \hat{n} -sized system (step 3) is given by the *difference* between the $e^{(\hat{n})}$ curve and the $e^{(n)}$ curves with $n > \hat{n}$. In this example $\hat{n} = 4$. While the step 1 error is plainly large (but less than ε), the step 3 error is negligible by comparison.

We have already pointed out that the error per molecule per time-step due to the restriction of the diffusion radius (step 1), for a system with *periodic* boundaries and $s = 1$, is $O((\kappa\Delta t)^2)$. The case of *reflective* boundaries is a little more difficult to analyze, but the answer turns out to be the same.

In a system with *reflective* boundaries, and $s = 1$, we recognize that we will have to consider three “classes” of subvolumes. *Class 1* contains the two subvolumes closest to the boundary (subvolumes 1 and n); *Class 2* contains the two subvolumes which are one subvolume removed from the boundary (2 and $(n - 1)$); *Class 3* contains the remaining subvolumes (subvolumes i with $3 \leq i \leq (n - 2)$), which we shall call “interior” subvolumes. Where the subscript i on the error $e_i^{(n)}$ previously denoted the subvolume of origin, we will now parenthesize ($e_{(i)}^{(n)}$, $i = 1, 2, 3$) it to denote the *class* of subvolume.

The probabilities of diffusing away from each class of subvolume are given by different formulas. The interior subvolumes (indexed $3, \dots, (n - 2)$) are assumed

to be sufficiently far from the boundary so that they do not “feel” its effect. Their diffusion probabilities will be taken to be those from a periodic system. Figure 3.2 has already shown us that $\hat{n} = 4$ for a periodic system. Thus, for the *class 3* (interior) subvolumes of a reflective boundary system we have error

$$e_{(3)}^{(4)} = 1 - (p_{11}^{(R4)} - 2p_{12}^{(R4)}) = p_{13}^{(R4)} \quad (3.34)$$

where the superscript *R4* denotes that the probabilities are taken from a periodic system (“R” stands for “ring”) with 4 subvolumes. We have already detailed in Subsection 3.3.2 that this error is $O((\kappa\Delta t)^2)$, and that to achieve an error level ε we must satisfy (3.30).

To decide on a value for \hat{n} for class one and class two subvolumes, we need to consult the error from reflective boundary systems with $n = 4$ and $n = 6$ subvolumes. These can be obtained analytically in the same way that we obtained the periodic boundary probabilities, using Mathematica to solve (3.18). The *class one* errors (for subvolumes indexed 1 and n) are given by:

$$e_{(1)}^{(4)}(1, \Delta t) = 1 - (p_{11}^{(L4)} - p_{12}^{(L4)}) = p_{13}^{(L4)} + p_{14}^{(L4)} \quad (3.35)$$

$$e_{(1)}^{(6)}(1, \Delta t) = 1 - (p_{11}^{(L6)} - p_{12}^{(L6)}) = p_{13}^{(L6)} + p_{14}^{(L6)} + p_{15}^{(L6)} + p_{16}^{(L6)} \quad (3.36)$$

where the superscripts $L4$ and $L6$ represent probabilities from the reflective boundary system (“L” stands for “line”) with 4 and 6 subvolumes, respectively. Performing a series expansion on these errors gives

$$e_{(1)}^{(4)}(1, \Delta t) = \frac{1}{2}(\kappa\Delta t)^2 - \frac{2}{3}(\kappa\Delta t)^3 + \frac{7}{12}(\kappa\Delta t)^4 - \frac{2}{5}(\kappa\Delta t)^5 + \frac{41}{180}(\kappa\Delta t)^6 + O((\kappa\Delta t)^7) \quad (3.37)$$

$$e_{(1)}^{(6)}(1, \Delta t) = \frac{1}{2}(\kappa\Delta t)^2 - \frac{2}{3}(\kappa\Delta t)^3 + \frac{7}{12}(\kappa\Delta t)^4 - \frac{2}{5}(\kappa\Delta t)^5 + \frac{11}{48}(\kappa\Delta t)^6 + O((\kappa\Delta t)^7) \quad (3.38)$$

Two things are notable. First, the errors differ in the sixth and higher order terms. This means that they are practically indistinguishable, and that we can take $\hat{n} = 4$. Second, the leading term is $O((\kappa\Delta t)^2)$, as it was for class three, but the coefficient is $\frac{1}{2}$, i.e. half that of the error for class three. We could have foreseen that using the following reasoning: molecules from class one subvolumes have half as many opportunities to leave their subvolume of origin as do molecules from interior subvolumes. From this observation we conclude that the error in class one subvolumes, being approximately half that of class three subvolumes, will not impose a further limitation on the time-step.

Our reasoning for *class two* subvolumes (indexed 2 and $(n - 1)$) is completely analogous. The errors for $n = 4$ and $n = 6$ are

$$e_{(2)}^{(4)}(1, t) = 1 - (p_{22}^{(L4)} - p_{21}^{(L4)} - p_{23}^{(L4)}) = p_{24}^{(L4)} = \frac{1}{2}(\kappa\Delta t)^2 + O((\kappa\Delta t)^3) \quad (3.39)$$

$$e_{(2)}^{(6)}(1, t) = 1 - (p_{22}^{(L6)} - p_{21}^{(L6)} - p_{23}^{(L6)}) = p_{24}^{(L6)} + p_{25}^{(L6)} + p_{26}^{(L6)} = \frac{1}{2}(\kappa\Delta t)^2 + O((\kappa\Delta t)^3) \quad (3.40)$$

Since they differ in higher order terms, we shall use $\hat{n} = 4$. Since the leading term in the error is half that of class three subvolumes, it will not restrict the time-step further.

Stability analysis for $s=1$

The per molecule per time-step error due to the diffusion radius restriction is a “local” error. In this section we show that the “global” error in the simulation mean is bounded and $O(\kappa\Delta t)$.

The expected value of a binomial random variable $\mathcal{B}(p, n)$ is np . Given x molecules in a subvolume, and probabilities $\hat{p}(\Delta t)$ of jumping either left and right in time Δt , we can say the following: The number of molecules that will jump to the right in the next Δt is $\mathcal{B}(\hat{p}(\Delta t), x)$. This implies that the mean number jumping to the right in the next Δt is $\hat{p}(\Delta t)x$. If we are *given* that r of the x molecules *do* jump to the right, then the number of the $(x - r)$ remaining

molecules that will jump to the left in the next Δt is $\mathcal{B}(\hat{p}(\Delta t)/(1 - \hat{p}(\Delta t)), x - r)$. This implies that the mean number of the x molecules that jump to the left, *given* that r of those x molecules have jumped to the right, is $[\hat{p}(\Delta t)/(1 - \hat{p}(\Delta t))][x - r]$. We can eliminate this conditioning by using the *iterated expectation* formula ($E(X) = E(E(X|Y))$). Then the mean number of molecules jumping to the left, unconditionally, reduces to $\hat{p}(\Delta t)x$, the same as jumping to the right, unconditionally.

Extending this idea the full system, we can obtain an update formula for the *mean population* evolving through multinomial diffusion with $s = 1$. The most condensed form of the update formula is

$$\mathbf{X}_{N+1} = (\mathbf{I} + \mathbf{B})\mathbf{X}_N \quad (3.41)$$

where \mathbf{X}_N is the state (as a column vector) at time-step N , \mathbf{I} is the identity matrix, and \mathbf{B} is the matrix with elements: $-(\hat{p}_{i,(i+1)}(\Delta t) + \hat{p}_{i,(i-1)}(\Delta t))$ on the i^{th} row of the diagonal; $\hat{p}_{(i-1),i}(\Delta t)$ on the $(i, (i-1))$ subdiagonal positions; $\hat{p}_{(i+1),i}(\Delta t)$ on the $(i, (i+1))$ superdiagonal positions; and zeros everywhere else. ¹

The next-nearest neighbor diffusion probabilities $\hat{p}_{i,(i+1)}(\Delta t)$ and $\hat{p}_{i,(i-1)}(\Delta t)$ can be series expanded, and shown to be $\kappa\Delta t + O((\kappa\Delta t)^2)$. Update formula (3.41)

¹For these definitions to work, we set: (a) for a reflective boundary system $\hat{p}_{1,0}(\Delta t) = \hat{p}_{n,(n+1)}(\Delta t) \equiv 0$, or (b) for a periodic boundary system, $\hat{p}_{1,0}(\Delta t) \equiv \hat{p}_{1,n}(\Delta t)$ and $\hat{p}_{n,(n+1)}(\Delta t) \equiv \hat{p}_{n,1}(\Delta t)$.

can then be approximated as $\mathbf{X}_{N+1} = (I + \Delta t \hat{\mathbf{B}} + O((\kappa \Delta t)^2)) \mathbf{X}_N$, where

$$\hat{\mathbf{B}} = \kappa \begin{pmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & -1 \end{pmatrix} \quad (3.42)$$

As we showed in the previous section, the error per molecule per time-step is $O((\kappa \Delta t)^2)$. Thus, the mean population satisfies a forward-time, centered-space approximation to the diffusion equation, as we would expect. Standard results from the numerical analysis of PDEs ([33]) yield stability as $\Delta t \rightarrow 0$ on a fixed time interval, and convergence to accuracy $O(\kappa \Delta t)$.

Diffusion radius $s=2$

Thus far we have mainly discussed the situation in which the diffusion radius is $s = 1$. Increasing the radius to $s = 2$ subvolumes on either side of the subvolume of origin, while maintaining same error $e_i(1, \Delta t) \leq 1\%$, will yield a longer $\kappa \Delta t_{max} \approx 0.4$, as is illustrated in Figure 3.4. For $s = 2$, $\varepsilon = 1\%$, and periodic boundaries, we have $\hat{n} = 6$, and $e^{(6)}(\Delta t) = \frac{1}{3}(\kappa \Delta t)^3 + O((\kappa \Delta t)^4)$. For reflective boundaries, there are four classes of subvolumes, numbered in increasing order from the closest

to the boundary (class one) to the interior subvolumes (class four). Class four subvolumes again dominate the error and determine the step-size restriction.

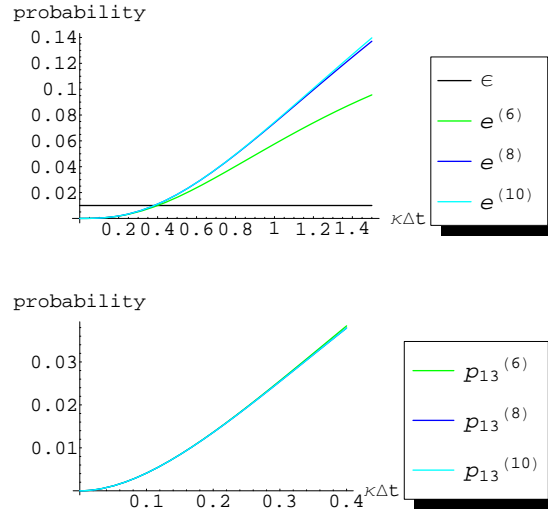


Figure 3.4: Plots of $e_i^{(n)}(s = 2, \kappa\Delta t)$ and $p_{13}^{(n)}(\kappa\Delta t)$ versus $\kappa\Delta t$, for a system with periodic boundaries, and for different values of n .

Because the choice $s = 2$ doubles the number of binomial samples required for a single time-step, it will also increase the computational time required per time-step, by two. In a diffusion-only setting, the $s = 2$ algorithm will take one quarter as many steps as the $s = 1$ algorithm, and use up half as much computational time. However, as we will show in the Simulation Results section, once reactions are added to the mix, the computational and accuracy advantages of the $s = 2$ algorithm will only manifest themselves in situations where reactions are spaced *overwhelmingly* farther apart than diffusive transfers.

3.3.3 Reaction-diffusion in one dimension

The algorithm

Our stated goal was to create an algorithm which will be faster than the ISSA for systems in which diffusion is much faster than reaction, and still accurately represent small population stochastic phenomena. We are now ready to describe this algorithm, which we call the Multinomial Simulation Algorithm. It incorporates Algorithm 3 for diffusion in one dimension as one element, while its other element is the firing of reactions according to the usual SSA scheme.

The system is divided into the usual n subvolumes of length l , but now contains more than one species. The state is given by the matrix \mathbf{X} , where X_{ij} is the population of the j^{th} species in the i^{th} subvolume. The diffusion coefficient of the j^{th} species is D_j . The reaction propensity functions α_{ir} give the propensity of the r^{th} reaction in the i^{th} subvolume, and a_0 is the total reaction propensity $\alpha_0 = \sum_{i=1}^n \sum_{r=1}^R \alpha_{ir}$. The variable \mathcal{U} represents a uniform random number in the interval $(0, 1)$. The time to the next reaction is given by τ , while the maximum time-step for diffusion is given by Δt_{max} .

Unlike the GMP method [31], which performs diffusion steps at time-points which are completely decoupled from the times at which reactions fire, the MSA couples the diffusion and reaction time-steps. First, the time τ to the next reac-

Algorithm 4 Reaction-diffusion in one dimension

Choose s and ε
 Calculate Δt_{max} as a function of ε , s , and $\max_i \{\kappa_i = D_i/l^2\}$
while $t \leq t_{final}$ **do**
 Calculate total reaction propensity $\alpha_0 = \sum_{i=1}^n \sum_{r=1}^R \alpha_{ir}$
 Pick time to next reaction as $\tau = -\ln(\mathcal{U})/\alpha_0$
 Pick indices i (subvolume) and j (reaction) of next reaction as in the SSA
 if $(t + \tau) \leq t_{final}$
 Remove reactants of reaction j in subvolume i
 reacted = True
 else
 $\tau = t_{final} - t$
 end if
 tmp = 0
 while $(\tau - \text{tmp}) \geq \Delta t_{max}$
 Take a Δt_{max} diffusion jump for all species (See Algorithm 1)
 tmp = tmp + Δt_{max}
 end while
 Take a $(\Delta t = \tau - \text{tmp})$ diffusion jump for all species (See Algorithm 1)
 if **reacted** is True
 Add products of reaction j in subvolume i
 reacted = False
 end if
 $t = t + \tau$
end while

tion, as well as the type and location of the reaction, is chosen. Then the reactants are immediately removed, and diffusive steps are taken until time τ is reached. Then the products of the reaction appear. This may seem somewhat strange, but it is the least complicated and most accurate strategy we have found. The alternative of both removing the reactants and producing the products at the beginning of the reaction step is less accurate. The alternative of doing both at the end of

the step does not guarantee that the reactants will still be at the same location after diffusion has occurred.

Simulation results and error analysis

We have three goals in this section: a) to establish that the multinomial method gives qualitatively correct results, b) to quantify the performance of the multinomial method compared to the ISSA, and c) to quantify the error between the multinomial method and the ISSA.

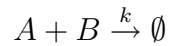
The MSA and ISSA codes on which the results in this paper are based are written in ANSI C, and the two methods are driven by a common problem description file. The ISSA implementation is based on the direct method, with only the most obvious optimizations: avoiding the recalculation of diffusion propensities for species which did not change in the previous time-step, and of reaction propensities in subvolumes which were not touched in the previous time-step. We use the shorthand $\text{MSA}(s)$ for the MSA with diffusion radius set to s .

We have already laid out the logic by which the probabilities $\hat{p}_{ij}(t; s = 1, \kappa\Delta t)$ were derived. The $s = 2$ probabilities were chosen completely analogously. These diffusion probabilities depend on the elements $p_{ij}(t)$ of the matrix $\mathbf{p}(t)$ from (3.18), which we obtained analytically using MATLAB's symbolic computation toolkit.

We would like to note here that this is not really a “scalable” way of developing an algorithm. but

The A+B annihilation problem

The *A+B annihilation problem*, introduced as a test problem for two implementations of the ISSA in reference [12], is given by the reaction



We consider a one-dimensional domain of length $L = 40$, with reflective boundaries at the ends, subdivided into $n = 100$ subvolumes. We set $k = 10$. Initially 1000 molecules of species A are evenly distributed across the system, while 1000 molecules of species B are located in the leftmost subvolume. Both species have diffusion coefficient $D = 5$. The maximum time-step is chosen to be consistent with error $\varepsilon = 1\%$, i.e. such that $(D/l^2)\Delta t_{max} = 0.1$ for the MSA(1), and $(D/l^2)\Delta t_{max} = 0.4$ for the MSA(2). We run ensembles of 1000 simulations to final time $t_f = 100$.

Figure 3.5 shows the population means for species A and B, vs time and space, of 1000 simulations of the A+B annihilation problem. The plots correspond to the ISSA (top), the MSA(1) (middle), and the MSA(2) (bottom). Recall that

in this problem B molecules in the left end of the system diffuse to the right and annihilate the uniformly distributed A molecules. It is plain to see that the qualitative agreement between the ensemble means of the methods is good.

To quantitatively assess the error between ISSA and MSA results, we use the Kolmogorov distance. For two cumulative distribution functions, $F_1(x)$ and $F_2(x)$ the Kolmogorov distance is defined as

$$K(F_1, F_2) = \max_{\text{all } x} |F_1(x) - F_2(x)| \quad (3.43)$$

and it has units of probability.

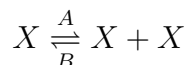
Figure 3.6 shows the Kolmogorov distance in space and time. The bottom plot gives what is known as the ISSA “self-distance” [9], which is the amount of “noise” we expect to see in an ensemble of a given size (here 1000 realizations) due to the natural fluctuations in the system. This is found by calculating the Kolmogorov distance between two ISSA ensembles of the same size which were run with different initial seeds.

It is interesting to note that the error is highest at the location of the wave-front of B first coming in contact with and annihilating A. That is where reactions are happening the fastest, in response to the molecules that have managed to diffuse the farthest. The MSA “corrals” molecules closer to their subvolume of

origin, introducing an error in the location of the molecules. But it also introduces another error by decoupling reaction from diffusion in a way that makes reactions happen later than they would by the ISSA. In the ISSA, molecules can move into a neighboring subvolume and begin being considered as reaction partners to the other molecules in that subvolume much earlier than in the MSA, according to which molecule transfers between subvolumes are lumped together into groups of preferably more than 10 (for $s = 1$) or 20 (for $s = 2$).

The Fisher problem

The *Fisher problem*, introduced as a test problem for the $S\tau$ - and $H\tau$ -leaping methods in reference [32], is given by the reversible reaction



X can represent, for example, an advantageous gene, in which case the Fisher equation models its spread. We initially place a total of X_0 molecules of species X in the left 10% of a reflective boundary system. If the reaction rate coefficients A and B are balanced with the diffusion coefficient D of X , then the system displays a wave-front which moves to the right. We use $A = 0.01$, $B = 0.00081$,

and $D = 50000$. The time-step is again chosen to achieve $\varepsilon = 1\%$ level for both MSA methods.

The ratio $r_{method} = (\text{diffusive steps (MSA) or jumps (ISSA)}) / (\text{reactions})$ is very informative. When we consider r_{ISSA} , corresponding to an ISSA simulation, we can get a sense of, on average, how much more frequent diffusive jumps are than reactions for a given problem. The MSA works by lowering the number of algorithmic steps necessary to perform diffusion, i.e. collecting many diffusive *jumps* into a single diffusive *step*. Thus the ratio $r_{MSA(s)}$ for an MSA simulation will be much smaller than for the corresponding ISSA simulation, and is approximately proportional to the speedup we expect to see when going from an ISSA simulation to an MSA simulation. Figure 3.7 gives this ratio for simulations of the Fisher system at varying initial populations \mathbf{X}_0 and subvolume number n .

The speedup observed in an MSA simulation compared to the corresponding ISSA simulation is computed as the ratio (CPU time taken for the multinomial simulation) / (CPU time taken for the ISSA simulation). Figure 3.8 shows the speedup for the Fisher problem. Note that the diffusion/reaction ratio of Figure 3.7 is a good predictor of the speedup. For the MSA(1), the diffusion/reaction ratio is approximately 10 times larger than the speedup. For the MSA(2), they differ by about a factor of 20. (The 20:10 ratio between the MSA(1) and MSA(2) is exactly as expected, since the MSA(2) requires the generation of twice as many

binomial samples). This means that the computational cost of performing a single multinomial diffusion step is approximately equal to the computational cost of performing 10 or 20 ISSA diffusive jumps. Thus we expect to see an improvement in performance due to using the MSA in cases where diffusive jumps outnumber reactions by more than an order of magnitude.

It is straightforward to compute the Kolmogorov distance between the ensemble distributions of a given species, in a given subvolume, at a given time. However, it is often the case that we need the distance between the ensemble distributions of a given species, at a given time, but over the entire spatial domain. In this case the random variable is a vector with as many elements as there are subvolumes. For this purpose we “average” the Kolmogorov distance over the spatial domain according to the formula

$$\langle K(F_1(\mathbf{x}), F_2(\mathbf{x})) \rangle_n = \frac{1}{n} \sum_{i=1}^n K(F_1(x_i), F_2(x_i)) \quad (3.44)$$

where n is the number of subvolumes. This “average” Kolmogorov distance satisfies two desirable properties: first, it has units of probability; second, it can be used for comparing across results on the same system with a different spatial discretization.

In Figure 3.9 we plot the space-averaged Kolmogorov distance (Eq. 3.44) for the ISSA (bottom), the MSA(1) (top), and the MSA(2) (middle). The simulations are of the Fisher problem, for increasingly fine discretization (i.e. increasing number of subvolumes n) and initial population density X_0 . We note that the error of the MSA(1) is approximately twice that of the MSA(2). We also note that, although the speedup from using the MSA is monotonically increasing as the number of subvolumes and the initial population increases (Figure 3.8), the space-averaged error presents no such monotonic behavior. In fact, based on the top plot, corresponding to the $s = 1$ method, one could argue that the error increases up to a point, and then shows a downward trend. This inflection point, the peak of the error curves, appears to be correlated with a population density per subvolume of about 50-100 molecules, for the $s = 1$ method.

3.3.4 Reaction-diffusion in two dimensions

We have also implemented the MSA(1) for two dimensional systems. The MSA(2) is considerably more complicated than the MSA(1) in two dimensions, so we did not implement it.

The test problem we used was a two-dimensional version of the $A + B$ annihilation problem. We considered a system with 30×30 subvolumes of side length $l = 0.04$, reaction rate $k = 10$, and the diffusion coefficients of A and B taken to be

$D = 2$. The time-step was chosen to satisfy the usual error level of $\varepsilon = 1\%$. The system was initialized with 9000 uniformly distributed A molecules, and 9000 B molecules placed in the lower left subvolume. Ensembles of 500 simulations were run to final time $t_f = 0.2$.

Figure 3.10 shows the qualitative agreement in the mean of the ISSA and MSA(1) ensembles at the final time. Figure 3.11 shows the (non-space averaged) Kolmogorov distance, a measure of error in the top plot and noise in the bottom plot, also at the final time.

3.3.5 Discussion

We have introduced a new method for efficient approximate stochastic simulation of reaction-diffusion problems. Where diffusion alone is concerned, the multinomial method has two sources of error: (a) the error from the truncation of the diffusion radius of molecules (step one), and (b) the error from the approximation of transition probabilities for systems of arbitrary size by the transition probabilities for systems of finite size (step three). The second source of error is negligible compared to the first, which is $O((\kappa\Delta t)^2)$.

When coupled with reactions, the multinomial method yields the Multinomial Simulation Algorithm. The MSA has an additional source of error, which is similar to that observed in τ -leaping methods. Like τ -leaping methods, the

MSA assumes that for specific intervals of time, while diffusion is, in fact, still occurring, the propensities of reactions are not changing. This is clearly an approximation. While τ -leaping methods constrain the size of their time-step via the “leap condition” in a way that ensures that the error from this approximation is below a certain level, the MSA has no such condition. In fact, the MSA’s computational efficiency hinges on leaping over as many diffusive transfers as possible. If those transfers are occurring in a system near diffusional equilibrium, the efficiency comes at no cost in accuracy. If, however, the diffusive transfers are contributing to the smoothing out of a sharp gradient, whose species can participate in reactions, then the MSA will incur an error from the assumption that, between the time when the reaction propensity is calculated and the time, location, and type of the next reaction are decided, and the time at which that reaction fires, reaction propensities have not changed. The $A + B$ annihilation problem and the Fisher problem were chosen because they represent this most challenging scenario for the MSA.

The MSA is efficient in situations where diffusive transfers substantially outnumber reaction events. The likelihood of this condition being satisfied can be easily assessed by comparing the magnitudes of the total diffusion propensity and the total reaction propensity. This simple criterion can serve as a reliable indicator for when the MSA should be used in an adaptive MSA-ISSA code.

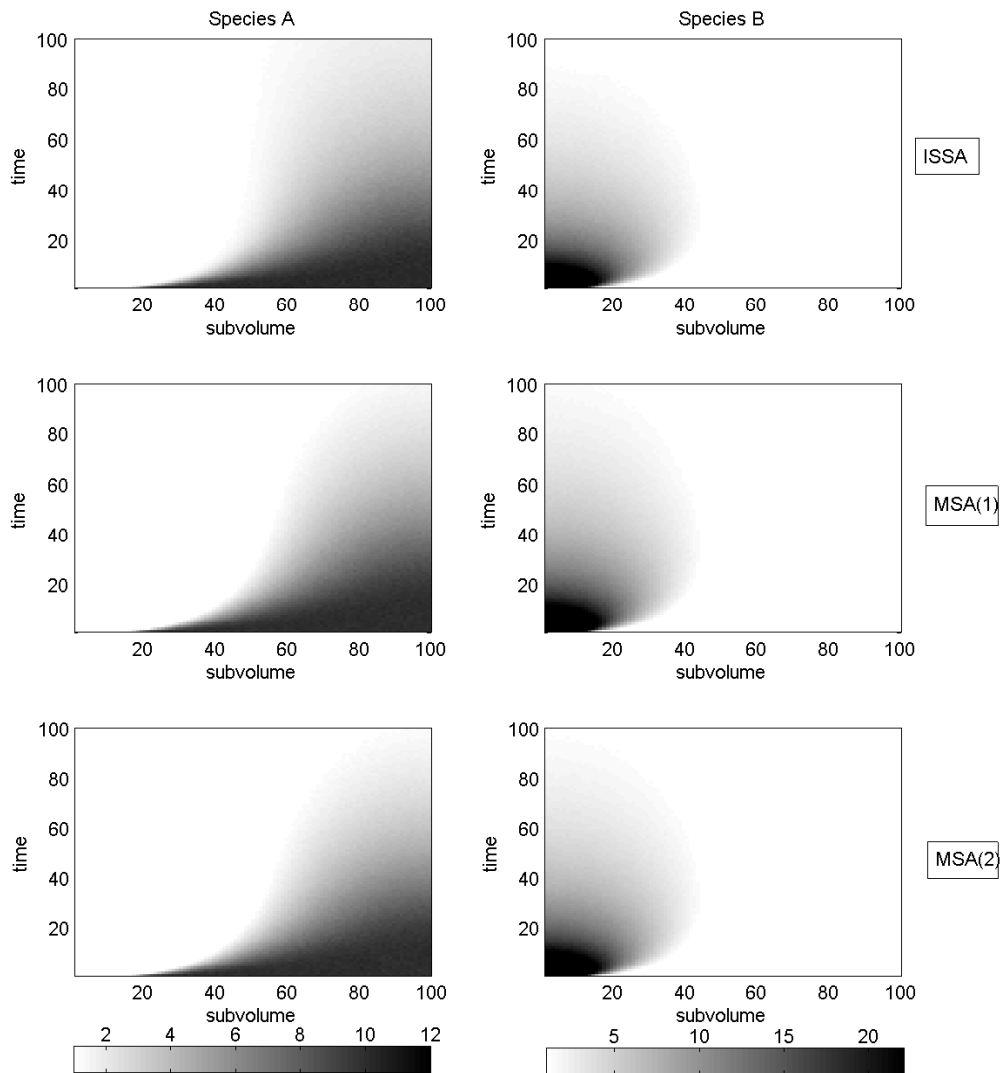


Figure 3.5: Mean population of 1000 runs vs time for the Annihilation problem. The methods are, from top to bottom, the ISSA, MSA(1), and MSA(2).

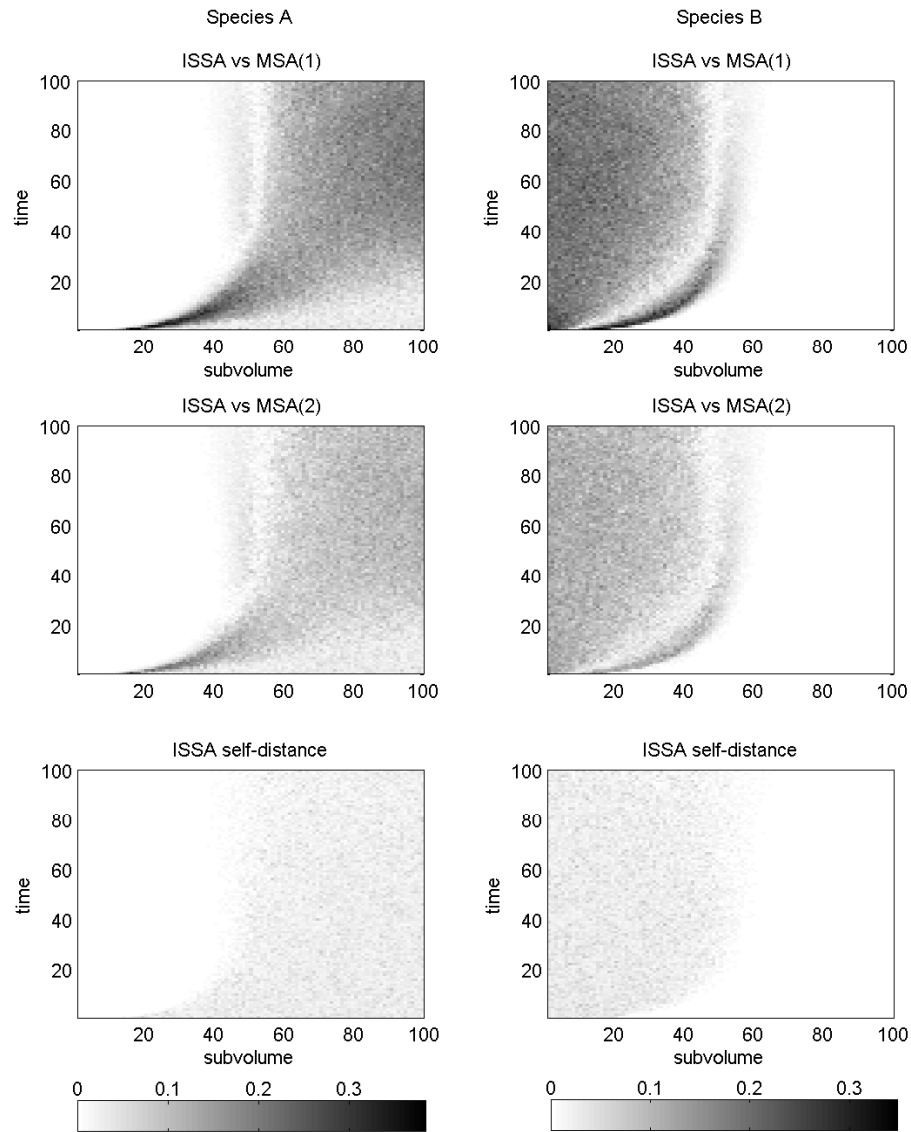


Figure 3.6: The Kolmogorov distance (units of probability) between ISSA and MSA(1) (top), ISSA and MSA(2) (middle), and the ISSA “self distance” (bottom), for the Annihilation problem.

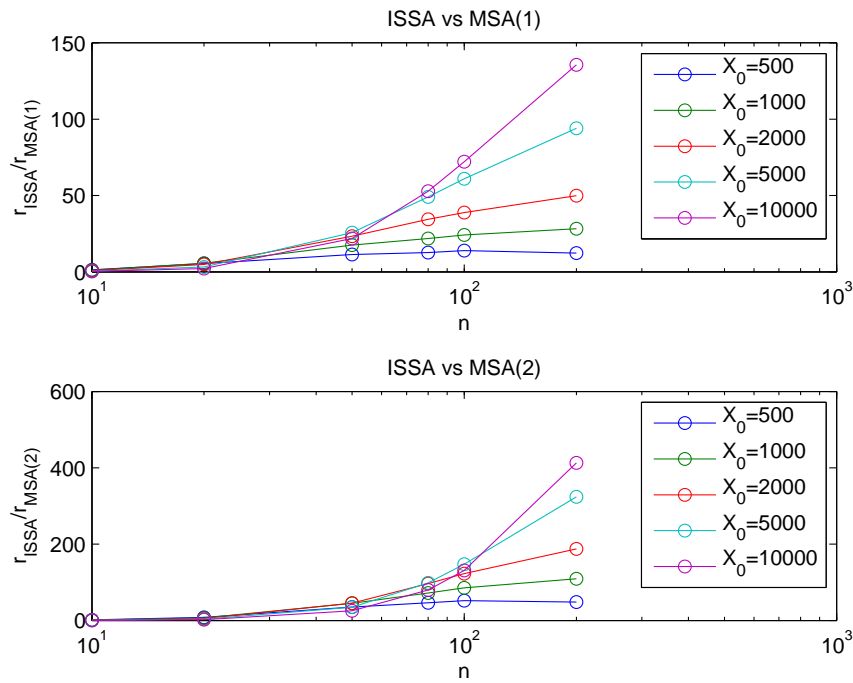


Figure 3.7: The ratio $r_{ISSA}/r_{MSA(s)}$, where $r_{method} = (\text{diffusive steps or jumps})/(\text{reactions})$, for the Fisher problem. The top plot gives $r_{ISSA}/r_{MSA(1)}$, while the bottom plot gives $r_{ISSA}/r_{MSA(2)}$. We vary the initial population X_0 and number of subvolumes n .

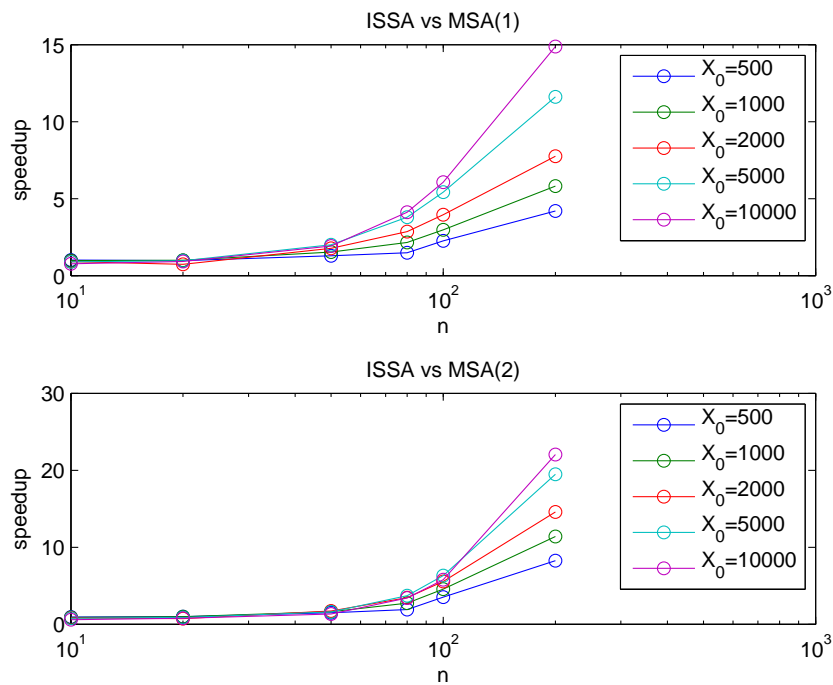


Figure 3.8: Speedup over the ISSA of MSA(1) (top) and MSA(2) (bottom), for the Fisher problem.

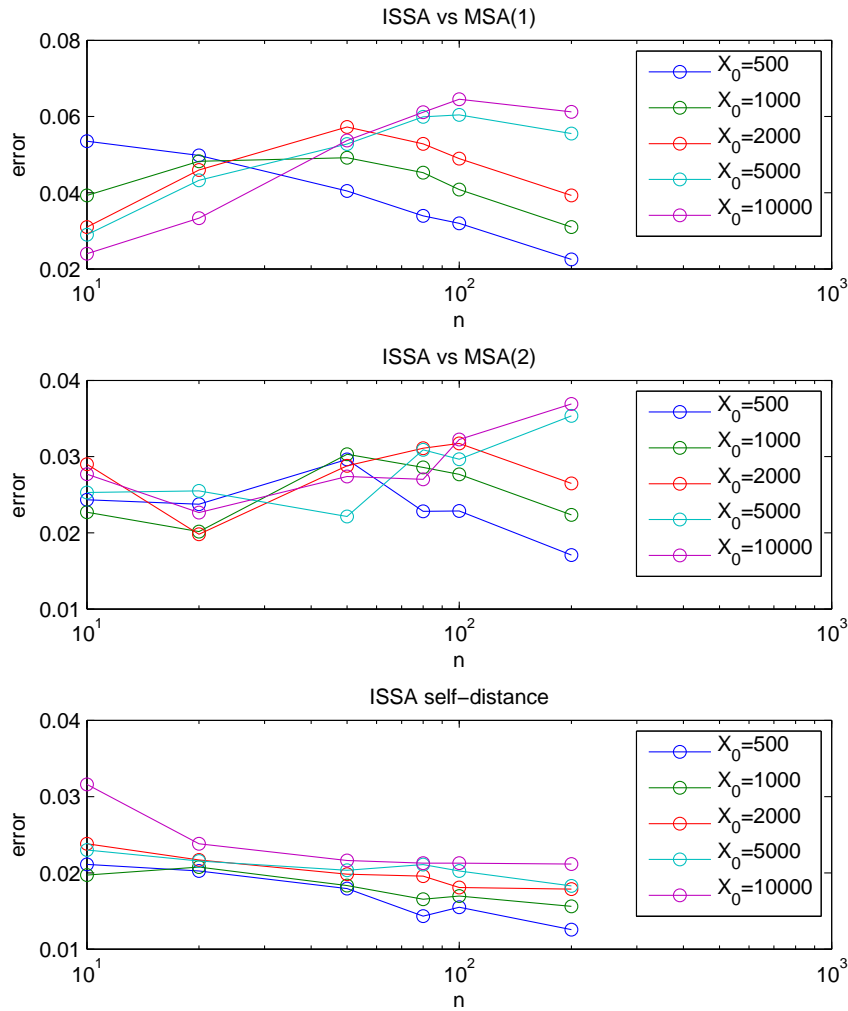


Figure 3.9: The space-averaged Kolmogorov distance between ISSA and MSA(1) (top), ISSA and MSA(2) (middle), and the ISSA “self distance” (bottom), for the Fisher problem. These results are based on ensembles of size 1000, varying the number of subvolumes n and initial population X_0 .

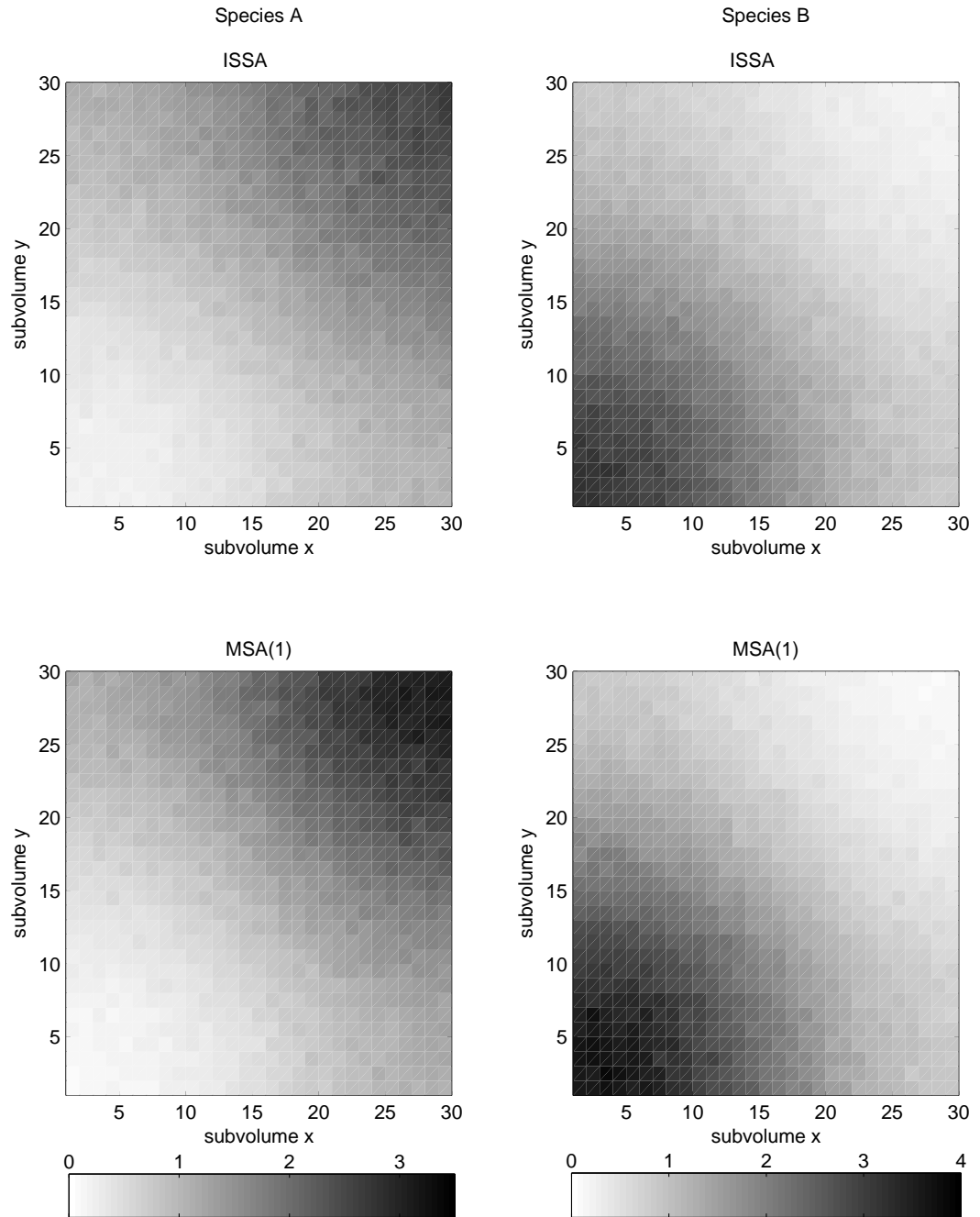


Figure 3.10: Plots of the mean of ensembles of 500 hundred simulations of the Annihilation problem in two dimensions at $t_f = 0.2$, obtained via the ISSA (top) and the two-dimensional MSA(1) (bottom). Recall that species A is uniformly distributed throughout the volume, while B is injected at the lower left hand corner, and diffuses throughout the volume.

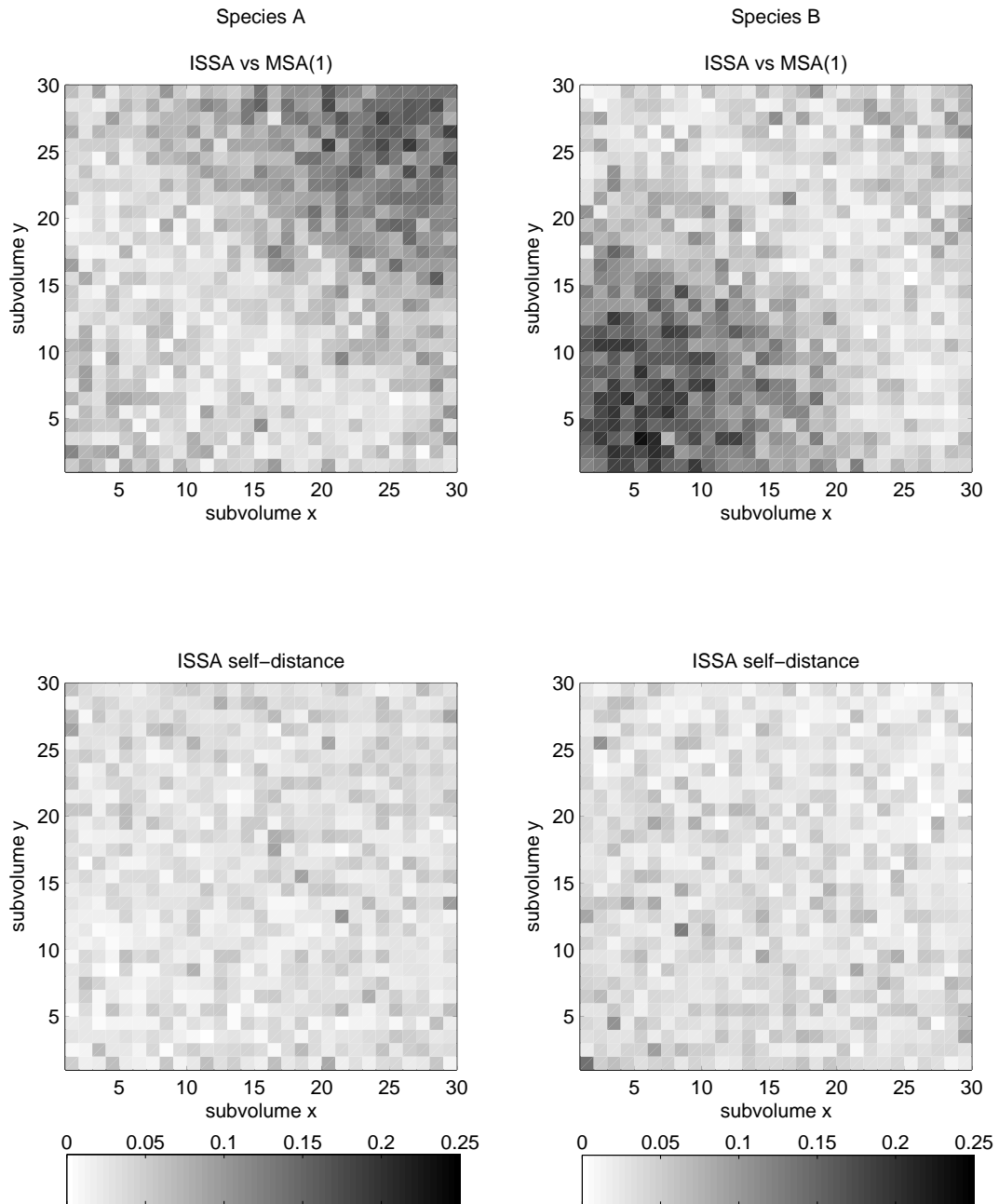


Figure 3.11: Plots of the (non-space averaged) Kolmogorov distance between ensembles of size 500, for the Annihilation problem in two dimensions. The top plot gives the error between the ISSA and MSA(1), while the bottom plot gives the ISSA self-distance.

Bibliography

- [1] L. W. Anacker and R. Kopelman. Fractal chemical kinetics: Simulations and experiments. *J. Chem. Phys.*, 81(12):6402–6403, 1984.
- [2] P. Argyrakis and R. Kopelman. Stirring in chemical reactions. *J. Phys. Chem.*, 93:225–229, 1989.
- [3] A. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics*, 149(4):1633–1648, 1998.
- [4] F. Baras and M. M. Mansour. Reaction-diffusion master equation: A comparison with microscopic simulations. *Phys. Rev. E*, 54:61396148, 1996.
- [5] N. N. Batada, L. A. Shepp, and D. O. Siegmund. Stochastic model of protein-protein interaction: Why signaling proteins need to be colocalized. *PNAS*, 101(17):6445–6449, 2004.
- [6] D. Bernstein. Simulating mesoscopic reaction-diffusion systems using the gillespie algorithm. *Phys. Rev. E*, 71:041103, 2005.
- [7] G. A. Bird. *Molecular Gas Dynamics*. Oxford University Press, Clarendon, Oxford, 1976.
- [8] Y. Cao, D. Gillespie, and L. Petzold. Avoiding negative populations in explicit tau leaping. *J. Chem. Phys.*, 123:054104, 2005.
- [9] Y. Cao and L. R. Petzold. Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems. *J. Comp. Phys.*, 212:6–24, 2006.
- [10] B. Chopard, A. Masselot, and M. Droz. Multiparticle lattice gas model for a fluid: Application to ballistic annihilation. *Phys. Rev. Lett.*, 81(9):1845–1848, Aug 1998.

- [11] J. Elf, A. Doncic, and M. Ehrenberg. Mesoscopic reaction-diffusion in intracellular signaling. In *Proceedings for SPIE's "First International Symposium on fluctuations and noise"*, volume 5110, pages 114–124, May 2003.
- [12] J. Elf and M. Ehrenberg. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *IEE Systems Biology*, 1(2):230–236, 2004.
- [13] R. J. Ellis. Macromolecular crowding: obvious but underappreciated. *TRENDS in Biochemical Sciences*, 26(10):597–604, 2001.
- [14] S. Engblom, L. Ferm, A. Hellander, and P. Lotstedt. Simulation of stochastic reaction-diffusion processes on unstructured meshes. Technical Report 012, University of Uppsala, 2008.
- [15] C. Gardiner. *Handbook of Stochastic Methods*. Springer-Verlag, Berlin, 1985.
- [16] C. Gardiner, K. McNeil, D. Walls, and I. Matheson. Correlations in stochastic theories of chemical reactions. *J. Stat. Phys.*, 14(4), 1976.
- [17] M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A.*, 104(9):1876–1889, 2000.
- [18] D. T. Gillespie. The monte carlo method of evaluating integrals. *Naval Weapons Center Technical Publication (NWC TP)*, 5714, Feb 1975.
- [19] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comp. Phys.*, 22:403, 1976.
- [20] D. T. Gillespie. *J. Stat. Phys.*, 16:311, 1977.
- [21] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361, 1977.
- [22] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188:404–425, 1992.
- [23] D. T. Gillespie, S. Lampoudi, and L. R. Petzold. Effect of reactant size on discrete stochastic chemical kinetics. *J. Chem. Phys.*, 126(3):034302, 2007.
- [24] A. M. L. J. M. Sancho, A. H. Romero and K. Lindenberg. *J. Phys.: Condensed Matter*, 19:065108, 2007.

- [25] N. V. Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, 1997.
- [26] R. Kopelman. Rate processes on fractals: Theory, simulations, experiments. *J. Stat. Phys.*, 42:185–200, 1986.
- [27] L.A.Bunimovich. On the ergodic properties of nowhere dispersing billiards. *Commun. Math. Phys.*, 65:295–312, 1979.
- [28] K. Lindenberg, P. Argyrakis, and R. Kopelman. Reaction-diffusion model for a+a reaction. *J. Phys. Chem.*, 99:7542–7556, 1995.
- [29] G. Nicolis and I. Prigogine. *Self-Organization in Nonequilibrium Systems*. Wiley-Interscience, 1977.
- [30] M. Rathinam and H. E. Samad. Reversible-equivalent-monomolecular tau: A leaping method for "small number and stiff" stochastic chemical systems. *J. Comp. Phys.*, 224:897–923, 2007.
- [31] J. V. Rodriguez, J. A. Kaandorp, M. Dobrzynski, and J. G. Blom. Spatial stochastic modelling of the phosphoenolpyruvate-dependent phosphotransferase (PTS) pathway in Escherichia coli. *Bioinformatics*, 22(15):1895–1901, 2006.
- [32] D. Rossinelli, B. Bayati, and P. Koumoutsakos. Accelerated stochastic and hybrid methods for spatial simulations of reaction-diffusion systems. *Chem. Phys. Lett.*, 451:136–140, 2008.
- [33] J. C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Wadsworth & Brooks/Cole Advanced Books and Software, 1989.
- [34] L. Tonks. The complete equation of state of one, two and three-dimensional gases of hard elastic spheres. *Phys. Rev. E*, 50:955–963, 1936.
- [35] S. Torquato. Nearest-neighbor statistics for packings of hard spheres and disks. *Phys. Rev. E*, 51:3170–3182, 1995.

Bibliography

Appendices

Appendix A

Uniform distribution of hard rods on a fixed interval

A.1 The distribution of centers

Let N rods, each of length l , be randomly and uniformly distributed over the x -axis interval $(0, L)$ subject to the condition that no rods overlap (the rods are thus “hard”). Let the rods be numbered from left to right, and let x_n , for $n = 1, \dots, N$, locate the center of the n^{th} rod (see Fig. A.1). Let $P(x_1, \dots, x_N)$ be the joint probability density function of these N position variables. We shall prove here that

$$P(x_1, \dots, x_N) = \prod_{n=1}^N P_{n|n-1}(x_n | x_{n-1}) \quad (\text{A.1})$$

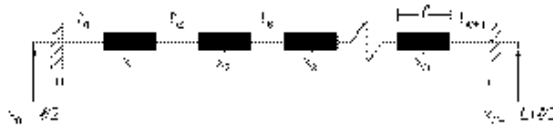


Figure A.1: Diagram of hard rods in a finite interval
Schematic of N hard rods, each of length l , distributed inside the x -axis interval $[0, L]$. The variables x_n locate the centers of the rods, and the variables ξ_n measure the gaps between the rods.

where the probability density function of x_n conditioned on x_{n-1} is given by

$$P_{n|n-1}(x_n | x_{n-1}) = \frac{N - n + 1}{(L_{n-1} - x_{n-1})^{N-n+1}} (L_n - x_n)^{N-n} \\ (x_{n-1} + l < x_n < L_n; n = 1, \dots, N) \quad (\text{A.2})$$

Here the parameter x_0 is defined by

$$x_0 \equiv -\frac{1}{2}l \quad (\text{A.3})$$

and the $N + 1$ parameters L_0, \dots, L_N are defined by

$$L_n \equiv L - \frac{1}{2}l - (N - n)l \quad (n = 0, \dots, N) \quad (\text{A.4})$$

or equivalently by the recursion relation

$$L_n \equiv \begin{cases} L - \frac{1}{2}l, & \text{for } n = N, \\ L_{n+1} - l, & \text{for } n = N - 1, N - 2, \dots, 0. \end{cases} \quad (\text{A.5})$$

A close examination of Eq. (A.5) in conjunction with Fig. A.1 will reveal that, for each $n = 1, \dots, N$, L_n is the upper limit on x_n , i.e., the value x_n would have if rods n through N were all shoved as far right in the interval $(0, L)$ as possible.

A systematic derivation of the above result was given in Ref. [18] (although there the rods were numbered from right to left instead of from left to right). Since that derivation is lengthy and formal, we shall give here a derivation that is shorter and more transparent. We begin by noting that any joint probability density function $P(x_1, \dots, x_N)$ can always be written in the conditioned form

$$P(x_1, \dots, x_N) = \prod_{n=1}^N P_n(x_n | x_{n-1}, x_{n-2}, \dots, x_1)$$

where $P_n(x_n | x_{n-1}, \dots, x_1)dx_n$ gives the probability that the n^{th} variable will have a value in $[x_n, x_n + dx_n)$ given that the $n - 1$ lower indexed variables have the respective values x_{n-1}, \dots, x_1 . But it is physically clear in our problem (see Fig. A.1) that once the position of rod $n - 1$ has been fixed, the positions of all the lower indexed rods supply no additional information about the possible position of rod n . Thus, for our problem, the above general conditioning formula reduces to the simpler form (A.1).

To establish Eq. (A.2), we shall derive an expression for the probability

$$P_{n|n-1}(x_n|x_{n-1})dx_n$$

that the center of the n^{th} rod from the left will lie in the infinitesimal interval $[x_n, x_n + dx_n)$, given that the center of the $(n - 1)^{\text{th}}$ rod from the left is at x_{n-1} . The “given” here implies that the n^{th} rod from the left, which we label n , will have its center somewhere in the interval $[x_{n-1} + l, L_n)$, and further that the first rod to its right, which we will label $n + 1$, will have its center somewhere in the interval $[x_{n-1} + 2l, L_{n+1})$. In general, the k^{th} rod to the right of rod n , which we label $n + k$, will have its center somewhere in the interval $[x_{n-1} + (1 + k)l, L_{n+k})$. Now we make the observation that, although the boundaries of these n confining intervals are all different, their lengths are all equal to $L_{n-1} - x_{n-1}$: Thus, the length of the interval containing the center of rod n is

$$L_n - (x_{n-1} + l) = (L_n - l) - x_{n-1} = L_{n-1} - x_{n-1}$$

where the last step invokes the second of relations (A.5). And the length of the interval for the center of rod $n + k$ for any $k = 1, \dots, N - n$ is

$$L_{n+k} - [x_{n-1} + (1 + k)l] = [L_{(n-1)+(1+k)} - (1 + k)l] - x_{n-1}$$

But since the second of relations (A.5) implies that $L_n = L_{n+k} - kl$, the quantity on the right in brackets is L_{n-1} , so the interval length is again $L_{n-1} - x_{n-1}$. So we see that, once the value of x_{n-1} is known, the values of every one of the higher-indexed variables x_n, \dots, x_N will be restricted to an interval of length $L_{n-1} - x_{n-1}$.

To compute the probability that the center of rod n will lie inside the infinitesimal interval dx_n at x_n , given that rod $n - 1$ has its center at x_{n-1} , we first choose at random any one of the $N - n + 1$ rods to the right of rod $n - 1$, without regard for its position or ordering. All we know about this randomly chosen rod is that its center must be confined to some interval of length $L_{n-1} - x_{n-1}$; therefore, the probability that its center will actually lie in a particular interval of length dx_n is just the ratio of those two interval lengths,

$$\frac{dx_n}{L_{n-1} - x_{n-1}}$$

If this dx_n subinterval is located at the value x_n , then our randomly chosen rod will be the left-most of the right-most $N - n + 1$ rods if and only if all the $N - n$ remaining rods lie to the right of x_n . That in turn means, by the result developed in the preceding paragraph, that the center of each one of those $N - n$ other rods

must be restricted to an interval of length $L_n - x_n$. Since all we know about any one of those other rods is that it is confined to an interval of length $L_{n-1} - x_{n-1}$, the probability that it will actually lie inside a smaller interval of length $L_n - x_n$ is the ratio

$$\frac{L_n - x_n}{L_{n-1} - x_{n-1}}$$

We conclude, then, that the net probability for the center of the chosen rod to lie in the infinitesimal interval dx_n at x_n and for all the $N - n$ other rods to lie to the right of x_n (thus making the chosen rod the left-most of those $N - n + 1$ rods) is

$$\frac{dx_n}{L_{n-1} - x_{n-1}} \times \left(\frac{L_n - x_n}{L_{n-1} - x_{n-1}} \right)^{N-n} = \frac{dx_n}{(L_{n-1} - x_{n-1})^{N-n+1}} (L_n - x_n)^{N-n}$$

Finally, since there are $N - n + 1$ ways in which we could have chosen that first rod, it follows that the probability $P_n(x_n | x_{n-1})dx_n$ must be equal to $N - n + 1$ times the above probability. This establishes the result (A.2).

A.2 The gap distribution

In the problem just addressed, number the gaps between adjacent rods so that gap n is bounded on the right by rod n and/or on the left by rod $n - 1$. The length ξ_n of gap n will thus be given by (see Fig. A.1)

$$\xi_n = x_n - x_{n-1} - l \quad (n = 1, \dots, N + 1) \quad (\text{A.6})$$

Here we have recalled definition (A.3) of x_0 , and additionally defined

$$x_{N+1} \equiv L + \frac{1}{2}l \quad (\text{A.7})$$

Counting the two gaps adjacent to the interval boundaries, there are a total of $N + 1$ gaps; however, only N of these gaps are algebraically independent, since the sum of all the gaps must be the unoccupied length $L - Nl$. Let us calculate the probability density function of the length of gap 1, the gap between the origin and the left edge of rod 1. From Eqs. (A.6) and (A.3) we see that the size of that gap is related to x_1 by (see Fig. A.1)

$$\xi_1 = x_1 - (-\frac{1}{2}l) - l = x_1 - \frac{1}{2}l$$

The density function of x_1 is given by Eq. (A.2) with $n = 1$:

$$\begin{aligned} P_{1|0}(x_1 | x_0) &= \frac{N}{(L_0 - x_0)^N} (L_1 - x_1)^{N-1} \quad (x_0 + l < x_1 < L_1) \\ &= \frac{N}{(L - Nl)^N} (L - Nl + \frac{1}{2}l - x_1)^{N-1} \quad (\frac{1}{2}l < x_1 < L - Nl + \frac{1}{2}l) \end{aligned}$$

where the last step has used the definitions (A.3) and (A.4). Note that since x_0 is a constant, this is really the unconditioned density function of the variable x_1 . The unconditioned density function of the variable ξ_1 can therefore be computed by applying the change-of-variable rule in probability theory:

$$\hat{P}_{1|0}(\xi_1 | x_0) = P_{1|0}(x_1 | x_0) \left| \frac{dx_1}{d\xi_1} \right|$$

Here, x_1 on the right is now understood to be the function $x_1(\xi_1) = \xi_1 + \frac{1}{2}l$. This immediately yields the result

$$\hat{P}_{1|0}(\xi_1 | x_0) = \frac{N}{(L - Nl)^N} (L - Nl - \xi_1)^{N-1} \quad (0 < \xi_1 < L - Nl) \quad (\text{A.8})$$

It turns out that all the other gaps $\xi_2, \xi_3, \dots, \xi_N$ have this same unconditioned density function. We shall prove this explicitly for the gap variable ξ_2 , and then consider the result for the remaining gaps to be obvious on grounds of symmetry.

To calculate the unconditioned density function for ξ_2 , we need to first compute the unconditioned joint density function for x_1 and x_2 :

$$\begin{aligned} P_{1,2|0}(x_1, x_2 | x_0) &= P_{1|0}(x_1 | x_0) P_{2|1}(x_2 | x_1) \\ &= \frac{N(N-1)}{(L_0 - x_0)^N} (L_2 - x_2)^{N-2} \\ &= \frac{N(N-1)}{(L - Nl)^N} (L - Nl + \frac{3}{2}l - x_2)^{N-2} \end{aligned}$$

The second step above follows from Eq. (A.2), and the last step follows from the definitions (A.3) and (A.4). The domain of definition of this joint density function (outside of which the function vanishes) is

$$\begin{aligned} (x_0 + l < x_1 < L_1; x_1 + l < x_2 < L_2) = \\ (\frac{1}{2}l < x_1 < L - Nl + \frac{1}{2}l; x_1 + l < x_2 < L - Nl + \frac{3}{2}l) \end{aligned}$$

where the last step follows from the definitions (A.3) and (A.4). Now we make the transformation of variables $(x_1, x_2) \rightarrow (x_1, \xi_2)$, where in accordance with Eq. (A.6), $\xi_2 = x_2 - x_1 - l$. The joint density function of the new pair of variables is

$$\hat{P}_{1,2|0}(x_1, \xi_2 | x_0) = P_{1,2|0}(x_1, x_2 | x_0) \left| \frac{\partial(x_1, x_2)}{\partial(x_1, \xi_2)} \right|$$

where x_2 on the right is now regarded as the function $x_2(x_1, \xi_2) = \xi_2 + x_1 + l$. The Jacobian here is easily shown to be unity, so we have

$$\hat{P}_{1,2|0}(x_1, \xi_2 | x_0) = \frac{N(N-1)}{(L-Nl)^N} (L - Nl + \frac{1}{2}l - \xi_2 - x_1)^{N-2}$$

The domain of definition of this function is found by replacing x_2 in the previous domain formula with the function $x_2(x_1, \xi_2)$. That gives

$$\begin{aligned} & (\frac{1}{2}l < x_1 < L - Nl + \frac{1}{2}l; 0 < \xi_2 < L - Nl + \frac{1}{2}l - x_1) \\ & = (0 < \xi_2 < L - Nl; \frac{1}{2}l < x_1 < L - Nl + \frac{1}{2}l - \xi_2) \end{aligned}$$

where the last step follows by interchanging the bounding order of the variables x_1 and ξ_2 . Now we can compute the unconditioned density function of ξ_2 by integrating out the variable x_1 :

$$\begin{aligned} \hat{P}_{2|0}(\xi_2 | x_0) &= \int_{\frac{1}{2}l}^{L-Nl+\frac{1}{2}l-\xi_2} \hat{P}_{1,2|0}(x_1, \xi_2 | x_0) dx_1 \\ &= \frac{N(N-1)}{(L-Nl)^N} \int_{\frac{1}{2}l}^{L-Nl+\frac{1}{2}l-\xi_2} (L - Nl + \frac{1}{2}l - \xi_2 - x_1)^{N-2} dx_1 \\ &= \frac{N(N-1)}{(L-Nl)^N} \int_0^{L-Nl-\xi_2} u^{N-2} du \\ \hat{P}_{2|0}(\xi_2 | x_0) &= \frac{N}{(L-Nl)^N} (L - Nl - \xi_2)^{N-1} \quad (0 < \xi_2 < L - Nl) \quad (\text{A.9}) \end{aligned}$$

Comparing the results (A.8) and (A.9) for ξ_1 and ξ_2 , we conclude by symmetry that the unconditioned density function of any gap length between adjacent rods is

$$P_{\text{gap}}(\xi) = \frac{N}{(L-Nl)^N} (L - Nl - \xi)^{N-1} \quad (0 < \xi < L - Nl) \quad (\text{A.10})$$

A.3 Monte Carlo placement of the rods

Equations (A.1) and (A.2) allow us to construct a Monte Carlo algorithm for dealing out the rods from left to right in the interval $0 \leq x \leq L$, successively specifying values for x_1, x_2, \dots, x_N , in such a way that the resulting configuration will be an unbiased sample of the uniform non-overlapping distribution [18]. The first step in constructing this generating algorithm is to compute the cumulative distribution function $F_{n|n-1}(x_n | x_{n-1})$ corresponding to the probability density function (A.2):

$$\begin{aligned} F_{n|n-1}(x_n | x_{n-1}) &\equiv \int_{x_{n-1}+l}^{x_n} P_{n|n-1}(x'_n | x_{n-1}) dx'_n \\ &= \frac{N-n+1}{(L_{n-1} - x_{n-1})^{N-n+1}} \int_{x_{n-1}+l}^{x_n} (L_n - x'_n)^{N-n} dx'_n \quad (\text{A.11}) \end{aligned}$$

The integration is easily performed using the variable transformation $u = L_n - x'_n$, and recalling from Eq. (A.5) that $L_n - l = L_{n-1}$. The result is

$$F_{n|n-1}(x_n | x_{n-1}) = 1 - \left(\frac{L_n - x_n}{L_{n-1} - x_{n-1}} \right)^{N-n+1} \quad (\text{A.12})$$

Now we apply the standard Monte Carlo inversion procedure to generate a set of values for x_1, x_2, \dots, x_N according to the joint probability density function (A.2): We pick N random numbers r_1, r_2, \dots, r_N from the unit-interval uniform distribution, and then solve the equation

$$F_{n|n-1}(x_n | x_{n-1}) = r_n \quad (\text{A.13})$$

for x_n , successively for $n = 1, 2, \dots, N$. Substituting Eq. (A.12) into Eq. (A.13), and for convenience replacing r_n by $1 - r_n$ (which is allowed since the latter will also be uniformly distributed in the unit interval), we get

$$\left(\frac{L_n - x_n}{L_{n-1} - x_{n-1}} \right)^{N-n+1} = r_n$$

Solving this for x_n gives us the generating formula:[18]

$$x_n = L_n - (L_{n-1} - x_{n-1}) r_n^{1/(N-n+1)} \quad (n = 1, 2, \dots, N) \quad (\text{A.14})$$

which is to be applied successively for increasing n . As before, $x_0 = -\frac{1}{2}l$, and the set of constants $\{L_0, L_1, \dots, L_N\}$ is as given in Eqs. (A.4).

It is shown in Ref. [18] that the alternative generating procedure of independently placing each rod center uniformly at random inside the interval $[\frac{1}{2}l, L - \frac{1}{2}l]$ and then accepting the final set if and only if no overlap occurs, has an associated acceptance probability of $[(L - Nl)/(L - a)]^N$. For $N = 100$ rods of length $l = 1$ on an interval of length $L = 200$, this probability of acceptance is only 1.3×10^{-30} . By contrast, the procedure (A.14) gives an acceptable configuration every time. The reason is that this procedure knows that the n^{th} rod laid down is to be the n^{th} rod from the left, so it biases the selection appropriately to the left, taking account of the already generated location of the $(n - 1)^{\text{th}}$ rod.