DIFFUSION IN FUZZY LATTICE SYSTEMS: EXPLORING THE ANOMALOUS REGIME, CONNECTING THE STEADY-STATE, AND FAT-TAILED DISTRIBUTIONS

NICHOLAS ILOW

Thesis submitted to The University of Ottawa
In partial fulfillment of the requirements
for the Master of Science in Physics

Ottawa-Carleton Institute of Physics
Department of Physics
Faculty of Science
University of Ottawa

SUMMARY

Diffusion and random walks have been studied for more than 100 years. However, there are still details in the methodology that are overlooked, and more information can be extracted from the typical data that is studied.

In this thesis, I simulate random walks on two dimensional lattices with immobile obstacles configured in a variety of ways: periodic, random, and "Fuzzy" (a cross intermediate state of disorder between periodic and random). The primary goal is to develop a deeper understanding of "Fuzzy" systems by designing different ways of generating tunable disorder. An example of this is the universal F_z parameter that we developed to unify the natural disorder parameters of the various disorder generation methods we developed.

Often times the importance of analysing the transient/anomalous regime with more precision and consistency is overlooked. In our work, we expand on random walk dynamics by applying non-standard probabilities, and justify our choice analytically and through a comparison of results. Furthermore we discuss how the transient regime should be analyzed so that there is consistency in the field.

Other than discussing semantics of algorithms and analysis, we study the connection between the transient regime and the steady-state. We introduce two measures of the width of the transient/anomalous regime, and compare them to the crossover time. Using the width of the transient/anomalous regime we are able to provide an estimate of the steady-state diffusion coefficient without access to the steady-state simulation data.

SOMMAIRE

La diffusion et les marches aléatoires sont étudiées depuis plus de 100 ans. Cependant, certains détails de la méthodologie sont encore négligés, et plus d'informations peuvent être extraites des données typiques qui sont produites.

Dans cette thèse, je simule des marches aléatoires sur des treillis bidimensionnels avec des obstacles immobiles configurés de différentes manières : périodique, aléatoire, et "Fuzzy" (un état de désordre intermédiaire croisé entre périodique et aléatoire). L'objectif principal est de développer une compréhension plus approfondie des systèmes "Fuzzy" en concevant différentes manières de générer un désordre synthonisable. Un exemple de ceci est le paramètre universel F_z que nous avons développé pour unifier les paramètres de désordre naturel des différentes méthodes de génération de désordre que nous avons développées.

On néglige souvent l'importance d'analyser le régime transitoire/anormal avec plus de précision et de cohérence. Dans notre travail, nous développons la dynamique de la marche aléatoire en appliquant des probabilités non standard, et nous justifions notre choix analytiquement et par une comparaison des résultats. En outre, nous discutons de la manière dont le régime transitoire devrait être analysé afin d'assurer la cohérence dans le domaine.

En plus de discuter de la sémantique des algorithmes et de l'analyse, nous étudions le lien entre le régime transitoire et l'état stationnaire. Nous introduisons deux mesures de la largeur du régime transitoire/anormal, et nous les comparons au temps de croisement. En utilisant la largeur du régime transitoire/anormal, nous sommes en mesure de fournir une estimation du coefficient de diffusion à l'état stationnaire sans avoir accès aux données de simulation à temps longs.

STATEMENT OF ORIGINALITY

I declare that the work presented in this thesis is novel, of my own construction and encompasses what I have done throughout my Masters degree under Prof. Gary Slater. This thesis consists of an introductory section (Chapter 1), a section on our methodology (Chapter 2) and three manuscripts (Chapters 3, 4, 5) one of which is under review (Chapter 4) and the others are soon to be submitted (Chapters 3, 5).

I am the primary author on two of the three manuscripts (Chapters 3 and 4), and as such the entirety of the simulations and data analysis has been done by me. Gary Slater contributed significantly in developing theories, ideas to test, improving the writing, and revising the thesis, and manuscripts themselves. In the third manuscript (Chapter 5) I contributed only the simulations and analysis contained in Section IV and the non-analytic portion of the text (where the analytic portion is my supervisors contribution); the rest of this manuscript is the work of Le Qiao.

It should also be noted that there are two ideas throughout the thesis that have been previously explored by Neo Nguiya Passi. First the idea of using non-standard probabilities in Section 2.13, however I differentiate from his work by extending the derivation to general probabilities and ultimately choosing a different set of hopping probabilities from his work. Secondly the concept of using the inflection point of the transient regime in Section 2.5 as a consistent region for fitting and analyzing our transient data is used throughout the thesis.

DEDICATION

I dedicate my thesis to my girlfriend, Taylor Gray, who helped me stay motivated throughout my degree. She always pushes me to be the best student and individual that I can be. She was why I looked into the University of Ottawa in the first place, and I could not have made a better choice.

LIST OF ABBREVIATIONS AND SYMBOLS

α	Anomalous Exponent
AR2P	Add Random to Periodic (Fuzzy System generation method)
β	Excess Diffusivity
CTRW	Continuous Time Random Walk
D_0	Free Diffusion Coefficient
D_α	Anomalous Diffusion Coefficient
D	Steady-State Diffusion Coefficient
F_z	Fuzzy Disorder Parameter
HPW	Harmonic Potential Wells (Fuzzy System generation method)
KMC	Kinetic Monte Carlo
lhs	Left Hand Side
LMC	Lattice Monte Carlo
MCMC	Markov Chain Monte Carlo
NoCR	No Cluster (Fuzzy System generation method)
PBC	Periodic Boundary Condition

ϕ	Obstacle Concentration
PTRS	Periodic Tiling of Random Subsystems (Fuzzy System generation method)
PUD	Pick Up and Drop (Fuzzy System generation method)
rhs	Right Hand Side
r^*	Crossover Length
Σ_α	Log-Log Space Anomalous Regime Width
Σ_t	Cartesian Space Anomalous Regime Width
t^*	Crossover Time

ACKNOWLEDGMENTS

I would like to acknowledge the University of Ottawa for providing me with an admission scholarship, Compute Canada for having resources available to me to increase productivity, and my supervisor Gary Slater for providing insight, teaching me, and for always helping me promptly whenever I had any questions.

CONTENTS

CHAPTER 1

INTRODUCTION

When we think of diffusion, we often think of something tangible to us: maybe you think of how your water becomes tea when you steep it with a teabag, or how smoke moves when you light an incense stick in your still room. These examples have in common a lack of impeding obstacles: the water in your mug does not have solid blocks in it, the air in your room does not have floating walls. So what happens when we do have walls or objects in the way? How would it affect our steeping tea? Taking this a step further, if we have many walls and we place them in our room we know we could change when we smell the incense and how much we smell it. What would happen to this behaviour if we place the walls randomly? How about periodically?

The incense sticks and teabags are both scenarios that are familiar with us, but what is less thought of is our cells. Diffusion is a key transport mechanism in our own body; our cells utilise it to transport important molecules within themselves, and it also plays a role in molecules crossing the cell membrane. Cells themselves have many immobile proteins within them, and if we investigate what happens on the cell membrane itself we are left with a two-dimensional plane, with some immobile and some mobile obstacles (proteins).

In Nature however, things are seldom periodic, nor are they completely random. One example are Penguins: Penguins are very territorial and thus do not allow other Penguins to settle in a close proximity to them, yet they still live in packs. The amount of territory that a Penguin will keep to itself will vary from Penguin to Penguin, resulting in something close to a periodic structure, but with small deviations. An example of randomness in Nature could be dandelions: dandelion seeds are sown by the wind, and their final resting position varies greatly depending on many factors.

The cell itself has a structure, and albeit mobile proteins are not tied to this structure directly, they are influenced by it. So we come to the conclusion that these obstacles (proteins) are spaced in a manner that is somewhere between periodic and random, which we will be calling "Fuzzy" from here on out.

Diffusion as a general topic has been extensively studied in biological systems as well as synthetic systems [19]. To put it in perspective one of the first findings of Anomalous Diffusion (diffusion that scales non-linearly with time) is from 1935, and used Single Particle Tracking) [8, 19] (the method of placing a tracer particle into a system and tracking the displacement as a function of time. The field is very active, with 100's of papers using Single Particle Tracking experimentally and random walks of tracer particles studied computationally every year [9, 15, 24, 25, 26, 32, 34]. Alongside advances in technology, and computing power becoming more accessible, computational techniques to model and investigate diffusion on a lattice are one of many ways the field continues to advance today [1, 11, 18].

The goal of this thesis is to build a deeper understanding of disorder and its effect on diffusion in systems containing immobile obstacles. We model diffusion of a point like particle traversing a two dimensional space. The details of the approach we take are found in Chapter 2. Every calculation performed throughout this thesis has been done in Python and designed to work on a single core; however, the calculations are scalable to multi core settings by scripts designed to combine data sets of like parameters. The primary exploration of disorder can be found in Chapter 3 where we generate tunable disorder in obstacle configurations using many different methods (described in Section 2.10) and compare the effects on the diffusive behaviour. Next we look at how we can further characterize the anomalous regime (the regime for which anomalous diffusion can be applied) in Chapter 4. We find that the anomalous regime offers a significant amount of information about the steady-state and discuss how to apply it. Finally, in Chapter 5 we investigate the displacement distribution functions resulting from diffusion in systems with randomly placed obstacles. We work with an empirical fitting function that is able to capture the fat tails that are a consequence of anomalous diffusion, but also able to capture the short range

Gaussian behaviour.

1.1 The Emergence of Diffusion

In the year 1827 Robert Brown began to study grains of pollen, that were filled with smaller granules, in water. He noticed that these smaller granules were in motion, and further noticed changes of form; he reported seeing contractions or changes in curvatures, and even rotation of the granules. These different forms of motion led Robert Brown to conclude that the motion was in fact not a result of currents in the water, nor the evaporation process, rather that the particles themselves were moving [3].

Louis Bachelier under the guidance of the famous mathematician Henri Poincaré would be next in discovering two key aspects of todays understanding of Brownian motion in his mathematical thesis titled "The Theory of Speculation" [2] where he would study the math behind the stock exchange. Bachelier would show us that Brownian motion has a Markovian character, i.e. prior information of the state of a system is not needed to predict the future states of the system. Secondly he would show the reflection principle, which connects the distribution of the particles max position $f(t)$ after a time $t = s$ to the distribution of the particles path for $t > s$.

In 1905 Albert Einstein approached the problem of Brownian motion using a probabilistic model [6], his goal being to explain, using the molecular-kinetic theory of heat, that microscopically small bodies suspended in a liquid will move due to the molecular motions caused by heat. Einstein began by considering a quantity of liquid in a volume V, with suspended particles retained in a smaller volume V^*, and divided from the rest of V by a semi-permeable membrane. Einstein showed that using the classical theory of thermodynamics no force would act on the membrane, however using the molecular-kinetic theory of heat that there is in fact a force, the cause of this force being the irregular movements of the suspended particles due to the movement of the liquid molecules. In more common terms the small suspended particles are bombarded by liquid molecules at random, and occasionally will be hit from one side more than the other, in turn causing random movement of the suspended particles. Einstein would then link this motion to the diffusion equation using a probabilistic approach by studying the number of suspended particles that would displace a small distance Δx over a small interval of time τ.

With the mathematical model given by Einstein, Jean Perrin was next to step up to the plate. He would use a microscope and a "camera lucida" (an optical technique for superimposing an object viewed through a microscope onto a surface for tracing) setup to mark the positions of suspended particles in a liquid on fixed time intervals. Perrin would collect data for 50 particles and find the mean squared displacement (the displacement from a particles starting position to its position after the fixed time interval squared), which would then be used in a calculation of Avogadro's number (first evaluated in 1860 by Stanislao Cannizzaro[13]) presented by Einstein[6]. Perrin confirmed the value found by Cannizzaro within error, thus affirming Einsteins mathematical model of Brownian motion.

1.2 Basics of Diffusion

Let us start investigating diffusion by deriving the diffusion equation. The simplest situation is steady-state, wherein individual particles move within the system, however the overall concentration in any section that we sub-sample remains constant; alternatively it can also be defined such that the concentration pattern is constant throughout when the total number of particles is not constant. In specific fields, such as drug delivery, we may be interested in the number of particles that leave or enter a given system. Fick's first law, which describes the number of particles crossing a cross-sectional area A in a fixed time step τ, allows us to build up mathematical tools to further understand non steady-state diffusion. Let us begin by considering Fig. 1.1. We will refer to particles to the left of area A within a distance Δx as being located at x, and particles to the right of area A within a distance Δx as being located at $x+\Delta x$. These particles may only move along the x-axis at random, i.e. the probability of moving to the left during the time step τ is equal to the probability of moving right which are equal to $\frac{1}{2}$. We will refer to the number of particles as N, in which case we have $N(x)$ particles on the left of our area A, and $N(x+\Delta x)$ particles on the right. Given our particle jumping probabilities of $\frac{1}{2}$, we expect $\frac{1}{2}N(x)$ particles to jump from the left side of A to

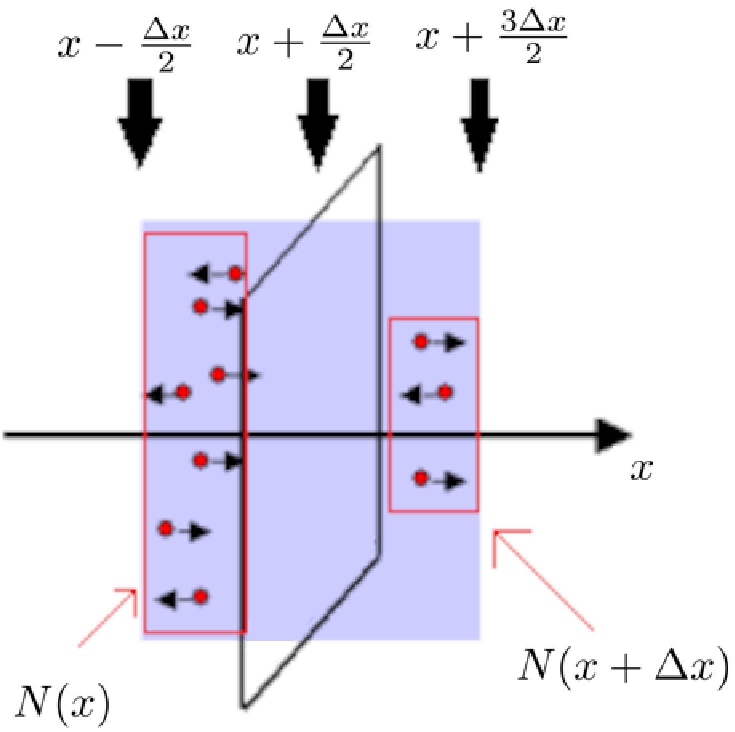


Figure 1.1: We consider particle motion only along the x-axis. We take a cross-sectional area, A, (black rectangle at position $x+\frac{\Delta x}{2}$) and investigate the number of particles that cross this area within a time step τ. The red rectangles encapsulate the particles on either side of the cross-sectional area through which particles move. The number of particles on the left of A is denoted by $N(x)$, and the number of particles on the right of A is denoted $N(x+\Delta x)$.

the right side during the time step τ. Similarily we expect $\frac{1}{2}N(x+\Delta x)$ particles from the right of A to jump across to the left of A during the time step τ. We can then calculate the net number of particles that have crossed the area A to the right during the time step τ to be $\frac{1}{2}[N(x)-N(x+\Delta x)]$. If we divide the net number of particles by the area A and the amount of time that has passed τ, we obtain the flux J:

$$J = \frac{\frac{1}{2}\left[N(x)-N(x+\Delta x)\right]}{A\tau} \,. \tag{1.1}$$

Rarely is the number of particles dealt with explicitly: we instead operate using concentration $C(x)$. First let us define the concentration: on the left of our area A we have $N(x)$ particles in a volume $A\Delta x$, which leaves us with $C(x) = \frac{N(x)}{A\Delta x}$. Substituting this into eq.1.1 leaves us with:

$$J = \frac{\frac{1}{2}\left[C(x)A\Delta x \,-\, C(x+\Delta x)A\Delta x\right]}{A\tau} = \frac{\Delta x}{2\tau}\left[C(x)-C(x+\Delta x)\right] = \frac{(\Delta x)^2}{2\tau} \times \frac{C(x)-C(x+\Delta x)}{\Delta x} \,. \tag{1.2}$$

In the limit of $\Delta x \to 0$ we can replace $\frac{C(x+\Delta x)-C(x)}{\Delta x}$ with $\frac{\partial C(x)}{\partial x}$, and we call $\frac{(\Delta x)^2}{2\tau}$ the diffusion coefficient (note that this does not vary spatially in our example, i.e. $D \neq D(x)$). This leaves us with Fick's law relating the flux to the derivative of the concentration:

$$J = -D\frac{\partial C(x)}{\partial x} \,. \tag{1.3}$$

Next consider Fig. 1.2, where we have a volume $A\Delta x$. The flux into the box is $J(x)$ (particles coming into the box from the left side) while the flux out of the box is $J(x+\Delta x)$ (particles leaving the box out the right wall). There is some concentration $C(x,t)$ in the box initially, and the change in concentration over a time interval τ is $C(x,t+\tau)-C(x,t)$. The difference in concentration between times $t+\tau$ and t is simply the difference between the influx of particles coming from the left ($J(x)A\tau$) and the amount of particles leaving out the right side ($J(x+\Delta x)A\tau$) divided by the volume of the box $A\Delta x$. We can convert these into rates by dividing both by the time step τ resulting in

$$\frac{C(x,t+\tau)-C(x,t)}{\tau} = \frac{\left[J(x)-J(x+\Delta x)\right]A\tau}{A\Delta x\tau} \,. \tag{1.4}$$

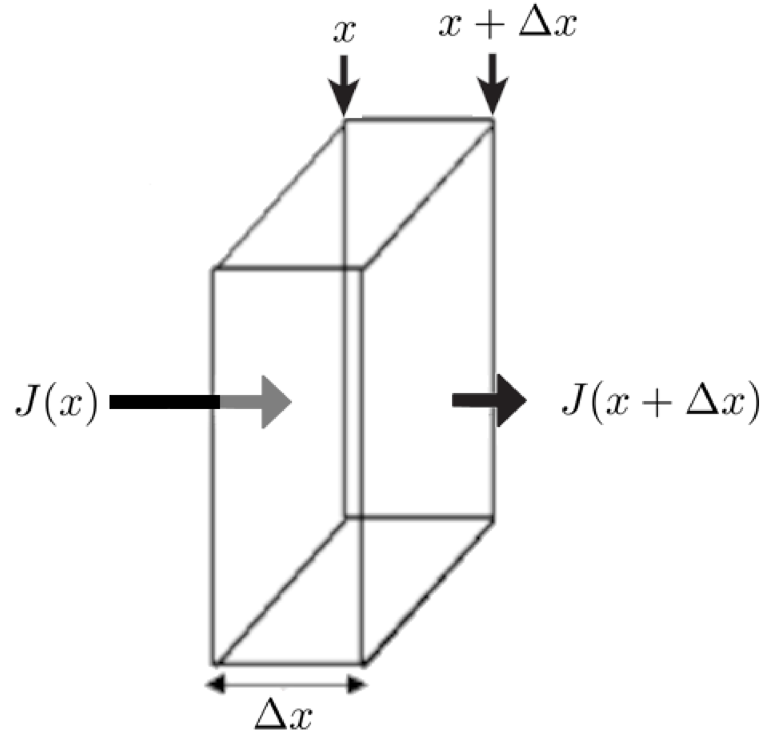


Figure 1.2: A rectangular volume $V = A\Delta x$ with cross sectional area A is shown with an incoming flux from the left wall, and outbound flux out of the right wall.

Taking the continuum limit of time t ($\tau \to 0$) simplifies the *lhs* to $\frac{\partial C(x,t)}{\partial t}$. Next we can take the continuum limit of space x ($\Delta x \to 0$) to retrieve $-\frac{\partial J(x)}{\partial x}$, resulting in the continuity equation:

$$\frac{\partial C(x,t)}{\partial t} = -\frac{\partial J(x)}{\partial x} \ . \tag{1.5}$$

With the continuity equation 1.5 we can simply substitute eq. 1.3 in for the flux (since we have already evaluated the continuum limit in x) resulting in the diffusion equation:

$$\frac{\partial C(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left(-D\frac{\partial C(x,t)}{\partial x}\right) = D\frac{\partial^2 C(x,t)}{\partial x^2} \ . \tag{1.6}$$

The derivations above have been using only one-dimensional motion in a three-dimensional environment; the equations are easily generalized since each dimension is independent of one-another. The spatial derivatives can thus be rewritten as gradient operators to account for dimensionality.

We however did not consider the possibility of a spatially varying diffusion coefficient $D(\vec{r})$, in which case eq. 1.3 would include the spatial dependence of the diffusion coefficient, leading us to not being able to extract D in front of the second derivative in eq. 1.6. The resulting general form of the diffusion equation with a spatially dependent diffusion coefficient is:

$$\frac{\partial C(\vec{r},t)}{\partial t} = \nabla \cdot \left(D(\vec{r}) \cdot \nabla C(\vec{r},t)\right) \ . \tag{1.7}$$

In this body of work, we will assign a fixed value of D for particles diffusing within the medium, i.e. it will not vary spatially $D(\vec{r}) = D$. Thus we can simplify the three-dimensional diffusion equation as follows:

$$\frac{\partial C(\vec{r},t)}{\partial t} = D\left[\nabla^2 \cdot C(\vec{r},t)\right] \ . \tag{1.8}$$

A portion of this thesis involves investigating the behaviour of a tracer particle moving within a uniform two-dimensional medium, corresponding to a delta peaked initial condition with the centre located at $\vec{r} = 0$, i.e. $C(\vec{r},t=0) \sim \delta(\vec{r})$. This is a standard initial value problem differential equation, and the solution is a Gaussian of the form

$$g(r,t) = Ae^{-r^2/\sigma^2} \; . \tag{1.9}$$

The variance is

$$\sigma^2 = 4Dt \tag{1.10}$$

in the two dimensional case. The normalization factor A can be evaluated

$$\frac{1}{A} = \int_0^{2\pi} \int_0^{\infty} r\, e^{-\frac{r^2}{4Dt}}\, dr\, d\theta \; , \tag{1.11}$$

to obtain the final form

$$p(r,t) = \frac{1}{4\pi Dt} e^{-\frac{r^2}{4Dt}} \; , \tag{1.12}$$

which we can multiply by the total number of particles N_{tot} in a system to convert from a probability distribution to a concentration:

$$C(r,t) = \frac{N_{tot}}{4\pi Dt} e^{-\frac{r^2}{4Dt}} \tag{1.13}$$

in two dimensions. The probability distribution 1.12 is shown for various times in Fig. 1.3. We show only positive values of r due to r being defined as the distance from the centre of the system.

We use the probability distribution $p(r,t)$ for calculating the continuum limit solutions of the second and fourth moments of our systems. Calculating these quantities is simply taking the expectation value of the displacement to the second, and fourth powers respectively as follows:

$$\langle r^2(t) \rangle = \int_0^{2\pi} \int_0^{\infty} r^2 p(r,t) r\, dr\, d\theta = 4Dt \tag{1.14}$$

This calculation can be generalized to any dimension with $\left\langle r^2(t) \right\rangle = 2nDt$ where n is the dimensionality of the system ($n = 1$ for one dimension, $n = 2$ for two dimensions, and so on...). Note that the second moment of the distribution is also the analytical variance of the solution to the diffusion

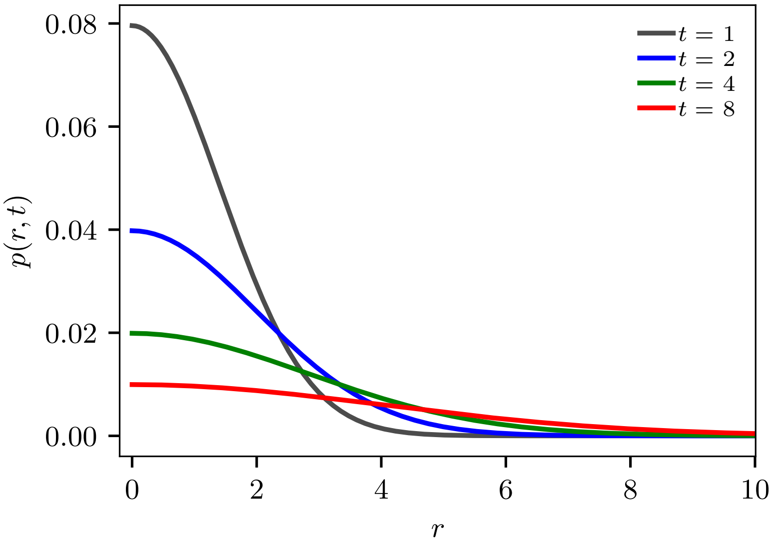


Figure 1.3: The probability distribution $p(r,t)$ is shown for a variety of times as a function of the radial distance r. The radial distance is defined as $r^2 = x^2 + y^2$; we use $D = 1$ and dimensionless variables for this example.

equation and thus dictates the width and broadening of our distribution. We introduce the fourth moment as well:

$$\langle r^4(t)\rangle = \int_0^{2\pi}\int_0^{\infty} r^4 p(r,t) r\,dr\,d\theta = 32D^2t^2 = 2\langle r^2(t)\rangle^2 \ . \tag{1.15}$$

Equation 1.15 will be useful in motivating some Lattice Monte Carlo choices in Section 2.1. We do not spend any time worrying about the odd moments of the distribution as the integrals yield 0 due to symmetry.

1.3 Anomalous Diffusion

Many diffusive processes do not behave in the same manner as described in eq. 1.14 for all times. Instead their transient (intermediate times) behaviour is often fitted using

$$\langle r^2(t)\rangle = 4D_\alpha t^\alpha \ , \tag{1.16}$$

and is referred to as anomalous diffusion, where α is the anomalous exponent that determines the scaling behaviour of the mean squared displacement $\langle r^2(t) \rangle$, and D_α is the anomalous diffusion coefficient [1] corresponding to the instantaneous observed diffusivity of the particle throughout the fitted time interval. As described, in the transient regime we expect eq. 1.14 not to apply, and thus $\alpha \neq 1$ (however α is a real positive number). We can split this into two different categories, $\alpha < 1$ (sub-diffusion) and $\alpha > 1$ (super-diffusion). The majority of cases in anomalous diffusion involves the prior, including our work with obstructing obstacles.

Some examples of anomalous diffusion models are Continuous Time Random Walks (CTRW) [17], Levy Flights [37], and diffusion in percolation systems [31]. A CTRW is a more general version of the typical random walk. In a CTRW the particles have both a distribution for how far they will jump, as well as how often they will jump. A Levy flight is similar in that the length of each step is chosen from a distribution, however with the condition that it is a heavy tailed distribution. Finally diffusion in percolation systems is the process of diffusion through a porous media consisting of percolation clusters (think coffee filter). In all of these examples we can imagine that we may observe diffusion that is faster, or slower than what we would normally expect. For Levy flights heavy tailed distributions yields to small step size clustering followed by large jumps. This results in strange behaviour in comparison to regular diffusion, and can in fact be applied to model animal hunting patterns [12, 28]. In the most general case of a CTRW, we are able to model both sub-diffusion ($\alpha < 1$) and super-diffusion ($\alpha > 1$). These different cases can be achieved by the choice of distributions for step sizes as well as the time between steps. In percolation systems the most common example is a random walker traversing a percolation cluster. The percolation cluster itself can be thought of as a maze, and in fact the diffusive behaviour will be anomalous for all times [31]. In all of these models we often expect the mean squared displacement (or second moment of our distribution) to not behave as we calculated it in eq. 1.14.

In our work we have a case that is more similar to percolation, in which we have obstacles that obstruct diffusion within our systems. This leads to three regimes: the free-diffusion regime where

---

[1] This is not a real diffusion coefficient as the units depend on α, the anomalous exponent, and thus are non standard.

our system follows the behaviour of eq. 1.14 with a free diffusion coefficient D_0 as tracer particles have not come into contact with obstacles; the transient regime (where anomalous-like diffusion may take place) where we have yet to explore the entire system ; finally the steady-state is reached, at which point the second moment returns to scaling linearly with time as in eq. 1.14 however with a reduced diffusion coefficient $D < D_0$.

CHAPTER 2

METHODS

2.1 Lattice Monte Carlo

Simulating diffusion in the presence of obstructing obstacles is challenging and computationally expensive in continuous space. As a result the problem is often mapped onto a lattice. In our case we will work with a 2-dimensional square lattice; other options exist such as a triangle lattice, or the honeycomb lattice, however a square lattice offers the simplest calculations. A square lattice leaves us with 4 natural directions (diagonal jumps are excluded; further discussion as to why can be found in Section 2.9) a tracer particle can move along, $\pm x$ and $\pm y$, as well as being able to remain at the same position. Figure 2.1 is a visual representation of our Lattice Monte Carlo (LMC) model to illustrate important system components. In this case we show the spacing between lattice sites to be a, the probabilities of hopping along either x or y, and the probability of staying on the same site. The standard set of jumping probabilities used are $p_{\pm x} = p_{\pm y} = \frac{1}{4}$ and $p_0 = 0$ [5, 7, 27]: however we will use $p_{\pm x} = p_{\pm y} = \frac{1}{8}$ with $p_0 = \frac{1}{2}$ (the motivation for this decision will be presented in Section 2.9).

Most commonly this model is used for single particle Monte Carlo simulations, whose position is tracked as a function of time. After each time step of duration τ the particle chooses to move in a random direction or stay still with the probabilities given. In Fig. 2.2 we present an example of a red tracer particle moving throughout a lattice. It is important to discuss the Monte Carlo rules that apply when an obstacle is adjacent to the tracer particle (see the blue particle in Fig. 2.2). If a hop is chosen towards an adjacent obstacle the particle is reflected and returns to its position prior

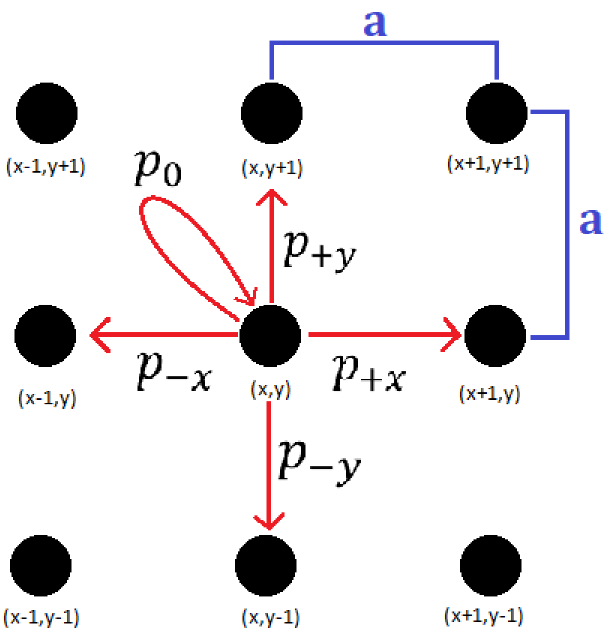


Figure 2.1: Square lattice positions separated by a lattice spacing a. Probabilities of jumping along any cardinal direction are present $(p_{\pm x}, p_{\pm y})$, as well as a probability p_0 of remaining on site (i, j).

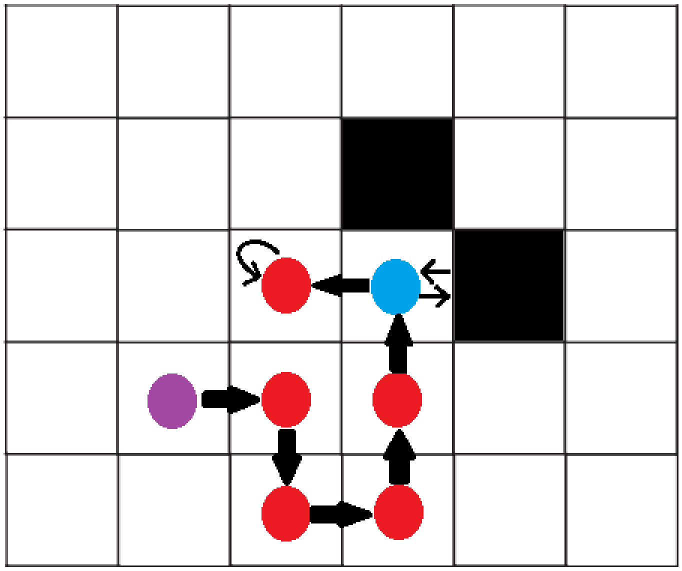


Figure 2.2: An example of a single particle random walking within a two-dimensional square lattice with two obstacles (in black) present. The starting position is labelled in purple, and the particle is highlighted in blue before and after the collision with an obstacle.

to the attempted hop. We can thus interpret the obstacles as reflecting boundary conditions. In other words it can be thought that p_0 increases to accommodate the probability of reflecting off an obstacle; however it should be noted that the time step τ does not change in this line of thinking while time is still incremented when jumps are rejected.

As discussed in Section 1.2 we are interested in the mean squared displacement $\langle r^2(t) \rangle$ as a function of time t. In the LMC tracer particle approach this is achieved by simulating a large ensemble of random walks as the one in Fig. 2.2. Let us examine this case after five time steps. In the example of Fig. 2.2 the tracer particle has moved from the purple position, to the blue position. The tracer particle has essentially moved two lattice sites in the $+x$ direction, and one lattice site in the $+y$ direction. The squared displacement after $t = 5\tau$ is $r^2(5\tau) = (2a)^2 + (a)^2 = 4a^2 + a^2 = 5a^2$ with a lattice spacing of a. In fact the same squared displacement is observed at $t = 6\tau$ as the particle then collides with an obstacle and returns to the blue position. Computing r^2 at every time step, for a large ensemble of obstacle systems and initial tracer particle positions, yields the necessary information for the parameter space we explore. Note the difference between r^2 in this example, and $\langle r^2(t) \rangle$ as we discussed prior. The $\langle ... \rangle$ notation implies an ensemble average over both disorder, i.e. obstacle configurations for any non-periodic systems (random or "Fuzzy"), and initial position of the tracer particle.

2.2 Markov Chain Monte Carlo

Markov Chain Monte Carlo involves converting the random walk Lattice Monte Carlo approach for a single tracer particle into a time-dependent concentration evolution. Rather than having a single tracer particle walk in a given direction that is associated with a probability, we instead evolve a concentration profile along each of the directions by the associated probabilities. This method thus immediately calculates the average behaviour of an infinite ensemble of tracer particle random walks. However as we average over disorder in the configuration of obstacles in our system, we must also average over the initial position of the tracer particle thus suggesting that for a single obstacle configuration we do not have the exact infinite ensemble size limit. Mathematically, on an empty lattice, the concentration $C_{x,y}(t)$ evolves as:

$$\begin{aligned}C_{(x,y)}(t+\tau) = {} & C_{(x,y+1)}(t)\times p_{-y} + C_{(x,y-1)}(t)\times p_{+y} \\ & + C_{(x+1,y)}(t)\times p_{-x} + C_{(x-1,y)}(t)\times p_{+x} \\ & + C_{(x,y)}(t)\times p_0 \; .\end{aligned} \tag{2.1}$$

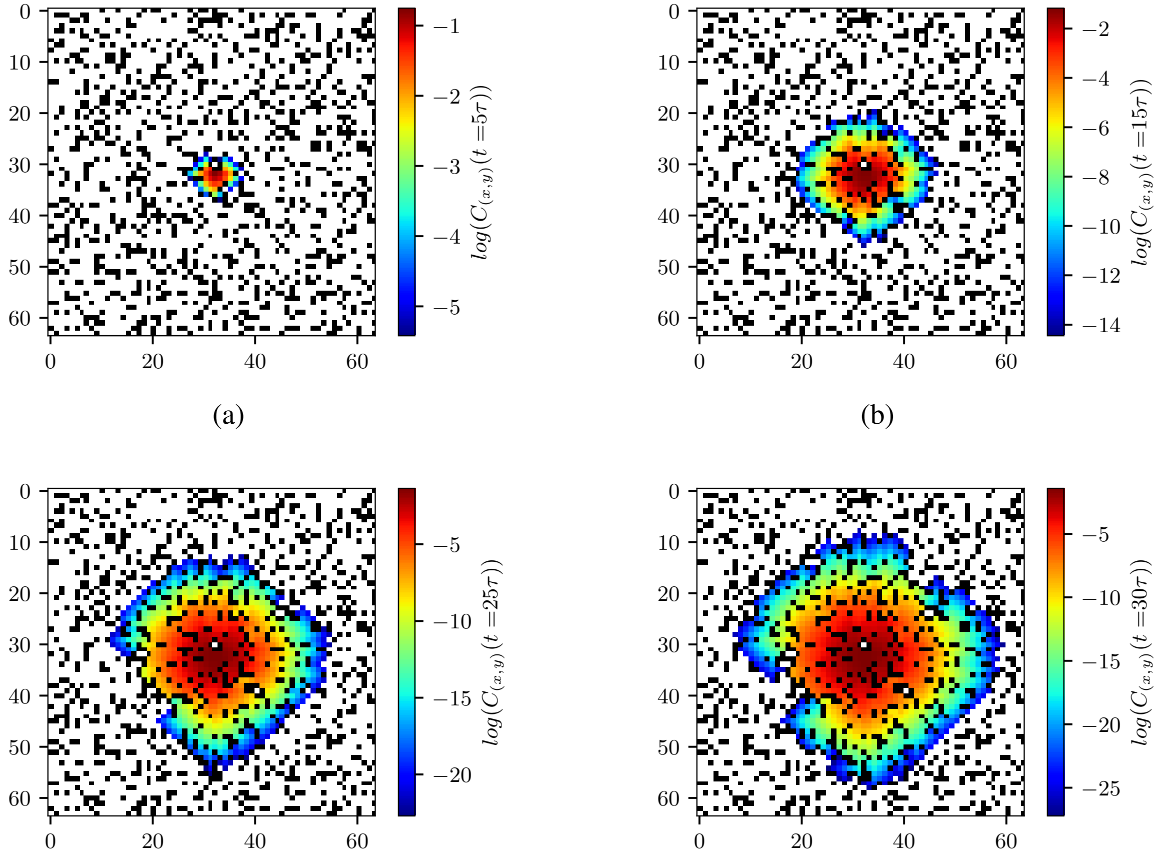
Our work includes obstacles and thus eq. 2.1 needs to be modified to include rejected hops due to obstacles that are adjacent to the site (x,y):

$$\begin{aligned}C_{(x,y)}(t+\tau) = {} & [C_{(x-1,y)}(t) + C_{(x,y)}(t)O_{(x+1,y)}]p_{+x} \\ & + [C_{(x+1,y)}(t) + C_{(x,y)}(t)O_{(x-1,y)}]p_{-x} \\ & + [C_{(x,y-1)}(t) + C_{(x,y)}(t)O_{(x,y+1)}]p_{+y} \\ & + [C_{(x,y+1)}(t) + C_{(x,y)}(t)O_{(x,y-1)}]p_{-y} \\ & + [C_{(x,y)}(t)p_0] \; ,\end{aligned} \tag{2.2}$$

where $O_{x,y}$ is 1 if there is an obstacle at (x,y) and 0 otherwise. we will briefly discuss the first term on the *rhs* of eq. 2.2. We can break this first term into two smaller terms:

- $C_{(x-1,y)}(t)p_{+x}$, where this term calculates the concentration that will move in the $+x$ direction from position $(x-1,y)$ to position (x,y) in a time step τ;
- $C_{(x,y)}(t)O_{(x+1,y)}p_{+x}$, where this term calculates the concentration that is reflected from an obstacle to the right (it will be 0 if there is no obstacle at $(x+1,y)$ because $O_{(x+1,y)} = 0$, otherwise we have $C_{(x,y)}(t)p_{+x}$).

We must assign probabilities of moving along any given axis, or staying in the same place; these probabilities have been discussed in Section 2.1 and we will use $p_{\pm x} = p_{\pm y} = \frac{1}{8}$ and $p_0 = \frac{1}{2}$. Equation 2.2 is particularly important as it is the rate equation for position (x,y) (e.g. it depicts the concentration that is incoming to position (x,y) from other positions and itself) and will be used extensively in Section 2.3.



(a) (b)

(c) (d)

Figure 2.3: Concentration evolution using Markov Chain Monte Carlo after 5 , 15 , 25 , and 30 steps for subfigures $a)$, $b)$, $c)$, and $d)$ respectively. The obstacles are configured completely randomly at an obstacle concentration $\phi = 0.25$, and shown as black squares. White square represent empty space, and the colour distribution throughout the lattice is a measure of the concentration profile $C_{(x,y)}(t)$.

An example of the concentration evolution for an initial Kronecker delta peaked concentration at the centre of the lattice is shown after various periods of time in Fig. 2.3. We can see the effects of obstacles on the distribution at intermediate times in Fig. 2.3 (c) where the concentration has not spread like an isotropic Gaussian because of rejections due to the presence of obstacles. After 30 time steps (Fig. 2.3 (d)) a fairly isotropic Gaussian is observed (a smoother perimeter is observed), suggesting Gaussian behaviour for long times.

2.3 Numerically Exact Diffusion Coefficient

Studying diffusion naturally involves measuring or estimating the diffusion coefficient of a particle within a system, whether it be the free diffusion coefficient, or the steady-state diffusion coefficient; the former being significantly easier to obtain, whereas the latter requires extensive times to be reached. The mean squared displacement, and how it scales with time, provides one avenue to learn about a diffusive system; however it only paints one part of the picture, primarily what happens at early times when very long times are inaccessible. Capturing the steady-state behaviour using MCMC requires simulating systems to very long times, and at best estimating the $t \to \infty$ limit information using an extrapolation scheme. We instead opt to calculate the steady-state diffusion coefficient beforehand using one of the methods introduced in Mercier, Slater, and Guo [18].

2.3.1 Theory

The method of Mercier, Slater, and Guo [18] begins by relating the steady-state diffusion coefficient D to the velocity of a walker in an obstructed system $v(\varepsilon)$, which is the mean global velocity over the lattice and ε is the amplitude of an applied bias along the x-axis (you could also choose to use the y-axis). The steady-state diffusion coefficient of our walker in a system with obstacles concentration ϕ is $D(\phi)$, and $D(0) = 1$ is the free diffusivity. The premise of the calculation lies within the following equation (a result of the fluctuation-dissipation theorem):

$$\frac{D(\phi)}{D(0)} = \lim_{\varepsilon \to 0} \frac{v(\varepsilon,\phi)}{v(\varepsilon,0)} \,. \tag{2.3}$$

A caveat of this calculation is that it measures the steady-state diffusion coefficient along the bias axis only. We will apply the bias along the x-axis, although as we primarily work with systems consisting of periodically or randomly placed obstacles, in the limit of an infinite ensemble size the diffusion coefficient along either axis will be the same.

On a discrete lattice, we can compute the mean global velocity v as

$$v = \langle v|n\rangle \tag{2.4}$$

where $|n\rangle$ encodes the steady-state probability of a walkers presence on each lattice site in a column vector (whether or not it is under a bias), and $\langle v|$ is the local velocity at each site encoded in a row vector. The entries of the local velocity vector are simply

$$v(i) = [p_{+x}(1-O_{+}(i)) - p_{-x}(1-O_{-}(i))]\,a/\tau \tag{2.5}$$

where $O_{\pm}(i)$ describes the presence of an obstacle in the given direction for the i'th element in the vector (e.g. $O_{+}(i) = 1$ if there is an obstacle to the right of i, and 0 if there is no obstacle), a is the lattice spacing, and τ is the time step for one lattice hop.

The first step in this calculation is evaluating what the steady-state concentration profile will be when a weak bias ε is applied. Most notably the bias will affect the local velocities by altering the probabilities of jumping to the right or to the left. In this calculation we will make use of $p_{\pm x} = p_{\pm y} = \frac{1}{4}$, and thus we use the simple linear form $p_{+x} = \frac{1}{4}(1+\varepsilon)$ and $p_{-x} = \frac{1}{4}(1-\varepsilon)$ given the bias is applied along the $+x$ axis. This is contrary to the $p_{\pm x} = p_{\pm y} = \frac{1}{8}$ choice we have made previously; however in the steady-state limit we reach the same result. To demonstrate this we must first realise that when changing between $p = \frac{1}{4}$ and $p = \frac{1}{8}$ we must adjust the time step τ

accordingly, where $\tau = {}^{pa^2}/_{D_0}$ (details for this formulation are found in Section 2.9) and D_0 is the free diffusion coefficient. In our case, and for simplicity, we use $a = 1$ and $D_0 = 1$, yielding $\tau = p$. If we consider a free system (no obstacles), we can compute v (it is site-independent in this case as all sites are equally probable and are not obstructed) for both $p = \frac{1}{4}$, and $p = \frac{1}{8}$. We will add subscripts to identify which probability choice the time step corresponds to. The associated local velocity is calculated with:

$$v_{p=1/4} = \left(\frac{1+\varepsilon}{4} - \frac{1-\varepsilon}{4}\right)\frac{a}{\tau_{p=1/4}} = \frac{a\varepsilon}{2\tau_{p=1/4}}\ , \tag{2.6}$$

and

$$v_{p=1/8} = \left(\frac{1+\varepsilon}{8} - \frac{1-\varepsilon}{8}\right)\frac{a}{\tau_{p=1/8}} = \frac{a\varepsilon}{4\tau_{p=1/8}}\ . \tag{2.7}$$

Recalling that $p = \tau$, we see that $\tau_{p=1/8} = \frac{1}{8} = \frac{\tau_{p=1/4}}{2}$, substituting this into eq. 2.7 we retrieve $v_{p=1/8} = \frac{a\varepsilon}{2\tau_{p=1/4}} = v_{p=1/4}$.

We can break up the velocity (v) and concentration (n) vectors into the sum of a constant with an ε dependent vector, such that $|n\rangle = |n_I\rangle + \varepsilon|n_\varepsilon\rangle$ and $\langle v| = \langle v_I| + \varepsilon\langle v_\varepsilon|$, where the subscript I denotes the constant portion, and subscript ε denotes the bias dependent portion. We can then expand the global mean velocity $v = \langle v|n\rangle = \langle v_I|n_I\rangle + \varepsilon\langle v_I|n_\varepsilon\rangle + \varepsilon\langle v_\varepsilon|n_I\rangle + \varepsilon^2\langle v_\varepsilon|n_\varepsilon\rangle$. Since we will be evaluating the mean global velocity in the limit $\varepsilon \to 0$ we establish that $\langle v_I|n_I\rangle = 0$ as there will be no velocity in the steady-state with no bias. Furthermore we neglect the $\langle v_\varepsilon|n_\varepsilon\rangle$ term as it is of order ε^2. We are then left with

$$v(\varepsilon) = \frac{a\varepsilon}{\tau}\left[\langle v_\varepsilon|n_I\rangle + \langle v_I|n_\varepsilon\rangle\right]\ , \tag{2.8}$$

where we have scaled out a factor of ${}^{a}/_{\tau}$ from the two $\langle v|$ terms so that the latter become dimensionless. Returning to eq. 2.3 we need $v(\varepsilon, \phi = 0)$. When $\phi = 0$ we have the result from eq. 2.7, and $n(i) = 1/N$ (where N is the number of accessible sites). Combining $v(\varepsilon, \phi = 0) = \frac{a\varepsilon}{2\tau}$, $D(0) = 1$, and eq. 2.8, with eq. 2.3 we retrieve the steady-state diffusion coefficient

$$D(\phi) = 2[\langle v_\varepsilon | n_I \rangle + \langle v_I | n_\varepsilon \rangle] \ , \tag{2.9}$$

where the only remaining unknown is $|n_\varepsilon\rangle$.

The first tool we will use in solving for $|n_\varepsilon\rangle$ is the transition matrix T, which evolves the probability vector forward in time by one time step $|n(t+\tau) = T|n(t)\rangle$ and contains all of the dynamics of the system. Since in the steady-state $|n(t+\tau)\rangle = |n(t)\rangle = |n(t \to \infty)\rangle$, we can write

$$T|n\rangle = |n\rangle \ . \tag{2.10}$$

We have access to an additional constraint as we always begin with a total probability of 1, i.e.

$$\sum_{i=1}^{N} n(i) = 1 \ , \tag{2.11}$$

where N is the number of accessible sites in the lattice. We can then rewrite the problem to solve for $|n\rangle$ analytically as $(T - \mathbb{I})|n\rangle = |0\rangle$, where $\mathbb{I}$ is the identity matrix and $|0\rangle$ represents the zero vector. All but one of the equations will be linearly independent, as such we are able to impose the normalization equation $\sum_{i=1}^{N} n(i) = 1$ ahead of time. The last element of the zero vector is replaced with a 1 and the bottom row of the $T - \mathbb{I}$ matrix is replaced with all 1's, which immediately imposes a normalized solution. We relabel this matrix, and edited $|0\rangle$ vector as A and $|b\rangle$ respectively. Similarly to our local velocity vector, and probability vectors, we split the A matrix into a constant component, and ε dependent component, $A = A_I + \varepsilon|A_\varepsilon\rangle$, however $|b\rangle$ will always be independent of ε, thus we write $|b\rangle = |b_I\rangle$. After splitting up the A matrix we expand $A|n\rangle = |b\rangle$:

$$A_I|n_I\rangle + \varepsilon[A_I|n_\varepsilon\rangle + A_\varepsilon|n_I\rangle] + \varepsilon^2 A_\varepsilon|n_\varepsilon\rangle = |b_I\rangle \ . \tag{2.12}$$

Again, neglecting the higher order terms in ε, we find that this equation is satisfied only if

$$A_I|n_I\rangle = |b_I\rangle \ , \tag{2.13}$$

and

$$A_I|n_\varepsilon\rangle = -A_\varepsilon|n_I\rangle \ , \tag{2.14}$$

are satisfied simultaneously. Equation 2.13 is straight forward as it corresponds to the free diffusion case where all sites are equally occupied and we have created A_I and $|b_I\rangle$ ourselves. From eq. 2.14, with access to A_I, A_ε, and $|n_I\rangle$, we are able to solve for $|n_\varepsilon\rangle$. To do so we simply multiply both sides by the inverse of A_I

$$|n_\varepsilon\rangle = -A_I^{-1}A_\varepsilon|n_I\rangle \ . \tag{2.15}$$

From here we have all the tools we need to calculate the steady-state diffusion coefficient from eq. 2.9.

2.3.2 Example

This subsection will walk through a sample calculation of the steady-state diffusion coefficient D for a particle in the 4×4 system with 6 obstacles and PBCs ($\phi = \frac{6}{16}$) shown in Fig. 2.4.

We start by generating the transition matrix (T) using rate equations (discussed in Section 2.2) corresponding to Fig. 2.4 (note that sites that are unoccupied yet inaccessible to a particle are not included in the calculation; see the red site as an example), with biased jumping probabilities at

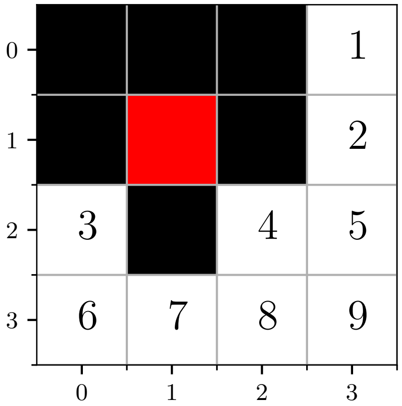


Figure 2.4: Example lattice with numbered accessible sites for transition matrix development. Black lattice sites correspond to obstacles, and the red lattice site corresponds to an unoccupied however inaccessible site. Periodic Boundary Conditions (PBCs) are present (e.g., site 5 is adjacent to sites $2, 3, 4$ and 9).

each given lattice site:

$$T = \frac{1}{4}\begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2+\varepsilon & 0 & 1+\varepsilon & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2-\varepsilon & 1-\varepsilon & 0 & 0 & 1 & 0 \\ 0 & 1 & 1-\varepsilon & 1+\varepsilon & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1-\varepsilon & 0 & 1+\varepsilon \\ 0 & 0 & 0 & 0 & 0 & 1+\varepsilon & 2 & 1-\varepsilon & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1+\varepsilon & 1 & 1-\varepsilon \\ 1 & 0 & 0 & 0 & 1 & 1-\varepsilon & 0 & 1+\varepsilon & 0 \end{pmatrix} \tag{2.16}$$

The matrix is of size $N \times N$ where $N = 9$ is the number of lattice sites accessible to the particle. Let us focus on the 1^{st} row of the transition matrix T to elaborate on how it is generated. The 1^{st} row in T corresponds to the rate equation for site #1:

$$n(1, t+\tau) = \left(\frac{1+\varepsilon}{4} + \frac{1-\varepsilon}{4}\right) n(1,t) + \frac{1}{4} n(2,t) + \frac{1}{4} n(9,t) \tag{2.17}$$

where site #1 corresponds to $(x, y) = (3, 0)$. The adjacent sites are $(2, 0)$ and $(0, 0)$ where obstacles reside (thus no labels on Fig. 2.4), and $(3,1)$, $(3,3)$ (sites #2 and #9 respectively) where the lattice sites are free. Each entry from left to right in the first row of T corresponds to the hopping probability to site #1 from sites #1 through #9 in order. Site #1 is adjacent to obstacles on both the right (due to PBC) and left side of it; all hops along x reflect back to site #1, the corresponding probability of staying is $\frac{1}{4}(1+\varepsilon)+\frac{1}{4}(1-\varepsilon)=\frac{2}{4}$, which is the first entry of row #1. The second entry means that site #1 is adjacent to site #2: the hopping probability for site #2 to site #1 is $\frac{1}{4}$ as it is a vertical hop (not along the bias axis). The 3^{rd},4^{th},5^{th},6^{th},7^{th}, and 8^{th} entries are empty as these sites are not adjacent to site #1. The final entry is the 9^{th} column, as site #9 is adjacent to site #1 through periodic boundary conditions by a downward hop, and is thus given a value in the transition matrix of $\frac{1}{4}$. Following the steps in Section 2.3.1 we develop $T-\mathbb{I}$

$$T-\mathbb{I}=\frac{1}{4}\begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2+\varepsilon & 0 & 1+\varepsilon & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2-\varepsilon & 1-\varepsilon & 0 & 0 & 1 & 0 \\ 0 & 1 & 1-\varepsilon & 1+\varepsilon & -4 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -3 & 1-\varepsilon & 0 & 1+\varepsilon \\ 0 & 0 & 0 & 0 & 0 & 1+\varepsilon & -2 & 1-\varepsilon & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1+\varepsilon & -3 & 1-\varepsilon \\ 1 & 0 & 0 & 0 & 1 & 1-\varepsilon & 0 & 1+\varepsilon & -4 \end{pmatrix}, \qquad (2.18)$$

which we modify to find A. To reach eq. 2.14 we split A into its components and write $|b_I\rangle$:

$$A_I=\frac{1}{4}\begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & -4 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -3 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & -3 & 1 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \end{pmatrix}, \qquad (2.19)$$

$$A_\varepsilon = \frac{1}{4}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.20}$$

and

$$|b\rangle = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}^T. \tag{2.21}$$

To solve for the bias dependent portion of the probability vector we need the bias independent portion (where each site is equally probable)

$$|n_I\rangle = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^T, \tag{2.22}$$

and the inverse of A_I

$$A_I^{-1} = \frac{1}{81 \times 5}\begin{pmatrix} -860 & -280 & 290 & 290 & 120 & 280 & 370 & 280 & 45 \\ -410 & -1000 & 110 & 110 & -150 & 190 & 280 & 190 & 45 \\ 130 & 80 & -916 & 56 & -150 & -242 & 10 & 82 & 45 \\ 130 & 80 & 56 & -916 & -150 & 82 & 10 & -242 & 45 \\ 40 & -100 & -70 & -70 & -420 & 100 & 190 & 100 & 45 \\ 220 & 260 & -142 & 182 & 120 & -584 & -170 & 64 & 45 \\ 220 & 260 & 20 & 20 & 120 & -260 & -980 & -260 & 45 \\ 220 & 260 & 182 & -142 & 120 & 64 & -170 & -584 & 45 \\ 310 & 440 & 470 & 470 & 390 & 370 & 460 & 370 & 45 \end{pmatrix}. \tag{2.23}$$

Using these matrices in eq. 2.15 we solve for $|n_\varepsilon\rangle$

$$|n_\varepsilon\rangle = \frac{1}{45}\begin{pmatrix} 0 & 0 & 6 & -6 & 0 & 2 & 0 & -2 & 0 \end{pmatrix}^T. \tag{2.24}$$

Next we need the local velocity vector corresponding to eq. 2.5, and we will need to split it into a field independent term v_I and field dependent term v_ε. For our example the velocity vectors are

$$\langle v_I| = \frac{1}{4}\begin{pmatrix} 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.25}$$

$$\langle v_\varepsilon| = \frac{1}{4}\begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 2 & 2 & 2 & 2 \end{pmatrix}. \tag{2.26}$$

We now have all of the vectors we need to compute the steady-state diffusion coefficient using eq. 2.9 and for this example case we find $D = \frac{8}{15}$.

2.3.3 Sparse Matrices

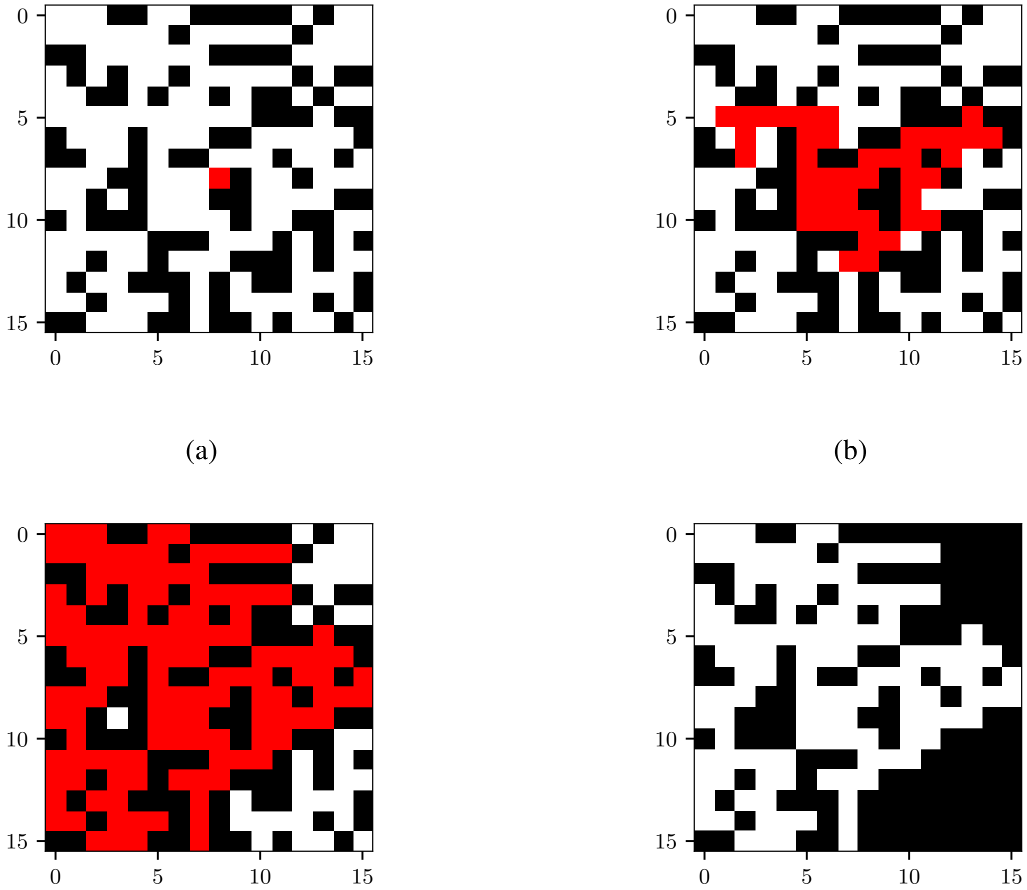
In practice for a lattice of size $L \approx 100$, low concentrations have $N \approx 100 \times 100 \approx 10,000$ accessible sites. This gives us a transition matrix of size $10^4 \times 10^4$, having on the order of 10^8 entries, and inverting this matrix to solve for the steady-state concentration profile is impossible. Fortunately any given site is adjacent to at most 4 other sites, thus each row of the transition matrix has at most 4 non-zero entries. We choose to encode our transition matrix into a Compressed Sparse Row (CSR) matrix, which comprises 3 arrays of size M, where M is the number of non-zero entries in the matrix of interest. The first array stores the values of the non-zero elements of the matrix, the second array contains the row indices of the corresponding values, and the third array contains the column indices. After building the transition matrix, we have access to A_I, A_ε, $|b_I\rangle$. Unfortunately the inverse of A_I (A_I^{-1}) is not a sparse matrix, thus we cannot use eq. 2.15. Instead, we use the SciPy library [16] to access a sparse matrix package, and use the linear algebra function from within this package that calls umfpack [4]. The algorithm in umfpack solves a linear system of form $Cx = d$ for x, where C is a sparse matrix (for our application we have $C = A_I$, $x = |n_\varepsilon\rangle$, and $d = -A_\varepsilon|n_I\rangle$), by computing a Lower Upper (LU) factorization of C into a lower triangular matrix L, an upper triangular matrix U, permutation matrices P and Q, and a diagonal matrix for row scaling R such that we have $PR^{-1}A_IQ = LU$. The preliminary solution to the linear system is computed with forward/backward substitutions, and then refined iteratively. After a solution is found we are able to return to eq. 2.9 with all of the components we need to calculate the steady-state diffusion coefficient.

2.4 Lake Filling and Connectivity

When generating obstacle configurations we must consider the possibility of lakes (sites that are surrounded by obstacles but not populated by an obstacle and thus are inaccessible to a tracer particle starting outside of the lake), and overall connectivity of a system. We are not interested in the diffusion coefficient if we start in a lake as it is simply 0 in the steady-state, so we fill these lakes with phantom obstacles.

This lake-detection procedure is simple: an initial position (the starting point of the filling algorithm) is chosen at random throughout the lattice, and a filling algorithm is executed (see Fig. 2.5). It is important to note that the resulting filled obstacle configuration can depend on the initial position. The filling algorithm can be simply described as allowing a delta peaked initial concentration to evolve with reflecting boundary conditions around the edges of the lattice. The concentration continues to evolve until no new sites are discovered (where discovered is defined as a site becoming filled). All sites that have been discovered are accessible to a tracer particle given the chosen initial position, and all sites that have no concentration are either obstacles, or inaccessible. Obstacles are thus placed on every site that does not have a concentration within it: consequently all lakes (inaccessible sites) are filled. It is important to note that it is possible for the filling algorithm to start in a lake. In this case the lake itself will be considered accessible, and the rest of the lattice will be filled with obstacles, meaning the PBCs will not be accessible resulting in a disconnected system (see next paragraph).

In Section 2.3 we make use of PBCs to treat an obstacle configuration as a patch of an infinite system. Figure 2.6 (b) is Fig. 2.6 (a) repeated on either side of the centre unit cell (this is a larger system of the unit cell created using PBCs). If we consider a tracer particle that starts in the centre unit cell (or any unit cell) of Fig. 2.6 (b) it is clear that we are in a lake as the tracer particle is



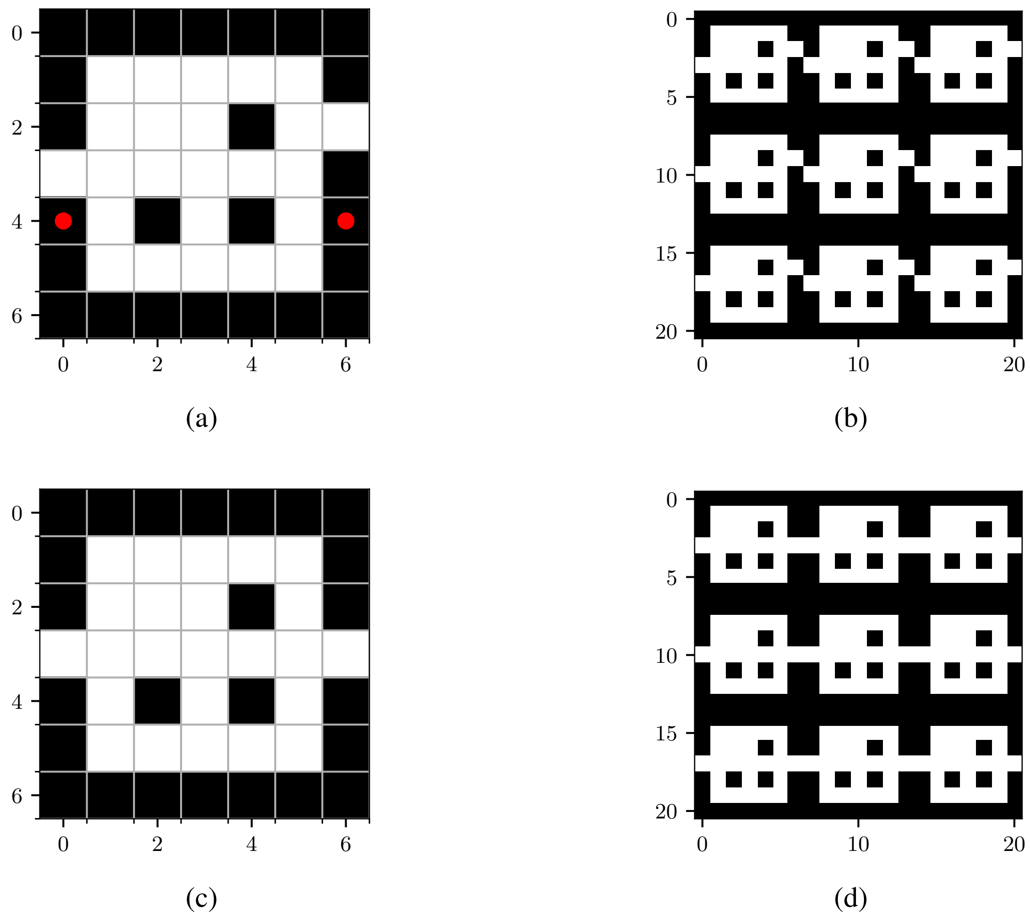
(a)

(b)

(c)

(d)

Figure 2.5: Example of the filling algorithm procedure on an initial random configuration at obstacle concentration $\phi = 0.375$. a) An initial position is chosen for the filling algorithm: for simplicity we use the centre in this example. The initial obstacle configuration is illustrated in black. b) The lattice after 40 accessible sites have been assigned the color red. c) The filling algorithm has completed. All accessible sites are painted red, leaving the initial obstacle configuration in black, and the white spaces are inaccessible to a tracer particle diffusing from the centre of the lattice. d) All sites that were not painted red in (c) are replaced with obstacles such that the initial obstacle configuration remains, and the (white) empty spaces are filled with phantom obstacles. We note that the system in (d) percolates along both x and y, and that ϕ remains unchanged independent of the addition of these "phantom" obstacles.



(a) (b)

(c) (d)

Figure 2.6: An illustration of the percolation constraint that we apply to more quickly filter for suitable obstacle configurations. a) A 7 × 7 unit cell system; black lattice sites represent obstacles. The two sites at $(0,4)$ and $(6,4)$ (with red circles) are adjacent sites through PBCs. b) The unit cell from (a) is repeated on either side of itself to emulate the PBCs applied in the numerical diffusion coefficient calculation. c) The empty sites from (a) have been shifted to $(0,3)$ and $(6,3)$ for adjacency through PBCs. d) The unit cell from (c) is repeated on either side of itself.

surrounded by obstacles. We call this a disconnected system, and Fig. 2.6 (d) is an example of a connected system (in this case along the x-axis only). We can test for connectivity by verifying that at least one set of neighbouring boundary sites (through PBCs along the x-axis) is accessible (after the filling algorithm has been applied). This percolation constraint allows a system such as Fig. 2.6 (a) to be rejected swiftly. We specifically test for connectivity along the x-axis for all obstacle configurations as the calculation in Section 2.3 calculates the steady-state diffusion coefficient along the x-axis. We apply this constraint for obstacle configurations used in MCMC simulations to remain consistent with the type of obstacle configurations that we accept for the calculation from Section 2.3 due to its use in supplementing the MCMC simulations. This constraint is particularly useful for optimizing the process of finding suitable obstacle configurations at high concentrations as lakes become increasingly more common for high concentrations. To ensure we do not reject an obstacle configuration due to an unfortunate initial position of the filling algorithm, we try up to 10 other random initial positions if the condition of a pair of adjacent PBCs is not met. After 10 failures to percolate the obstacle configuration is rejected.

2.5 Analysing the Diffusivity Ratio vs. Time

The mean squared displacement $\langle r^2(t)\rangle$ is computed from the probability distribution function $p(x,y,t)$ which we collect with MCMC simulations, and generally we investigate the short to intermediate time behaviour of $\langle r^2(t)\rangle$. Plotting $\langle r^2(t)\rangle$ *vs* t would show 3 regimes, but analysis of short and intermediate times is challenging. By dividing both sides of eq. 1.14 by $4t$ we are able to amplify the features at short and intermediate times compared to long times; we define this quantity as the Diffusivity Ratio:

$$D_r(t) = \frac{\langle r^2(t)\rangle}{4t}\ . \tag{2.27}$$

The Diffusivity Ratio has two clear limits: at short times (before collisions with obstacles) we have the free diffusion limit

$$\lim_{t\to 0} D_r(t) = D_0 \ , \tag{2.28}$$

and for long times we have steady-state behaviour

$$\lim_{t\to\infty} D_r(t) = D \ . \tag{2.29}$$

Due to the presence of obstacles we observe the particles to diffuse faster for short times than long times ($D_0 > D$). We thus adjust the second moment at steady-state (long times) with a term β^2 that we call the excess diffusivity

$$\langle r^2(t)\rangle = 4Dt + \beta^2 \ , \tag{2.30}$$

to keep track of how much more the mean squared displacement has grown due to early dynamics. This behaviour is apparent in Fig. 2.7 as the Diffusivity Ratio is larger at short times than at long times. As we reach long times β^2 becomes negligible compared to $4Dt$ thus we could find the steady-state diffusion coefficient by extrapolating the Diffusivity Ratio to $t \to \infty$. This is not needed throughout this thesis because of the method for calculating the steady-state diffusion coefficient introduced in Section 2.3.

We now consider how the Diffusivity Ratio changes in the transient regime. In Section 1.3 we introduced the idea that the transient regime is fit using eq. 1.16. Since we are going to be plotting the Diffusivity Ratio we can again divide both sides of eq. 1.16 by $4t$ to make the *lhs* the Diffusivity Ratio:

$$D_r(t) = \frac{\langle r^2\rangle}{4t} = D_\alpha t^{\alpha-1} \ . \tag{2.31}$$

We work in log-log space as this is the natural way to study and fit power laws like eq. 2.31. Taking the log of both sides of eq. 2.31 results in

$$log\left(\frac{\langle r^2\rangle}{4t}\right) = (\alpha-1)log(t) + log(D_\alpha)\ . \tag{2.32}$$

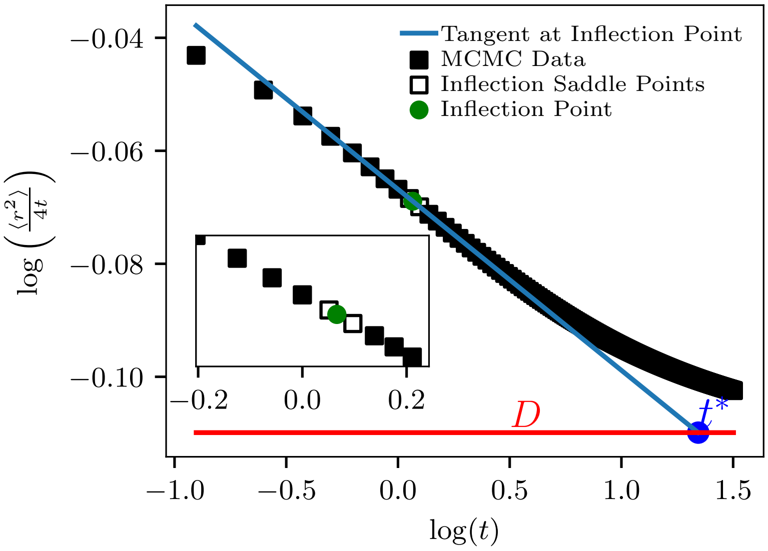


Figure 2.7: The logged Diffusivity Ratio as a function of logged time for random obstacle configurations at an obstacle concentration $\phi = 0.1$. A numerical inflection point is found in between the two hollow squares, and the interpolated inflection point is shown in green. The crossover time is the time at which the tangent line through the inflection point intersects the horizontal line giving the steady-state diffusion coefficient D. The set of probabilities used for these MCMC simulations is $p_{\pm x} = p_{\pm y} = \frac{1}{8}$, with $p_0 = \frac{1}{2}$, with an ensemble size $\Omega = 2000$.

We know that our system passes through three separate regimes: a free diffusion regime, the transient regime, and finally the steady-state regime. We expect the Diffusivity Ratio to plateau for short (eq. 2.28) and long (eq. 2.29) times; the beginning of both plateaus can be seen in Fig. 2.7. The transient regime must then connect the free diffusion and steady-state regimes smoothly, implying the presence of an inflection point.

The slope of the data is sensitive to the fitting region as the transient regime is rarely a straight line. Independent of this, it is still useful to fit the transient regime with eq. 1.16. However a robust method for determining the fitting region is necessary. We propose to use the region surrounding the inflection point in log-log space, because the inflection point by definition has no curvature, and higher order derivatives are of less importance implying the surrounding region may resemble a straight line. Note that this idea has been previously explored in the thesis of Neo Nguiya Passi [21].

The simplest way to locate the inflection point is to fit all of the data with a polynomial, and evaluate the second derivative of the fit analytically to find the inflection point. It is important to note that a fit of a complete data set will have a bias towards long time behaviour as there is more points when approaching the steady-state regime in log-log space compared to the number of points in the transient regime. Instead, we start by determining an approximate region where the inflection point resides by implementing a numerical second derivative in log-log space. To calculate the numerical second derivative we apply the forward difference method [22] two times in succession. To compute these derivatives in log-log space is not as simple as computing the numerical derivative in Cartesian space, due to equally spaced Cartesian times no longer being equally spaced in log-log space. We begin by introducing

$$d\bar{t}_i = log(t_{i+1}) - log(t_i) \ , \tag{2.33}$$

and

$$\bar{D}_{r_i} = log\left(\frac{\langle r_i^2 \rangle}{4t_i}\right) \ . \tag{2.34}$$

The first derivative is then computed as

$$\bar{D}'_{r_i} = \frac{\bar{D}_{r_{i+1}} - \bar{D}_{r_i}}{d\bar{t}_i} \ . \tag{2.35}$$

It is important to note that the times at which these derivatives are evaluated are not $log(t_i)$ but instead what we will refer to as

$$\bar{t}_i = \frac{log(t_{i+1}) + log(t_i)}{2} \ . \tag{2.36}$$

Since our times are still in log-log space but have changed to $\bar{t}_i$ we must again evaluate the time step for the second forward derivative using the same procedure as eq. 2.33. With our new times corresponding to $\bar{D}'_{r_i}$ we reapply eq. 2.35 using $\bar{D}'_{r_i}$ in place of $\bar{D}_{r_i}$. We can find the times that $\bar{D}''_{r_i}$

correspond to using the formulation from eq. 2.36 replacing $log(t_i)$ with $\bar{t}_i$.

With access to the second derivative $\bar{D}''_{r_i}(t)$ as a function of time we are able to check for a sign change (i.e., the inflection point of $\bar{D}_r$). We then fit $\bar{D}_r$ with a third order polynomial near the proposed inflection region with 5 data points on either side of the inflection point. Occasionally the inflection point will be too early such that there is not 5 data points prior to the inflection point; in this case we fit to the earliest data point available, and continue using only 5 data points beyond the inflection region. A third order polynomial is chosen to allow us to evaluate up to the third derivative of our data without overfitting a short transient regime. Using this third order polynomial we evaluate an analytic second derivative, from which we can find an interpolated inflection point. The time at which the inflection point occurs will be referred to as t_I (or $log(t_I)$ in log-log space). We find that this value reliably falls within our numerical second derivative inflection saddle points (e.g. the green data point always falls between the hollow squares in Fig. 2.7, and this is the case for all data sets).

With a precise inflection point we are able to evaluate the first derivative of our fit at t_I to then extract α and D_α from eq. 2.31. Having found the value of α, and with a value of D obtained using the method from Section 2.3, we are able to evaluate the crossover time t^*. To do so we find where $(\alpha - 1)log(t) + log(D_\alpha)$ (blue line in Fig. 2.7) intersects the steady-state diffusion coefficient D (red line in Fig. 2.7); mathematically, this correspond to

$$log(t^*) = \frac{log(D) - log(D_\alpha)}{\alpha - 1} \ . \tag{2.37}$$

The crossover length r^* follows quite simply by evaluating eq. 1.14 at the crossover time:

$$r^* = \sqrt{4Dt^*} \ . \tag{2.38}$$

To find the excess diffusivity β we naturally need to have reached the steady-state, i.e. we need to reach $\langle r^2(t) \rangle \propto 4Dt$ after the $\langle r^2(t) \rangle \propto t^\alpha$ regime. Figure 2.7 clearly shows that we do not quite reach the steady-state. Since we have access to the steady-state diffusion coefficient we can rearrange eq. 2.30 as

$$\beta^2(t) = \langle r^2(t) \rangle - 4Dt \tag{2.39}$$

to define a "time-dependent $\beta(t)$", which tracks the accumulation of excess diffusivity as a function of time. In Fig. 2.8 we show $\beta(t)$ as a function of $1/t$, from which we should be able to extrapolate to $\beta(t \to \infty)$ by using the limit $\frac{1}{t} \to 0$. Looking at the periodic data set (black curve) there is a clear plateau, meaning we have reached the steady-state, and thus can easily extract the true value of $\beta(t \to \infty)$. For random obstacle configurations we do not see a plateau, and we are unable to predict when the plateau will occur. This is partially due to the errors on the steady-state diffusion coefficient D discussed in Sections 2.7 and 2.8 as well as MCMC limited times. Instead we propose to use $\beta(t = t^*)$ for all of our data sets as a proxy for $\lim_{t\to\infty} \beta(t)$. This can be justified with the Diffusivity Ratio in Fig. 2.7 (similarily for other obstacle systems): the saturation of the Diffusivity Ratio around and beyond t^* signals we are near steady-state meaning $\beta(t = t^*)$ will not be far from the plateau value of $\beta(t \to \infty)$. Note that we will always undershoot the true value of β. The choice of $\beta(t = t^*)$ works very well for periodic data, in Fig. 2.8 $\beta(t = t^*)$ resides very close to

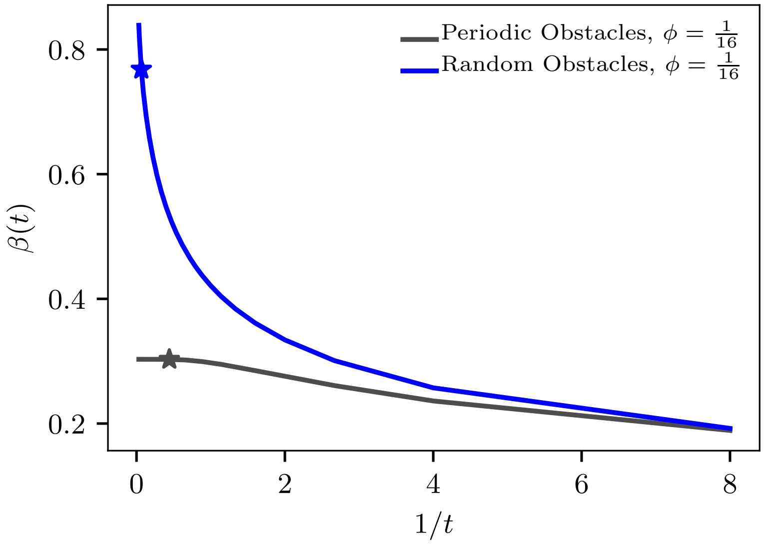


Figure 2.8: Accumulation of excess diffusivity $\beta(t)$ against inverse time. Data for a periodic, and a random configuration are shown to demonstrate the difference between them. A plateau is also expected for random configurations, but late enough time is not reached to discover it, and we cannot predict its position. The star markers correspond to the value $\beta(t = t^*)$.

the plateau, and thus returns $\beta(t = t^*) = 0.3029$ which is only 0.03% from the extrapolated value $\beta(t \to \infty) = 0.3030$.

2.6 The Width of the Anomalous Regime

The anomalous regime is typically described using eq. 1.16, where the anomalous exponent α is the prime parameter characterizing the regime. We expand on the description of the anomalous regime by introducing the anomalous regime width (Σ_t in Cartesian space and Σ_α in log-log space). The anomalous regime width describes the width of the linear portion of the transient regime, i.e. the region for which eq. 1.16 can be considered a good approximation. In Chapter 4 we show that Σ_t also provides significant information about the steady-state.

Typically the curvature of a function would be described using the second derivative; however, as we have discussed in Section 2.5, we use the inflection point $t = t_I$ as the center of the transient/anomalous regime (for a general explanation we will use $x = x_I$ briefly). To be consistent we estimate the anomalous regime width at the inflection point; consequently a general function at the inflection point returns $f''(x = x_I) = 0$. Note that since $f'' = 0$ the radius of curvature tends to ∞, i.e. a straight line; f''' then informs us of how fast f'' changes, from which $1/f'''$ can be used as a measure of the width of this straight region. We thus test two different estimates of the anomalous regime width that are of the form

$$\Sigma_\alpha^y \propto \frac{1}{\partial^3 f(x) / \partial x^3} , \tag{2.40}$$

where y is an integer exponent. Naturally the width should be in units of x (the *rhs* of eq. 2.40 has units $x^3/f(x)$); we thus use $f(x)$ and/or $f'(x)$ to fix the units (primarily to remove the remaining $f(x)$ factor that resides). Our simplest candidates (there are infinite possibilities) are

$$\Sigma_\alpha^3 = \left. \frac{f(x)}{\partial^3 f(x) / \partial x^3} \right|_{x = x_I} , \tag{2.41}$$

$$\Sigma_\alpha^2 = \left. \frac{\partial f(x) / \partial x}{\partial^3 f(x) / \partial x^3} \right|_{x=x_I} . \tag{2.42}$$

Note that we operate in log-log space with our data sets, thus our function $f(x)$ refers to

$$f(x) = \log \left(\frac{\langle r^2(t) \rangle}{4t} \right) , \tag{2.43}$$

and $x = log(t)$. We need to convert this estimate of the width to Cartesian space as follows:

$$\Sigma_t = 10^{[log(t_I)+\Sigma_\alpha]} - 10^{[log(t_I)-\Sigma_\alpha]} . \tag{2.44}$$

This can be simplified further as Σ_α is always (independent of walker/obstacle size) large compared to $log(t_I)$; thus the second term is small compared to the first and we can write

$$\Sigma_t \approx 10^{[t_I+\Sigma_\alpha]} . \tag{2.45}$$

In Fig. 2.9 we compare the anomalous regime width Σ_t as a function of obstacle concentration ϕ using both candidate Σ_α equations. The behaviour of Σ_t is similar and independent of the choice of Σ_α, however we find that using eq. 2.41 for Σ_α provided a very good estimate for the crossover time t^*. Further information on the anomalous regime width and its application can be found in Chapter 4.

2.7 Ensemble Size Effects

Ensemble size effects are a statistical feature stemming from the random nature of our obstacle configurations. Sampling random obstacle configurations we find that some have large lakes, where others have none, some have large clusters of obstacles, and others are sparse. It is clear how these factors could play a large role on the diffusion coefficient (further discussed in Chapter 3), and thus choosing a suitable ensemble size Ω is important.

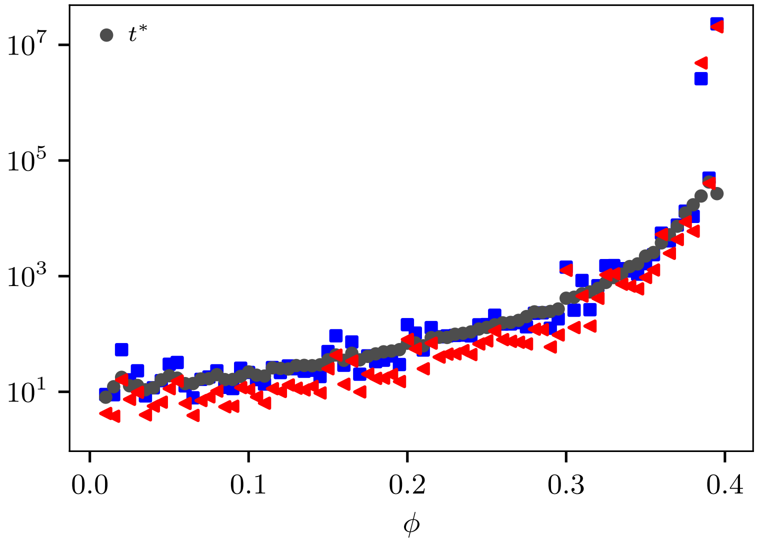
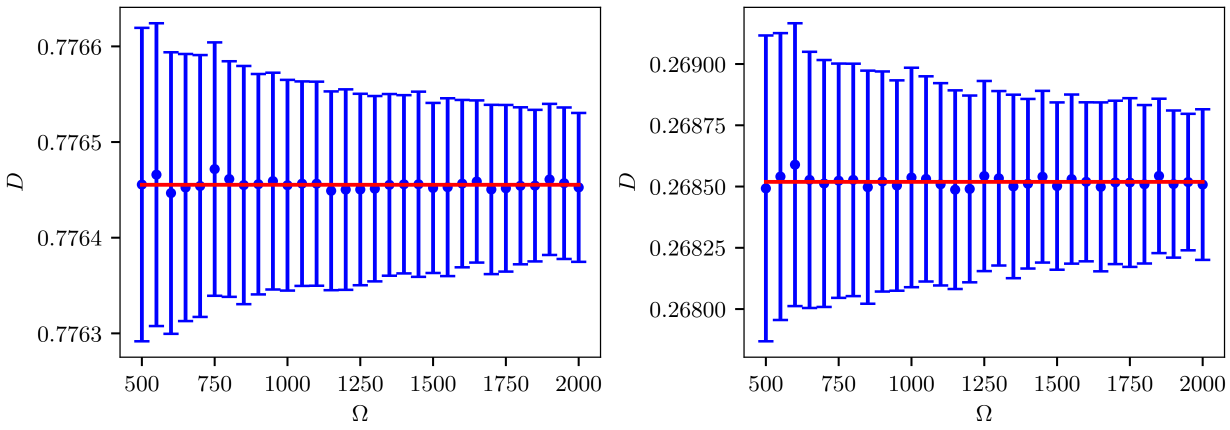


Figure 2.9: A comparison between the value of the cross-over time t^* (black circles) and our width estimates Σ_t of the transient regime as a function of obstacle concentration ϕ. The data points (black circles) correspond to random systems with a walker of size 1×1 and obstacles of size 1×1. The blue squares use Σ_α from eq. 2.41, and the red left pointing triangles use Σ_α from eq. 2.42.

Understanding how parameters we extract from the Diffusivity Ratio, as well as the steady-state diffusion coefficient, scale as a function of the ensemble size Ω, and their error bars, will allow us to justify our choice of ensemble size throughout this thesis ($\Omega = 2000$ in most cases, with some exceptions using $\Omega = 1000$).

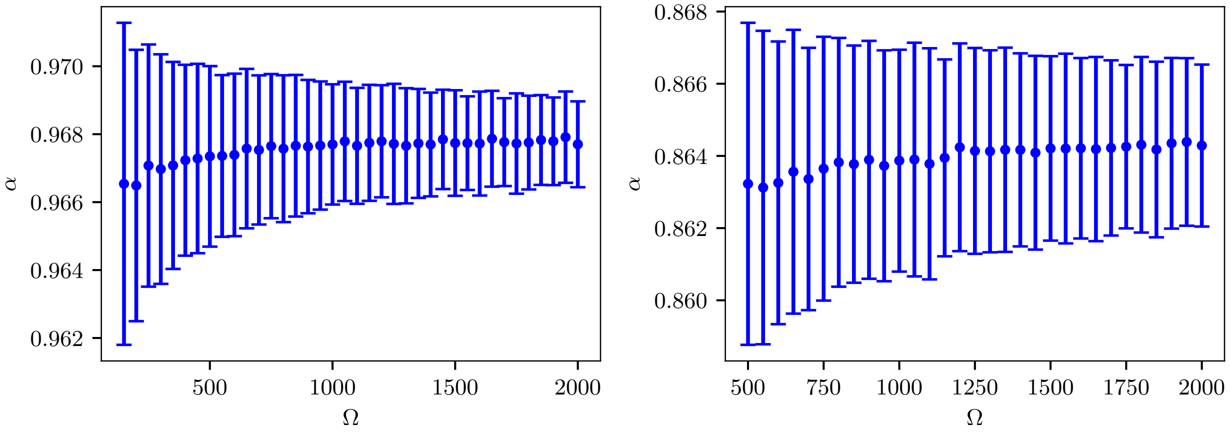
In Fig. 2.10 we compute the mean steady-state diffusion coefficient D (using the method from Section. 2.3) for different ensemble sizes of random obstacle configurations varying in size (from $\Omega = 50$ to $\Omega = 2000$ on intervals of $\Delta\Omega = 50$) at an obstacle concentration $\phi = 0.1$ in Fig. 2.10 (a) and $\phi = 0.3$ in Fig. 2.10 (b) with a system size $L = 128$. We find that there is a negligible effect on the mean value of the steady-state diffusion coefficient, furthermore the error bars scale as $\Omega^{-1/2}$, a trivial result in the case of evaluating the standard error on a mean value (where the standard error is $^{\sigma}/_{\sqrt{\Omega}}$, and σ is the standard deviation). We conclude from the convergence in Fig. 2.10 and the error bar scaling and magnitude ($\frac{\Delta D(\phi=0.1,\Omega=2000)}{D(\phi=0.1,\Omega\rightarrow\infty)} \approx 1.4 \times 10^{-4}$ and $\frac{\Delta D(\phi=0.3,\Omega=2000)}{D(\phi=0.3,\Omega\rightarrow\infty)} \approx 1.1 \times 10^{-3}$) that for the steady-state diffusion coefficient an ensemble size of $\Omega = 2000$ is more than enough to ensure reliable data independent of obstacle concentration.



(a) (b)

Figure 2.10: The steady-state diffusion coefficient D as a function of the ensemble size Ω for an obstacle concentration $\phi = 0.1$ in (a) and $\phi = 0.3$ in (b). Each data point consists of its own set of obstacle configurations. The solid red horizontal line represents the mean value of all data points.

Next we look at the effects of ensemble size on the anomalous exponent α as obtained using MCMC simulations. We choose to investigate α as it directly influences the rest of the parameters we extract from the MCMC data (some examples being D_α, t^*, r^*, Σ_t). To retrieve a single value of α we analyse the mean Diffusivity Ratio for a large ensemble of obstacle configurations, and to capture the mean value and corresponding error bars we bootstrap[1]our data set. The ensemble size will naturally introduce noise in the Diffusivity Ratio and alter the position of the inflection point, thus affecting α. In Fig. 2.11 (a) we find that for $\Omega < 1000$ there is a small positive slope to the mean value of α, whereas the mean value stabilizes for $\Omega > 1000$. In Fig. 2.11 (b) the slope persists until $\Omega \approx 1250$. The slope is smaller than the error on α itself, and we find that $\Omega > 1000$ is sufficient to evaluate α independent of obstacle concentration as the difference in α when comparing $\Omega = 1000$ and $\Omega = 2000$ ensemble sizes is within error ($\alpha(\phi = 0.1, \Omega = 1000) = 0.968(1)$, and $\alpha(\phi = 0.1, \Omega = 2000) = 0.968(1)$, $\alpha(\phi = 0.3, \Omega = 1000) = 0.864(3)$, and $\alpha(\phi = 0.3, \Omega = 2000) = 0.864(2)$). Similarly to D the error bars scale as $\Omega^{-1/2}$, again a trivial result. We conclude that $\Omega = 2000$ is more than sufficient for the ensemble size, and $\Omega = 1000$ is suitable given the trade off in computational time for a minor difference in error $\frac{\Delta\alpha(\phi=0.3,\Omega=1000)}{\alpha(\phi=0.3,\Omega\to\infty)} \approx 3.6 \times 10^{-3}$ compared to $\frac{\Delta\alpha(\phi=0.3,\Omega=2000)}{\alpha(\phi=0.3,\Omega\to\infty)} \approx 2.5 \times 10^{-3}$, where the errors are the standard deviations of the set of values gathered by bootstrapping the data set.



(a)

(b)

Figure 2.11: The anomalous exponent α as a function of the ensemble size Ω. Random systems are simulated at: a) $\phi = 0.1$, b) $\phi = 0.3$. Each data set corresponds to 2000 obstacle configurations at the respective obstacle concentrations that are bootstrapped. Analysing the data at high obstacle concentrations with very low ensemble size is not possible; this is why the x-axis in (b) starts at $\Omega = 500$ instead of $\Omega = 150$, as in case (a).

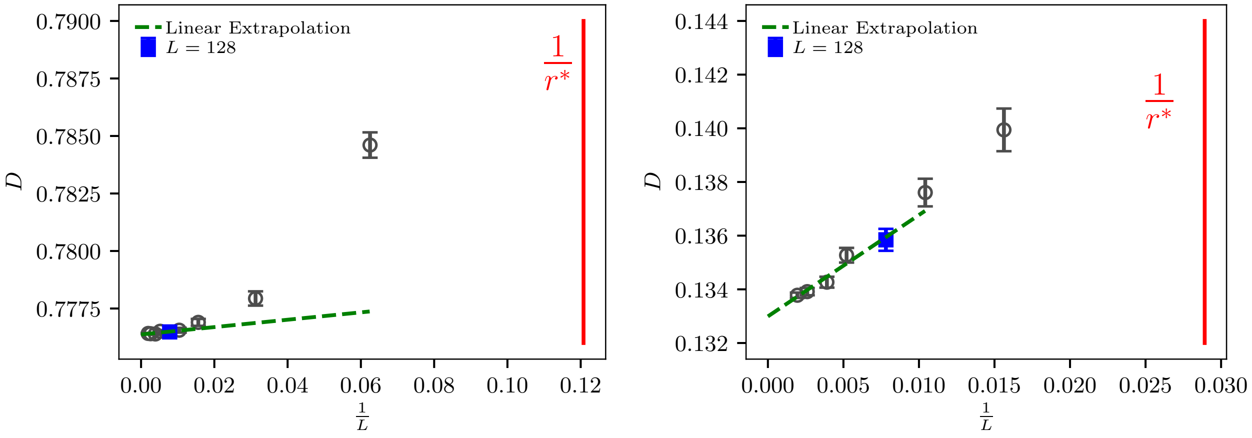
2.8 Finite Size Effects

Finite size effects are a result of the system size, and we can develop an understanding of their origin using the correlation length ξ (closely related to the crossover length r^* introduced in eq. 2.38). In percolation theory ξ is a critical length in the distribution of large cluster sizes $p(r) \propto e^{-r/\xi}$ [31]; thus ensuring $L \gg \xi$ is important to avoid cutting the tails of the cluster size distribution. In two-dimensional systems it is known that $\xi \propto |\phi^* - \phi|^{-4/3}$ whereas $r^* \propto |\phi^* - \phi|^{-4/3+5/72}$ [24], where ϕ^* is the percolation threshold which is $\phi^* = 0.407$ [14] for two dimensional square lattices with 1×1 obstacles. This shows ξ and r^* diverge in a similar way near percolation, we thus interchange the $L \gg \xi$ condition with $L \gg r^*$ (we can do this because the constraint only comes into

---

[1] Bootstrapping here refers to the process of randomly re-sampling our ensemble of 2000 data sets, 600 times (a standard choice according to Wilcox[38]). A re-sample size of 600 is further justified by studying the Gaussianity of the resulting parameters, as there no significant improvements with higher re-sample sizes. By doing this we are able to generate 600 values of each parameter we study (i.e. r^*, α, β) and calculate their standard errors.

play for large ϕ). It is important to note that finite size effects do not affect MCMC simulations and only apply to the steady-state diffusion coefficient calculation. This is because MCMC simulations do not reach the boundaries of the systems (any system that is simulated continues to appear infinite to the tracer particles, i.e. the tracer particles have no knowledge of cluster size truncation; instead finite size purely limits access to long time MCMC data).



(a) (b)

Figure 2.12: The mean steady-state diffusion coefficient D (ensemble size $\Omega = 2000$) as a function of the inverse system size $1/L$, for random systems of 1×1 obstacles (and walkers) at obstacle concentration (a) $\phi = 0.1$ and (b) $\phi = 0.35$. The inverse of the crossover length is shown as a solid red line, $r^*(\phi = 0.1) \approx 8.3$ and $r^*(\phi = 0.35) \approx 35$), while the $L = 128$ data point is shown as a blue square.

To test for finite size effect we investigate the behaviour of the steady-state diffusion coefficient D as a function of the inverse system size $^1/_L$ using ensembles of size $\Omega = 2000$. From Fig. 2.12 we indeed conclude that our system size needs to be much larger than r^* for an accurate calculation of the steady-state diffusion coefficient (the point nearest the vertical red line is significantly different than for large L). Chapters 3 and 4 show that r^* increases very quickly as a function of obstacle concentration, and that we have $r^*(\phi = 0.35) \approx 35$, for 1×1 walkers and obstacles. We expect that for concentrations below $\phi = 0.35$ the finite size effects will be much smaller as $r^*(\phi)$ diverges as ϕ approaches the percolation threshold ϕ^* [24]. Hence at lower concentrations we have $r^*(\phi < 0.35) \ll r^*(\phi > 0.35)$ leading to the system size better reflecting the $L \gg r^*$ requirement. We have chosen to use $L = 128$ proxy systems (the term proxy here refers to the difference between

$L = 128$ for the steady-state D calculation compared to $L = 512$ used for MCMC simulations) when evaluating the steady-state diffusion coefficient for a balanced trade off between computational time and accuracy at high obstacle concentrations. By extrapolating the data in Fig. 2.12 we find $lim_{L\to\infty} D(\phi = 0.35, L) = 0.1336$, and our data point shows $D(\phi = 0.35, L = 128) = 0.1358$ corresponding to an error of approximately 1.5%, whereas for a lower concentration $lim_{L\to\infty} D(\phi = 0.1, L) \approx 0.7763$ and $D(\phi = 0.1, L = 128) = 0.7764$ corresponding to an error of approximately 0.01% (the error bars from averaging are $\Delta D(\phi = 0.1, L = 128) = 1 \times 10^{-4}$ and $\Delta D(\phi = 0.35, L = 128) = 4 \times 10^{-4}$).

2.9 Comparing $p = \frac{1}{4}$ and $p = \frac{1}{8}$

We have discussed the probabilities of jumping along a given axis in Sections 2.2 and 2.1 ($p_{\pm x} = p_{\pm y} = p$ and $p_0 = 1 - 4p$), however proper justification for choosing $p = \frac{1}{8}$ over the standard hopping probability of $p = \frac{1}{4}$ has not been given. In this section we discuss the analytical repercussions of the $p = \frac{1}{4}$ choice, and we show that $p = \frac{1}{8}$ falls out of the math naturally. A similar analytical calculation was done by Neo Nguiya Passi [21], although specifically for $p = \frac{1}{4}$. We introduce the basics for the derivation, and extend it to a general value of p.

We begin by introducing the second moment of a random walker on a two dimensional square lattice as

$$\langle r_N^2 \rangle = \langle x_N^2 \rangle + \langle y_N^2 \rangle \tag{2.46}$$

where $\langle x_N^2 \rangle$ is the mean squared displacement along x after N total jumps on the lattice (not necessarily along x). This definition extends to $\langle y_N^2 \rangle$ analogously. The second moment along x is given by

$$\begin{aligned} \left\langle x_N^2 \right\rangle &= \left\langle p(x_{N-1} + a)^2 + p(x_{N-1} - a)^2 + (1-2p)x_{N-1}^2 \right\rangle \\ &= \left\langle x_{N-1}^2 \right\rangle + 2pa^2 , \end{aligned} \tag{2.47}$$

where a is the lattice spacing. It is important to discuss the $(1-2p)x_{N-1}^2$ term corresponding to the probability of not jumping along x: this probability of not jumping along x can be written as

$1-2p = p_0 + 2p$, where $p_0 + 2p$ is the sum of the probabilities of not jumping and jumping along y, i.e. the only options other than moving along x. Using eq. 2.47 recursively we obtain

$$\left\langle x_N^2 \right\rangle = 2pNa^2 + \left\langle x_0^2 \right\rangle \ , \tag{2.48}$$

where $\left\langle x_0^2 \right\rangle = 0$ since the starting position of the random walker is chosen to be $x = 0$. Since x and y are symmetric and independent we expect to find the same result for $\left\langle y_N^2 \right\rangle$. Using eq. 2.46 we can simply sum our results to find

$$\left\langle r_N^2 \right\rangle = 4pNa^2. \tag{2.49}$$

Comparing eqs. 1.14 and 2.49, we have

$$4pNa^2 = 4Dt \ . \tag{2.50}$$

Thus we are able to link the model parameters p,a, and τ to the diffusion coefficient:

$$D = \tfrac{pa^2}{\tau} \ , \tag{2.51}$$

where $\tau = {}^{t}\!/\!{}_{N}$ is the time step corresponding to a single hop. This simply comes from the fact that a time t passes after N equally temporally spaced hops are completed.

Similarly, the fourth moment is defined as

$$\left\langle r_N^4 \right\rangle = \left\langle x_N^4 \right\rangle + \left\langle y_N^4 \right\rangle + 2\left\langle x_N^2 y_N^2 \right\rangle . \tag{2.52}$$

The equation for the fourth moment in the x direction reads

$$\begin{aligned} \left\langle x_N^4 \right\rangle &= \left\langle p\left(x_{N-1}+a\right)^4 + p\left(x_{N-1}-a\right)^4 + (1-2p)x_{N-1}^4 \right\rangle \\ &= \left\langle x_{N-1}^4 \right\rangle + 12pa^2\left\langle x_{N-1}^2 \right\rangle + 2pa^4 \ . \end{aligned} \tag{2.53}$$

Applying the expression recursively we get

$$\left\langle x_N^4 \right\rangle = 2Npa^4 + 12N^2p^2a^4 - 12Np^2a^4 \; , \tag{2.54}$$

and the same result is found for y_N^4 again due to symmetry and independence.

In order to calculate the cross term $\left\langle x_N^2 y_N^2 \right\rangle$, we must take into account the actual number of steps attempted in each direction. Let's consider a particle that stays on the same site during n_0 of the N time steps, but takes n steps along x, and $N-n-n_0$ steps along y. The cross term is an ensemble average over all possible values of n and n_0:

$$\begin{aligned}
\left\langle x_N^2 y_N^2 \right\rangle &= \left\langle \left\langle x_n^2 y_{N-n-n_0}^2 \right\rangle_{ensemble} \right\rangle_{\{n,n_0\}} \\
&= \left\langle na^2(N-n-n_0)a^2 \right\rangle_{\{n,n_0\}} \\
&= \left[N\left\langle n \right\rangle_n - \left\langle nn_0 \right\rangle_{\{n,n_0\}} - \left\langle n^2 \right\rangle_n \right] a^4 \; .
\end{aligned} \tag{2.55}$$

The probability of a particle making n of the N jumps along x is given by the binomial distribution

$$q_N(n) = \frac{N!}{n!(N-n)!} (2p)^n (1-2p)^{N-n} . \tag{2.56}$$

Using

$$\left\langle n^j \right\rangle_n = \sum_{n=0}^{N} q_N(n) n^j \tag{2.57}$$

we find that

$$\left\langle n \right\rangle_n = 2pN \tag{2.58}$$

and

$$\left\langle n^2 \right\rangle_n = 2pN\,(1-2p+2pN). \tag{2.59}$$

To calculate $\left\langle nn_0 \right\rangle_{\{n,n_0\}}$ we need the probability of having n jumps along x *and* n_0 steps without moving in a set of N steps

$$q_N(n,n_0) = \frac{N!\,(2p)^{N-n_0}\,(1-4p)^{n_0}}{n!n_0!\,(N-n-n_0)!}, \tag{2.60}$$

where the probability of moving along x is $2p$ (p along $+x$, and p along $-x$), the probability of moving along y is thus also $2p$, and the probability of not moving is $1-4p$. Averaging the product nn_0 over all combinations of n and n_0,

$$\langle nn_0 \rangle_{\{n,n_0\}} = \sum_{n=0}^{N} \sum_{n_0=0}^{N-n} q_N(n,n_0)\, n\, n_0 \ , \tag{2.61}$$

gives

$$\langle nn_0 \rangle_{\{n,n_0\}} = 2p(1-4p)(N^2-N) \ . \tag{2.62}$$

The cross term is thus given by

$$\begin{aligned} \left\langle x_N^2 y_N^2 \right\rangle =& \{N(2pN) + 2(-p+4p^2)(-N+N^2) \\ & - \left[2p(N-2pN+2pN^2)\right]\}a^4 \\ & = (4N^2p^2 - 4Np^2)a^4 \ . \end{aligned} \tag{2.63}$$

This leaves us with

$$\left\langle r_N^4 \right\rangle = 32N^2p^2a^4 + 4Npa^4 - 32Np^2a^4 \ ; \tag{2.64}$$

using eq. 2.50 we then find that

$$\left\langle r^4(t) \right\rangle = 32D^2t^2 + (4-32p)Dta^2 = 32D^2t^2 \left(1 + \frac{4-32p}{8} \cdot \frac{a^2}{4Dt}\right) \ . \tag{2.65}$$

It is easy to see that $p = \frac{1}{8}$ will cause the correction term to be 0, whilst $p = \frac{1}{4}$ leads to:

$$\left\langle r^4(t) \right\rangle = 32D^2t^2 \left(1 - \frac{a^2}{8Dt}\right) \ . \tag{2.66}$$

It is important to note that for long times (i.e. $t \gg a^2/8D$) the correction term in eq. 2.66 is negligible.

The correction term corresponds to a derivation assuming free diffusion, thus we can estimate its effect only until the first collision with an obstacle; the nearest obstacle distance is $q \approx a/\sqrt{\phi}$ and the corresponding time is $t_1 \approx {}^{q^2}\!/_{4D} = {}^{a^2}\!/_{4D\phi}$. Substituting the time to first collision into eq. 2.66

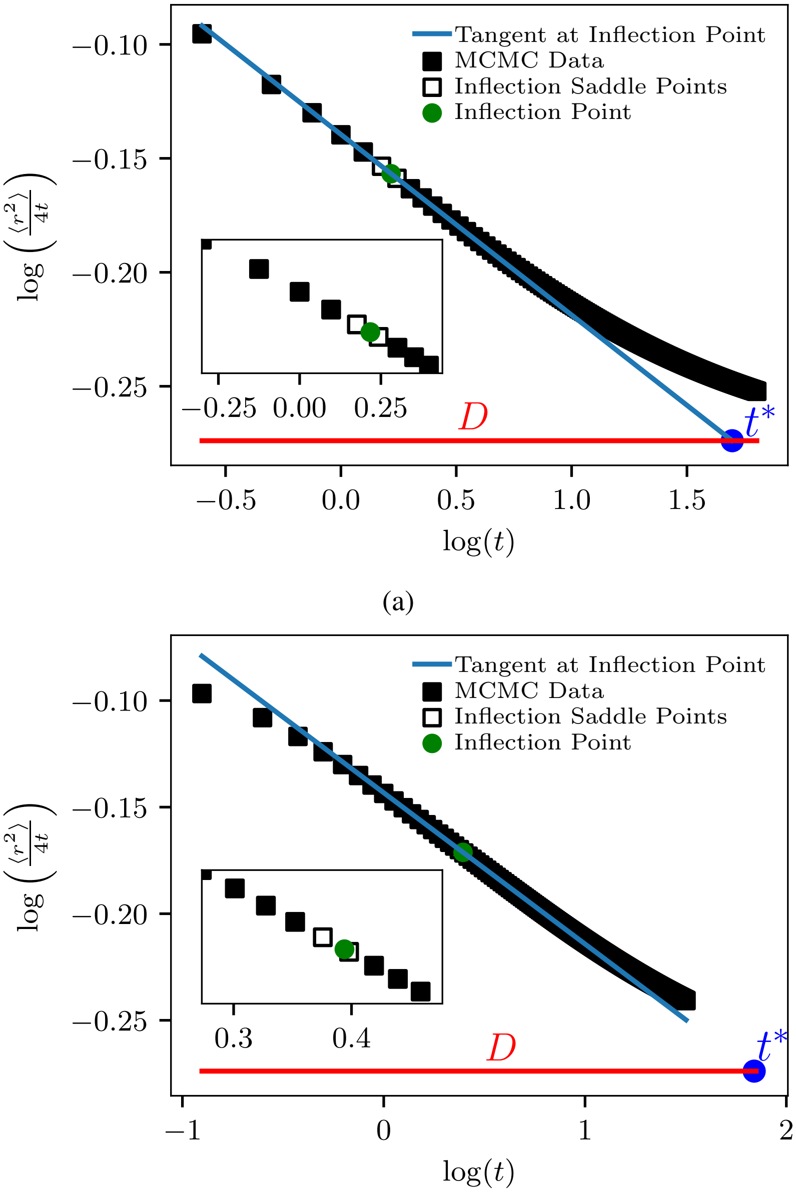
$$\left\langle r^4(t_1) \right\rangle \approx 32 D^2 t_1^2 \left(1 - \frac{\phi}{2}\right) \,, \tag{2.67}$$

we see that the correction term is dependent on the obstacle concentration ϕ. However we expect the correction term to be negligible in the limits $\phi \to \phi^*$ (particle will very rapidly collide with an obstacle) and $\phi \to 0$ (the time to collision is long enough for the correction term to be negligible).

Inevitably moments higher than the fourth moment also have an impact on dynamics, however in two-dimensions only allowing hops of length a along x and y we may only optimize one other moment beyond the second moment. It is possible to include diagonal jumps, and jumps of length longer than a to optimize higher moments in free diffusion, but it is unclear how diagonal jumps or longer jumps should interact with obstacles. Although there are many schemes that can be implemented to handle diagonal jump interaction with obstacles, there is no clear method to test these schemes; thus we cannot choose a "correct" scheme.

The difference between the two values of p can be directly investigated in the context of our work. Significant differences can be seen when examining the Diffusivity Ratio *vs.* time plots in Fig. 2.13. The most obvious difference is the position of the interpolated inflection point (green circles). In Fig. 2.13 (a) we see the inflection point is earlier in time in comparison to Fig. 2.13 (b); this is also illustrated in Fig. 2.14 where the sign change of the second derivative occurs earlier for the $p = \frac{1}{4}$ data set (black circles) compared to the $p = \frac{1}{8}$ data set (blue squares). This will have a direct effect on the values of different parameters, as we shall see next. When comparing the error bars of the results for $p = \frac{1}{4}$ and $p = \frac{1}{8}$ in Table 2.1, we see that this behaviour does not agree with our prediction in eq. 2.67. It is unclear why this is the case.

We now look at some of the parameters we extract and how they differ for identical systems when using $p = \frac{1}{4}$ compared to $p = \frac{1}{8}$. The key differences between $p = \frac{1}{4}$ and $p = \frac{1}{8}$ in Table. 2.1 are seen in the position of the inflection point t_I, and the anomalous diffusion coefficient D_α. The inflection point is affected both by the difference in the fourth moment as it effects the dynamics



(a)

(b)

Figure 2.13: The Diffusivity Ratio is shown as a function of time using two different MCMC algorithms for random systems consisting of 1×1 obstacles and walker with obstacle concentration $\phi = 0.2$. a) The standard set of LMC probabilities are used in the MCMC algorithm, i.e. $p_{\pm x} = p_{\pm y} = \frac{1}{4}$ with $p_0 = 0$. b) The probabilities chosen for our MCMC algorithm are $p_{\pm x} = p_{\pm y} = \frac{1}{8}$ with $p_0 = \frac{1}{2}$.

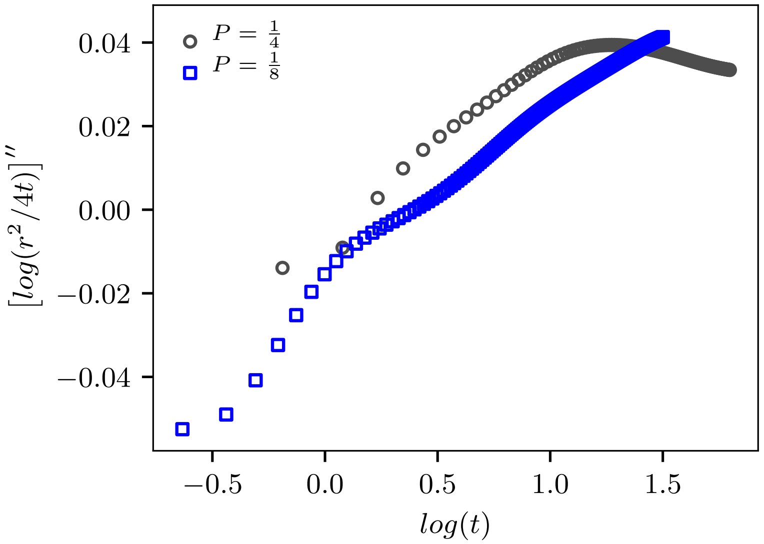


Figure 2.14: The numerical second derivative of the Diffusivity Ratio (outlined in Section 2.5) is plotted as a function of $log(t)$ for the two cases shown in Fig. 2.13. The two MCMC schemes ($p = \frac{1}{4}$, $p_0 = 0$ and $p = \frac{1}{8}$, $p_0 = \frac{1}{2}$) are shown simultaneously for comparison.

of diffusion when obstacles are present on the lattice, as well as the time resolution being less refined when $p = \frac{1}{4}$ is chosen. We find a difference in D_α as a result of the difference in t_I since D_α depends on t_I (D_α is discussed in Section 2.5). The similar values for t^*, r^*, and α in general support our choice to evaluate parameters at the inflection point as discussed in Section 2.5.

One practical drawback to using $p = \frac{1}{8}$ rather than $p = \frac{1}{4}$ is access to very late time data. We are limited in time by the size of our systems (in MCMC simulations we cannot reach the boundary), and thus the number of Markov steps. The time-step using $p = \frac{1}{8}$ is smaller than $p = \frac{1}{4}$ given eq. 2.51. Thus the maximum time $t = N\tau$ is smaller for the $p = \frac{1}{8}$ scheme. However the data is more time resolved. The time axes in Fig.2.13 corroborate this as Fig. 2.13 (b) has data in the approximate range $-1 < log(t) < 1.5$ and Fig. 2.13 (a) has data in the approximate range $-0.5 < log(t) < 1.9$. The consequence of lacking long time data is rather insignificant in the work this thesis presents, as the primary use for late time Markov data is to extrapolate the plateau of the Diffusivity Ratio for an estimate of the steady-state diffusion coefficient. Instead we use a numerical method for calculating the steady-state diffusion coefficient in Section 2.3.

In general the error bars using $p = \frac{1}{8}$ are smaller than those of $p = \frac{1}{4}$ (see Table 2.1): access

$p = 1/4$	$\phi = 0.1$	$\phi = 0.2$	$\phi = 0.35$
t^*	22(4)	$7(1) \times 10^1$	$2.2(2) \times 10^3$
r^*	8(1)	12(1)	34(2)
α	0.969(2)	0.926(5)	0.81(3)
D_α	0.91(3)	0.75(1)	0.57(1)
t_I	1.1(1)	1.7(1)	14.2(1)

$p = 1/8$	$\phi = 0.1$	$\phi = 0.2$	$\phi = 0.35$
t^*	22(2)	69(5)	$2.1(2) \times 10^3$
r^*	8.2(4)	12.0(5)	35(2)
α	0.968(1)	0.928(2)	0.81(2)
D_α	0.862(3)	0.714(5)	0.579(7)
t_I	1.14(6)	2.47(6)	21.77(6)

Table 2.1: Comparing transient regime parameters obtained when MCMC is implemented using either $p = \frac{1}{4}$ or $p = \frac{1}{8}$ (the value of p can be identified using the top left corner of the table) for random obstacle configurations and various ϕ. The numbers in parentheses at the end represent error bars on the last digit, e.g. $12(1) = 12 \pm 1$. The error bars are generated through bootstrapping our data sets.

to earlier times with higher precision around the inflection point is an asset for our analysis of the anomalous regime. It is clear that $p = \frac{1}{8}$ is a better choice compared to $p = \frac{1}{4}$.

2.10 Creating and Characterizing Obstacle Configurations

We have mentioned the presence of obstacles on our two dimensional square lattice, but not how the obstacle configurations are generated. Random and periodic obstacle configurations have been studied before [24, 29, 30, 31], however the region in between random and periodic, which we call "Fuzzy", has not. We wish to explore this region because there are clear qualitative differences in the diffusive behaviour between random and periodic systems; thus we expect some transition from one behaviour to the other when tuning the amount of disorder. These differences and effects are explored in Chapter 3. In this section we build the basis for generating tunable disorder and comparing disordered systems generated in different ways.

Important parameters used in this section include: the surface area occupied by obstacles ϕ, the system size L (i.e. the number of lattice sites in a column or row of the lattice), the ensemble size Ω (number of obstacle configurations used to average over disorder), and the first neighbour distance in a periodic system $r_p(\phi)$ (with obstacle concentration ϕ).

When generating disordered systems using different methods the natural description of disorder varies with the parameter used to control the disorder (e.g., the fraction of obstacles that are moved, the width of a distribution function for placement, or an acceptance probability). Therefore we need a way to compare the results with a single (universal) parameter that quantifies disorder. We propose to introduce the disorder parameter $0 \leq F_z \leq 1$, where $F_z = 0$ corresponds to a periodic system, and $F_z = 1$ is random. To define F_z we must first define $s(r)$, the number of sites that are within a distance r of a given obstacle. The first neighbour distance of obstacles in a periodic system with obstacle concentration ϕ is $r_p = \frac{a}{\sqrt{\phi}}$ ($a = 1$ is used throughout). We will use this as an upper limit of distance in our definition of F_z to ensure $F_z = 0$ when a system is periodic. Next we define $n(r)$ as the number of obstacles within a distance r of a given obstacle (excluding the latter).

We can now define F_z:

$$F_z = \frac{\left\langle \left\langle n(r < r_p) \right\rangle_{obstacles} \right\rangle_{ensemble}}{\phi \; s(r < r_p)} \;, \tag{2.68}$$

where the denominator corresponds to the number of obstacles we expect to be within a distance $r < r_p$ if the obstacles were placed randomly. The first set of angled brackets represent an average taken over all obstacles in a given obstacle configuration, and the second set of angled brackets corresponds to an average over the ensemble of obstacle configurations studied. It is clear then that if we apply this definition to a random system we expect $\left\langle \left\langle n(r < r_p) \right\rangle_{obstacles} \right\rangle_{ensemble} = \phi s(r < r_p)$ and thus $F_z = 1$. When we apply it to a periodic system $n(r < r_p) = 0$ since all other obstacles are at least a distance r_p away from any given obstacle, resulting in $F_z = 0$. With F_z in our arsenal, we can discuss how obstacle configurations are generated (and characterized) and their respective natural disorder control parameters.

2.10.1 Random Systems

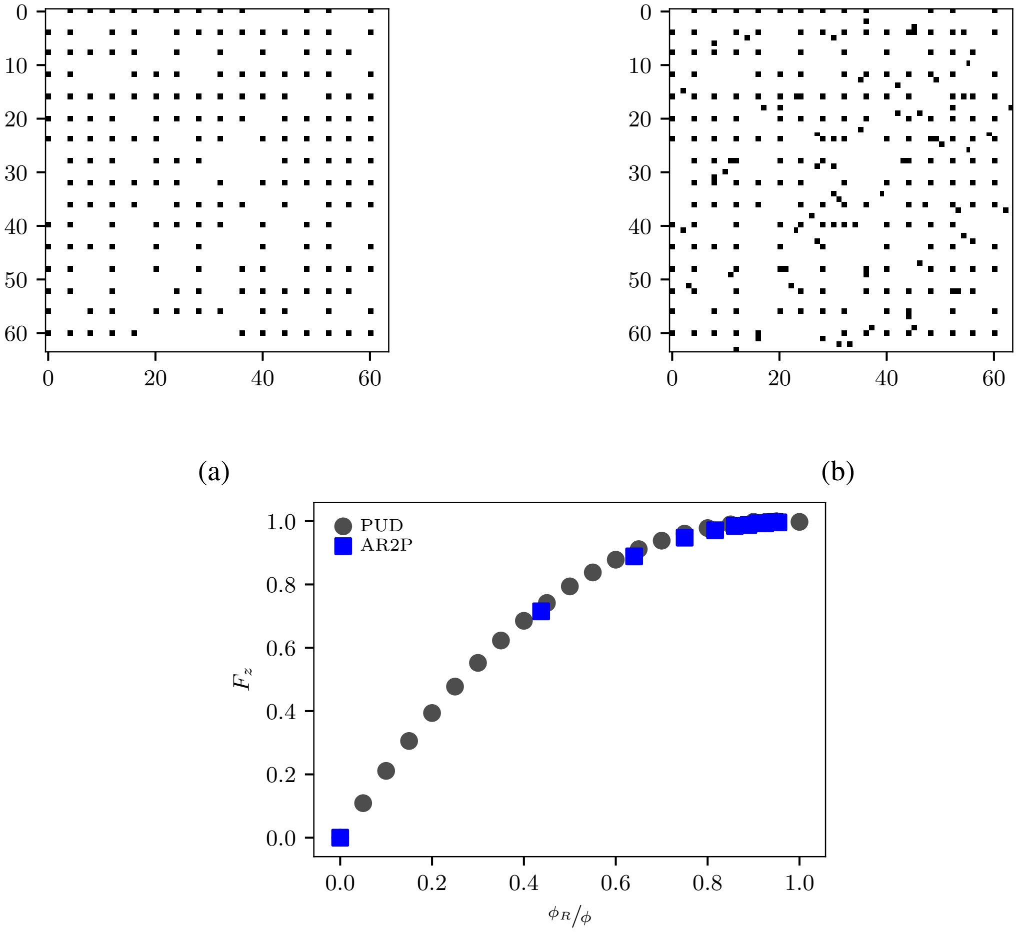
There are two ways to approach building a random system. One option is the grand canonical ensemble: in this case we treat the obstacle concentration ϕ as an explicit probability of placing an obstacle on a given site. Going through every site we determine whether an obstacle is placed at that site by generating a random number from the uniform distribution such that $0 \leq r \leq 1$, if $r < \phi$ we place an obstacle. As a result of this method any obstacle configuration can have an obstacle concentration ϕ within the range $\phi - \varepsilon \leq \phi \leq \phi + \varepsilon$ where ε is the deviation from the desired obstacle concentration. The deviation ε is naturally a function of the system size L as the number of lattice sites is the number of random numbers we draw from the uniform distribution, and by Central Limit Theorem we expect $\lim_{L \to \infty} \varepsilon = 0$ where $\varepsilon \propto \frac{1}{L}$. The alternative approach is to place a fixed number of obstacles at the desired obstacle concentration in every single obstacle configuration that we average over. In this case we simply fill the lattice with the desired number of obstacles and take a random permutation of them within the lattice. In practice we use NumPy [10], a numerical package for the programming language Python [33], to generate a one dimensional array of size L^2 (permutations of two-dimensional arrays are not possible). We populate the first N entries of the array with obstacles, where $N = \phi \times L^2$. From here we simply use Numpys permutation function to generate randomness.

2.10.2 Periodic Systems

There are many ways to generate periodic lattices, for example a body centred cubic lattice, a skewed square, or a simple square lattice. We will only investigate a simple square lattice, as we require isotropy in our systems to ensure the diffusivity along the x and y axes are the same. When generating a periodic system it is important we make proper use of PBCs. For instance, we need the system size to be a multiple of the period at which we will place obstacles. This is simply done by taking $L_{New} = floor(\frac{L}{r_p}) \times r_p$. We can then place obstacles with spacing r_p throughout the lattice of size $L_{New} \times L_{New}$.

2.10.3 Pick Up and Drop (PUD) Fuzzy Systems

To generate a Pick Up and Drop system we first begin with a periodic square lattice of obstacles at a desired obstacle concentration ϕ. As the name suggests, we will simply "Pick Up", i.e. remove, a fraction of the total obstacles, and then place them back onto the lattice randomly. In practice we remove obstacles one at a time at random until the desired fraction of obstacles is removed. Afterwards we place these obstacles back onto the lattice one at a time randomly, until the lattice has been repopulated. We make this choice to ensure there is no obstacle overlap. The percentage of obstacles that are removed and replaced randomly is a clear metric of the disorder generated in the system. Figure 2.15 demonstrates the process visually. Furthermore using our definition of F_z, we see it is a non-linear function of the percentage of replaced obstacles; we will see the repercussions of this interesting result in Chapter 3.



(a)

(b)

(c)

Figure 2.15: The process of generating a Fuzzy System using the Pick Up and Drop method: a) The initial periodic obstacle configuration at concentration $\frac{1}{16}$ with 25% of its obstacles removed at random. b) The removed obstacles are replaced at random on the configuration seen in (a). c) The disorder parameter as a function of the fraction of periodic obstacles replaced randomly (black circles); interestingly, this is a non-linear function, and it spans the full range $0 \leq F_z \leq 1$. The blue squares represent the analogous curve using the *AR2P* method, both are shown for simple side by side comparison.

2.10.4 Add Random to Periodic (AR2P) Fuzzy Systems

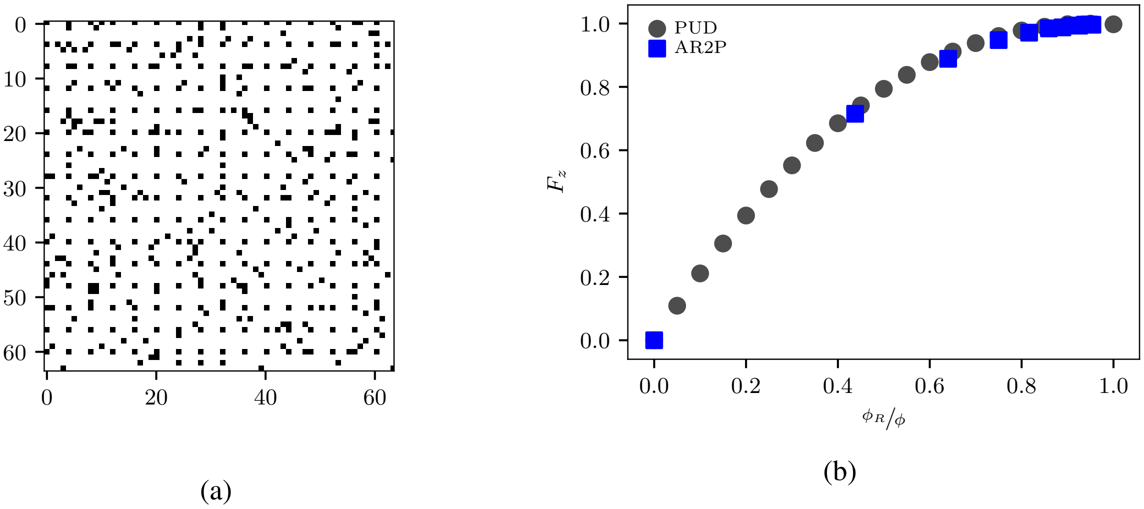
Arguably the simplest of the methods for creating disorder, AR2P consists of generating a "Fuzzy" system by adding obstacles randomly to an existing periodic system. Without using our disorder parameter, the natural metric for disorder in this system is the fraction of randomly placed obstacles to the total number of obstacles $R = {}^{\phi_R}/_{\phi}$, where $\phi = \phi_R + \phi_P$ and ϕ_P is the periodic obstacle concentration. An example system is shown in Fig. 2.16 (a) where $\phi_P = \frac{1}{16}$, $\phi = \frac{1}{9}$, and thus $\phi_R = \phi - \phi_P$, the underlying periodic lattice is clearly visible with random obstacles placed throughout. There are two types of problems we are interested in investigating:

1) systems with equal ϕ but different F_z values (examining the transition from the periodic limit to the random limit);

2) systems with similar F_z values but different obstacle concentrations ϕ (the concentration-dependent effects of disorder).

To achieve the same value of ϕ, and different F_z values, we simply change the periodicity of the base periodic lattice (i.e., we change ϕ_P), and adjust ϕ_R such that ϕ remains constant. This process is represented visually in Fig. 2.17 where constant ϕ lines are plotted in the $[\phi_R, \phi_P]$ phase space: any cross between a data point and a constant ϕ line corresponds to a different system at constant ϕ. Adjusting ϕ_R to satisfy a constant value of ϕ when varying ϕ_P will naturally cause F_z to change governed by the change in the concentration ratio R (this can be seen in Fig. 2.17, specifically looking the $\phi = 0.04$ we see an intersect with a red triangle ($0.6 < F_z < 0.8$) a green triangle ($0.4 < F_z < 0.6$) and a grey circle ($0 < F_z < 0.2$) giving us three systems in the phase space with varying F_z values yet identical overall obstacle concentration ϕ. Figure 2.16 (b) shows the non-linearity, and limitation of F_z values for $F_z(R, \phi = \frac{1}{9})$. The sparsity in low F_z values is a consequence of the requirement of a underlying periodic lattice, i.e. the lowest possible F_z value will always correspond to the closest periodic concentration ϕ_P to the desired concentration ϕ. Take for example $\phi = \frac{1}{9}$, the nearest periodic obstacle concentration (other than $\phi_P = \frac{1}{9}$ itself) is $\phi_P = \frac{1}{16}$, from which we find $\phi_R = \frac{1}{9} - \frac{1}{16} = \frac{7}{144} \approx 0.0486$ resulting in $R \approx 0.45$ and $F_z \approx 0.7$ (in fact it is the left most non-zero data point in Fig. 2.16 (b)).

For systems with varying obstacle concentrations ϕ but similar F_z we keep R constant. This is done by making ϕ_P an independent parameter, with a fixed value of R, and thus calculating the necessary value of ϕ_R. The result is $\phi_R = \frac{R\phi_P}{1-R}$. It is important to note that constant R curves are not quite equal to constant F_z curves since there is a ϕ dependent effect, however it is relatively small (see Fig. 2.17 where the teal dashed lines is a constant $F_z = 0.6$ curve).



(a)

(b)

Figure 2.16: The process of generating a Fuzzy System using the Add Random to Periodic method: a) A final system is shown with a underlying periodic system of periodicity 4 ($\phi_P = \frac{1}{16}$) with random obstacles overlaid to reach a total obstacle concentration of $\phi = \frac{1}{9}$. b) The disorder parameter as a function of the fraction of obstacles that are randomly placed within the system, the blue squares correspond to the data for *AR2P* whereas the black circles are for comparison to the *PUD* methods data. Non-linearity is present, with an asymptotic approach to $F_z = 1$ as ϕ_R begins to dominate ϕ_P in the total obstacle concentration $\phi = \phi_P + \phi_R$.

2.10.5 Harmonic Potential Wells (HPW) Fuzzy Systems

The base positions of obstacles in a periodic system can also be seen as the centres of Harmonic Potential Wells. There are two pieces to generating an obstacle configuration in this manner: we first choose which well we place an obstacle within; and then we determine where the obstacle is placed. We begin by choosing the position of a periodic lattice site (the centre of a well) at random. The position of this well is temporarily assigned $(x = 0, y = 0)$ (i.e. we recentre the lattice about the current well) and we generate a Gaussian probability distribution (representing

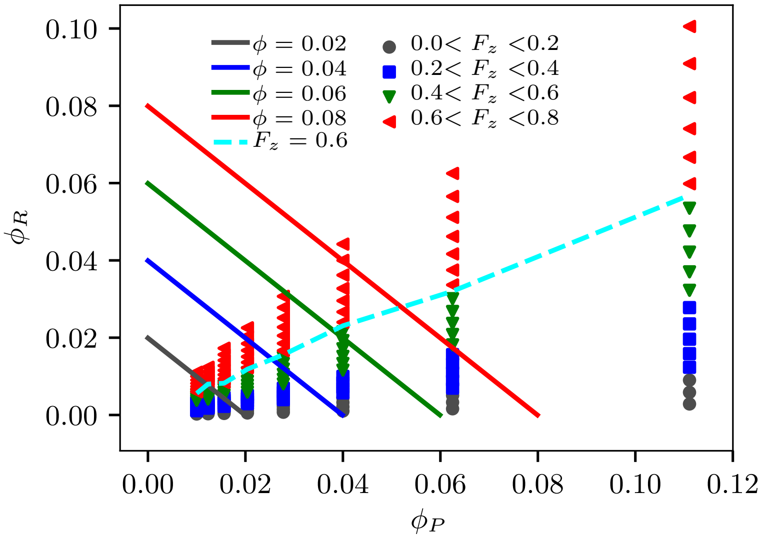
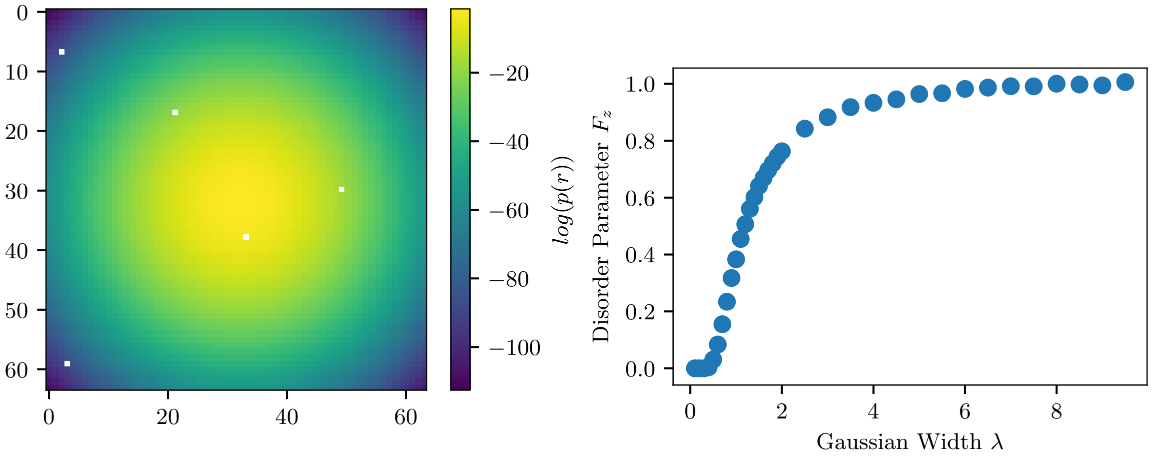


Figure 2.17: Different combinations of ϕ_R and ϕ_P are plotted in the ϕ_R-ϕ_P phase space. Each data point corresponds to a different ensemble of systems for which the mean value of the disorder parameter, $\langle F_z \rangle$, is calculated. We bin the data set into 4 groups to show approximate equi-F_z boundaries with an example curve shown as a teal dashed line (note that equi-R curves would be linear with a positive slope on this figure). Constant ϕ curves are shown as solid lines to illustrate the many combinations of systems that we can create at constant ϕ (overlaps between solid lines and data points).

the potential of the well) across the lattice $g(x,y) = \exp\left[-(x^2+y^2)/2\lambda^2\right](1-O(x,y))$, where $O(x,y) = 1$ represents the presence of an obstacle at (x,y). Next we normalize $g(x,y)$ such that $P(x,y) = g(x,y)/\sum_x \sum_y g(x,y)$. Then we choose a lattice site according to $P(x,y)$ to place the next obstacle. In theory we could choose to place multiple obstacles per well, however we will only explore this method with a single obstacle per well. This process of choosing a well, generating the normalized probability distribution given the centre of the well and the updated obstacle configuration, then placing an obstacle accordingly is repeated until every well has had exactly one obstacle placed within it. Figure 2.18 (a) illustrates the normalized probability distribution $P(x,y)$ for a well in the centre of the system after a few obstacles have already been placed.

The natural metric for disorder here is λ, where λ is the width of the Gaussian. There are clear limits that correspond to periodic $\lambda \to 0$ (e.g. all obstacles will be placed in the centre of a well) and random $\lambda \to \infty$ (e.g. all obstacles will be placed randomly throughout the system) systems. Figure 2.18 (b) shows the non-linear behaviour of F_z as a function of λ. We can see qualitative changes at $\lambda \approx \frac{1}{2}$ (fast increase) and $\lambda \approx 1$ (curvature changes sign) and the value of F_z begins to asymptotically approach 1. Furthermore when $\lambda > 4$ we have nearly reached the random limit; this corresponds to the mean displacement of an obstacle being larger than the first neighbour distance between well centres, thus making it impossible to differentiate which well centre was used to generate the position of an obstacle, in turn making the system approach an entirely random system.



(a)

(b)

Figure 2.18: The process of generating a "Fuzzy" system using the Harmonic Potential Wells method at an obstacle concentration of $\phi = \frac{1}{16}$: a) The Gaussian distribution due to a well located in the centre of the lattice with $\lambda = 2$ (the white spaces correspond to the positions of previously placed obstacles, i.e. to $P(x,y) = 0$). b) The disorder parameter F_z is a non-linear function of the Gaussian width λ.

2.10.6 Periodic Tiling of Random Subsystems (PTRS) Fuzzy Systems

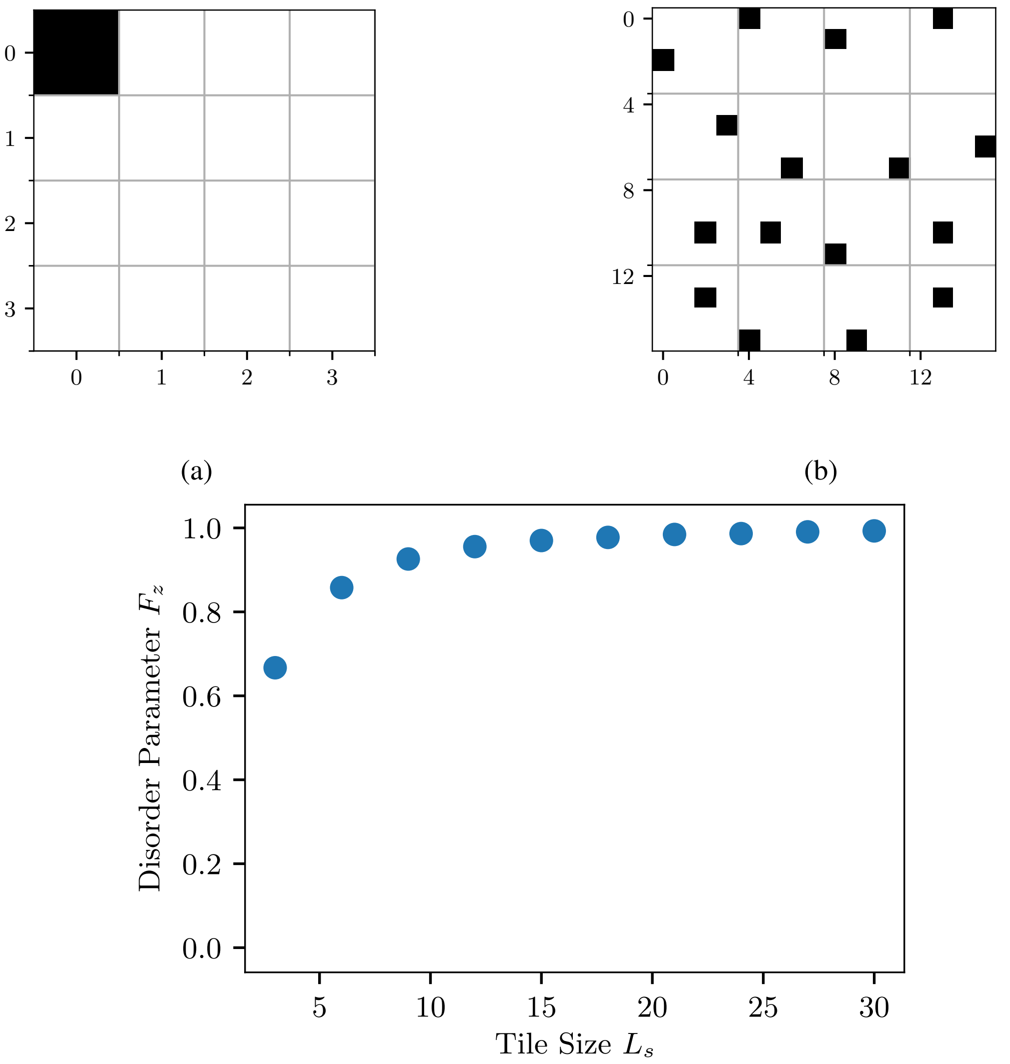
This method does not allow for the widest variety of F_z values as we will see in Chapter 3 and can be seen for $\phi = \frac{1}{9}$ in Fig. 2.19(c). However it does introduce an interesting constraint in the limitation of local concentration fluctuations. These "Fuzzy" systems are generated by first choosing a subsystem (or tile) size L_s, then a desired number of obstacles per subsystem N. The resulting obstacle concentration of one subsystem (as well as our entire lattice) is $\phi = \frac{N}{L_s}$, where the N obstacles are placed randomly within the subsystem. We then populate a larger lattice of size L, where L must be a multiple of L_s, with many sub lattices (each with N obstacles randomly placed within them) creating an overall system of size $L \times L$ with obstacle concentration $\phi = \frac{N}{L_s}$. An example of generating such a system is provided in Fig. 2.19 where Fig. 2.19 (a) corresponds to an example subsystem with $L_s = 4$ and $N = 1$; we then generate many random subsystems as in Fig. 2.19 (a) to create a $L = 16$ system in Fig. 2.19 (b).

The main line of thinking we follow to understand how disorder is generated in these systems is to examine the effect of changing L_s whilst holding $\phi = {}^{N}/_{L_s}$ constant. When L_s is small the fluctuations in the position of the N obstacles are limited and F_z tends towards the periodic limit, e.g. for $N = 1$ and $L_s = 2$ we find $\langle F_z \rangle \approx 0.6$ (this is the lowest possible F_z value for PTRS). As we increase N and L_s to keep ϕ constant F_z increases, e.g. $N = 16$ and $L_s = 64$ results in $F_z \approx 0.97$ (note that these example values have been averaged using $\Omega = 2000$ and $L = 512$); the general trend can be seen in see Fig. 2.19(c). When we approach the limit of the $L_s \approx r^*$ (where $r^*(\phi = \frac{1}{9}) \approx 9$) we reach the limit of a completely random system, hence the rapid saturation and limited values of $F_z(\phi = \frac{1}{9})$.

2.10.7 NoCluster (NoCR) Fuzzy Systems

The NoCluster method limits clustering by imposing an acceptance probability $0 \leq Q \leq 1$ for the placement of new obstacles adjacent to one or more obstacles. Note that Q has clear limits: $Q = 0$ imposes a hard-exclusion, i.e. no obstacles can be adjacent to one another, and $Q = 1$ yields completely random systems. This method is of interest as clustering plays an important role in percolation [31], thus varying Q allows us to investigate the role of clustering in the context of our work directly.

The simple computational approach to generating these systems would be to place obstacles randomly one by one until a desired obstacle concentration ϕ is reached. We would keep track of lattice sites that are adjacent to existing obstacles, and if a site is chosen to have an obstacle placed adjacent to an existing obstacle, a random number between 0 and 1 would be chosen to determine whether the obstacle is successfully placed. In the case where an obstacle placement is rejected,



(a)

(b)

(c)

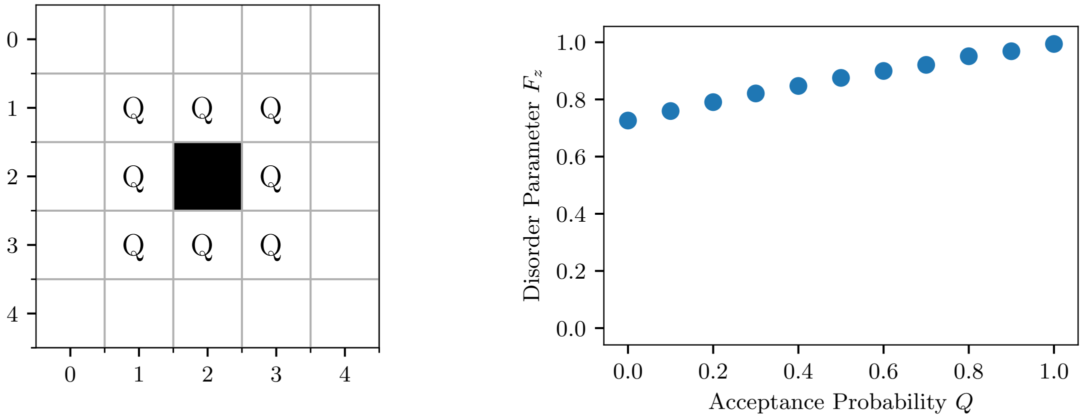
Figure 2.19: The process of generating a "Fuzzy" system using the PTRS method at a concentration of $\phi = \frac{1}{16}$: a) An example of a tile of size $L_s = 4$ with $N = 1$ obstacle. b) A $L = 16$ system is populated with many tiles of size $L_s = 4$ with $N = 1$. c) We see non-linear behaviour of F_z as a function of the sub lattice system size L_s at a fixed obstacle concentration $\phi = \frac{1}{9}$. We have chosen to use $N = 1$ and $L_s = 4$ in (a) and (b) to better illustrate the fluctuations, whereas $\phi = \frac{1}{9}$ is chosen in (c) to show a larger range of F_z values (this also correspond to results that are presented in Chapter 3).

the algorithm will choose a new position to attempt to place the same obstacle, and continue to do so until it is placed. An exception to this is when $Q = 0$, and no remaining lattice sites are not adjacent to obstacles. In this case the desired concentration ϕ cannot be reached, and thus the system is rejected, and the process begins anew.

The main problem with this "simple" algorithm is when Q is small and we attempt to generate a high obstacle concentration ϕ system because the majority of remaining lattice sites have adjacent obstacle already. Attempting to place an obstacle results in a loop where attempted placements are repeatedly rejected until a random number smaller than Q is drawn. In order to reach the desired obstacle concentration ϕ, the computation becomes unnecessarily computationally intensive.

In practice we have implemented a "smart" obstacle placement algorithm. In this scheme we initially assign a value of 1 to all lattice sites that are not adjacent to an obstacle. For all lattice sites adjacent to one or more obstacle we assign a value of Q, and all lattice sites occupied by an obstacle are assigned 0. We interpret this as a probability distribution, which can be normalized such that the sum is equal to 1. Placing obstacles according to this normalized probability distribution ensures that every obstacle placement is accepted. The drawback is that the probability distribution must be recalculated after every obstacle placement; however this is still more effective than the "simple" algorithm. Similarly to the "simple" algorithm, when $Q = 0$ it is still possible to reject a system; if all unoccupied lattice sites are adjacent to obstacles we will assign a value of 0 to all unoccupied lattice sites, meaning no more obstacles can be placed and the desired obstacle concentration cannot be met. In this case we will once again reject the system and begin again.

The placement of Q on lattice sites adjacent to an obstacle is shown in Fig. 2.20 (a). The value of 1 should be assigned to all sites that are not occupied by an obstacle or adjacent to an obstacle; we omit this to improve clarity of the figure. Figure 2.20 (b) shows that the disorder parameter is relatively linear in terms of Q compared to the rest of the natural disorder parameters for other methods. It should be noted however that Q is not a true disorder parameter as $Q = 0$ does not create periodic systems except at $\phi = \frac{1}{4}$ (the maximum concentration for the $Q = 0$ limit) and thus we cannot explore the entire "Fuzzy" region $0 \leq F_z \leq 1$.



(a)

(b)

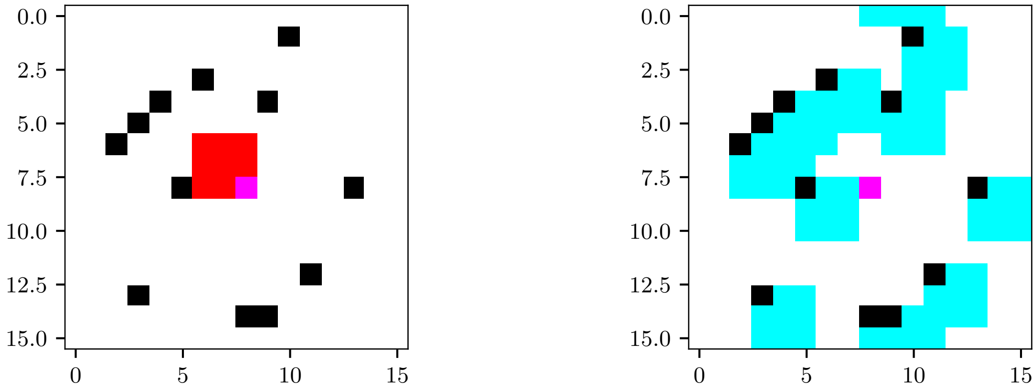
Figure 2.20: The process of generating a Fuzzy System using the No Cluster method: a) A singular obstacle is placed at the center of the lattice for demonstration purposes, and the acceptance probability Q is assigned to all lattice sites adjacent to 1 or more obstacles. b) This figure shows a relatively linear relationship between the disorder parameter F_z and the natural measure of disorder Q (the only method which does so).

2.10.8 Larger Obstacles and Walkers

In Chapter 4 we investigate the ability to estimate the steady-state diffusion coefficient using data only from the anomalous regime, specifically data near the inflection point. In this paper we extend our estimate from 1×1 obstacles and walkers, to a variety of obstacle sizes and walker sizes. As a result we must discuss in more detail how we create systems with larger obstacles, and how we treat larger walkers.

The creation of larger obstacles is simple: we place obstacles of size $l \times l$ one at a time on the lattice. We use the upper left corner of the obstacle as the anchor for generating the rest of the $l \times l$ obstacles. We reject the placement of an obstacle were there to be any overlap with another obstacle. Furthermore it is important that we think of our lattice not as a standalone system, but as a sample of an infinite size system with the desired obstacle concentration. Consequently we treat the edges of our systems as PBCs, such that if we place a 2×2 obstacle with our top left anchor flush with the right edge of our system, it will wrap around to the left side of the lattice.

We do not explicitly create walkers of size $l \times l$ with 1×1 obstacles, instead we consider that a larger walker has an excluded region around it which it cannot access. We can thus grow the obstacles to the size of the desired walker size such that they account for the exclusion zone of a larger walker, and replace the larger walker with a 1×1 walker. To do this we assign an anchor position to the walker, i.e. the position of our equivalent 1×1 walker, and the anchor position for obstacle growth is then the relative position of the walkers anchor mirrored along both axes; e.g. if the walker anchor is in the bottom right of the larger walker then the corresponding obstacle anchor position is the top left corner, i.e. the initial obstacle will be in the top left corner of the grown version in this example. When growing obstacles in this case we continue to allow periodic boundary conditions as we did when we created larger obstacles, however we do not impose the non overlapping criteria as before due to the nature of the exclusion zone of a large walker. An example is shown in Fig. 2.21 where Fig. 2.21 (a) shows an initial obstacle configuration in black, with a 3×3 walker shown in red and a pink site representing the walkers anchor position. In Fig. 2.21 (b) we see the anchor position in pink again which now acts as a 1×1 walker in a system with the initial obstacles in black, and the growth regions in teal, from which it is clear that the black obstacles are in the top left of their respective growth regions.



(a)

(b)

Figure 2.21: A visual guide to understanding the obstacle growth procedure when converting a larger walker system into an equivalent 1 × 1 walker system. a) A random obstacle configuration (with 1 × 1 obstacles shown in black) is shown, with an initial 3 × 3 walker at the centre of the lattice (shown in red), the bottom right corner is our walkers anchor site and is shown in pink. b) We grow the obstacles according to the exclusion region resulting from the 3 × 3 walker where the teal region is the growth region and the initial obstacle positions are shown in black. After we have grown the obstacles we have just the anchor position of the 3 × 3 walker remaining which now acts as a 1 × 1 walker. The result is independent of our choice of anchor position by a shift in the resulting obstacle configuration.

CHAPTER 3

OBSTRUCTED DIFFUSION IN SYSTEMS WITH TUNABLE DISORDER: WHEN DOES A FUZZY SYSTEM LOOK LIKE A PERIODIC (OR A RANDOM) ONE?

Obstructed Diffusion in Systems with Tunable Disorder:
When does a fuzzy system look like a periodic (or a random) one?

Nicholas Ilow[1] and Gary Slater[a)]
University of Ottawa, Department of Physics, Ottawa, Ontario, K1N 6N5, Canada

(Dated: 10 January 2022)

In real world cases such as biomembranes or drug delivery systems, obstructed diffusion never occurs in the presence of perfectly periodic or perfectly random distributions of obstacles. We present a study of transient and steady-state diffusion in ”Fuzzy” systems that bridge the gap between these two ideal limits. In particular, we examine whether there are ”diffusion phase transitions”, *i.e.*, abrupt quantitative and/or qualitative changes at some critical degree of disorder. For instance, while the crossover length r^* decreases when the concentration of obstacles ϕ increases in a periodic system, it is the opposite for random systems. The short-time anomalous diffusion parameters, the transition to the steady-state and the asymptotic diffusion coefficient D are studied as a function of both the degree of disorder and the obstacle concentration ϕ. Different ways of creating tunable disorder are explored, and a disorder parameter F_z is introduced to quantify disorder in a method-independent manner. One of our methods provides data which falls on a linear interpolation between the periodic and random limits for all parameters (i.e. r^*, β, D, α). Interestingly, we find that r^* can become a very weak function of ϕ in some fuzzy systems, and that a critical disorder parameter of $F_z \approx 2/3$ can be defined. We also propose a new connection between the properties of the transient and steady-state regimes.

[a)] gary.slater@uOttawa.ca

I. INTRODUCTION

How the diffusivity of a particle is impacted by obstruction has been explored extensively using Lattice Monte Carlo methods for random[1–5] and periodic[6–8] systems of immobile and passive obstacles. However real systems are neither periodic nor totally random[9]. The object of our work is diffusion in the intermediate "Fuzzy" regime between these two limits. We focus our attention on changes of qualitative behaviour to investigate how a system transitions from quasi-periodic to quasi-random behavior as we increase the amount of disorder in a controlled (tunable) way. As we shall see, our results (obtained for two-dimensional lattice systems) suggest non-trivial intermediate regimes.

There are three primary regimes for obstructed diffusion. Particles first undergo free Brownian motion with diffusion coefficient D_o. Once they start colliding with obstacles they enter a transient phase during which they slow down. Finally, after moving a critical distance r^* (or, equivalently, after a diffusion time t^*), called the *crossover length*, they reach the steady state with a diffusion coefficient $D < D_o$ and their mean square displacement (MSD) then scales linearly with time[10–12],

$$\left\langle r^2(t) \right\rangle = \beta^2 + 4Dt, \tag{1}$$

where β is the excess diffusion due to the fact that diffusion is actually faster in the first two regimes. In the first regime, we simply have $\beta = 0$ and $D = D_o$.

Because the MSD does not scale linearly with time in the transient regime[12–14], this regime is often described using the concept of anomalous diffusion and the relation

$$\left\langle r^2(t) \right\rangle = 4D_\alpha t^\alpha, \tag{2}$$

where α is called the anomalous exponent. Since obstacles hinder diffusion, we have $\alpha < 1$, *i.e.* subdiffusion. The parameter D_α is sometimes called the anomalous diffusion coefficient, a misleading name since its units depend on α. However, eq. 2 is rarely satisfied over a wide time interval; obtaining a reliable value for α is then challenging since the time interval used to fit the data is often chosen arbitrarily.

We study this problem by analysing the time-dependence of the average mean squared displacement $\langle r^2(t) \rangle$ of a tracer particle on a two-dimensional lattice with a wide range of obstacle configurations (a detailed description of the methodology can be found in Ilow and Slater[15]). The tracer particle is allowed to either stay still or hop along the x and y axes (diagonal jumps are excluded) of length a where a is the lattice spacing. In other words, we use a version of lazy random walk dynamics where $p = p_{\pm x} = p_{\pm y} = \frac{1}{8}$, and $p_0 = \frac{1}{2}$ is the probability of not moving. The fraction of the system's surface occupied by obstacles is denoted by ϕ. Short time dynamics is studied using a Markov Chain Monte Carlo algorithm that provides high-precision data. The steady state diffusion coefficient $D(\phi)$ is calculated using a numerical method outlined in ref.[16] (note that we fix $D_o = D(0) = 1$ for simplicity). We extract key parameters such as the crossover length r^*, the crossover time t^*, the excess diffusivity β, the anomalous diffusion coefficient D_α and the anomalous exponent α by plotting $log\left(\langle r^2\rangle/4t\right)$ $vs.$ $log(t)$, as described in ref.[15].

This manuscript will first introduce a universal disorder parameter F_z to connect different methods of generating disorder (we propose to study 5 unique methods) within obstructed lattice systems in Section II. We then study key diffusive and system properties such as r^*, β, D, and α in two contexts: Concentration dependent behaviour in Section III A where we investigate how these parameters change for fixed F_z values as a function of obstacle concentration ϕ, and F_z dependent behaviour in Section III B where a fixed concentration is chosen and the transition between periodic and random limits is studied.

II. CREATING AND QUANTIFYING DISORDER

We now describe the five methods we use to create systems that are between Periodic and Random. We characterize these systems using a "disorder parameter" $0 \leq F_z \leq 1$ defined as the ratio of the mean number of obstacles neighbouring an obstacle to the expected number in a random system. To ensure a periodic system returns a F_z value of 0, we define neighbouring as within a distance $< 1/\sqrt{\phi}$, which corresponds to the distance between periodically placed obstacles (if the latter form a simple square geometry – see Fig. 1). In practice we use a neighbouring distance $= 3$ for all concentrations $\phi < \frac{1}{9}$, and a neighbouring distance of $= 2$ for $\phi \geq \frac{1}{9}$. Note that some of the methods described below do not give access to the full range of F_z values, while others will use a smaller ensemble size compared to our standard $\Omega = 2000$ due to computational limitations. Figure 1 shows an example of the transition of a periodic system to a random system, generated using the method described in Section II A.

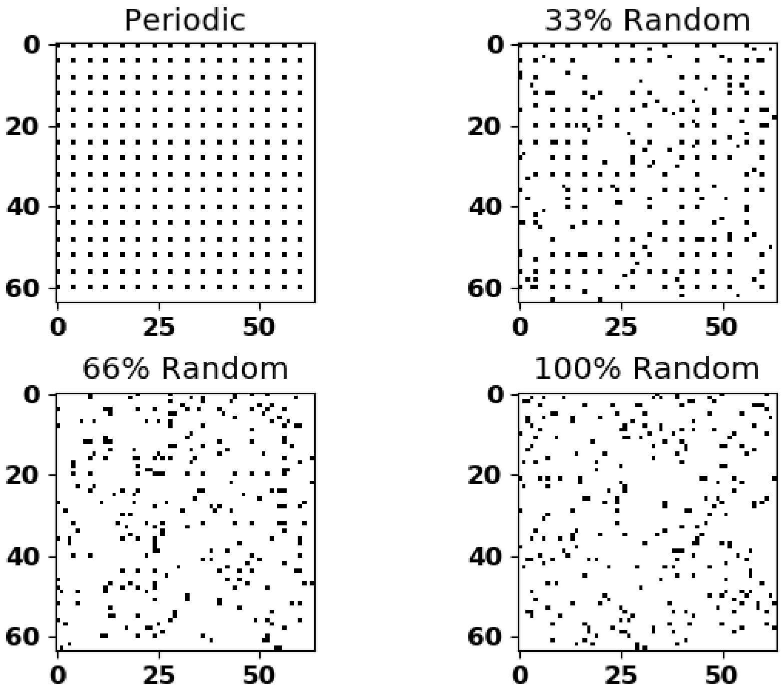


FIG. 1: Four 64×64 systems with an obstacle concentration $\phi = \frac{1}{16}$ showing the progression of *Pick Up and Drop* (PUD) algorithm, from periodic to random obstacle placement. For these cases we find $F_z(periodic) = 0$, $F_z(33\%) = 0.5984(5)$, $F_z(66\%) = 0.9196(7)$ and $F_z(random) = 0.9994(8)$.

A. Pick Up and Drop (PUD)

This algorithm begins with periodic obstacles at concentration ϕ. A fraction of the obstacles are then randomly moved to available lattice sites. This method lacks the ability to access intermediate concentrations since ϕ is not a continuous variable (moreover, $\phi \leq \frac{1}{4}$). An illustration of PUD is shown in Fig. 1 where the transition from a periodic configuration of obstacles to a completely random configuration of obstacles is shown. In Fig. 1 (a) the pre-existing periodic lattice of obstacles can be clearly seen, and in Fig. 1 (b) it becomes less clear due to the presence of disorder.

B. Adding Random Obstacles to Periodic Systems (AR2P)

Disorder can also be increased by randomly adding obstacles (concentration ϕ_R) to a system that is already populated with periodic obstacles at concentration ϕ_P. Since the same final concentration $\phi = \phi_P + \phi_R$ can be reached via different combinations of ϕ_P and ϕ_R, this method allows for multiple values of F_z for a given ϕ.

C. Harmonic Potential Wells (HPW)

Harmonic potential wells are periodically placed on the lattice with a concentration ϕ_P, and obstacles are then randomly "dropped" in these wells. The probability of accepting a placement in a well is given by the Gaussian distribution

$$P(r) \sim \exp\left(-r^2/2\lambda^2\right), \tag{3}$$

where r is the distance between the selected lattice site and the centre of the chosen potential well, and λ measures the width of the well.

The HPW method allows for a wide range of disordered systems since the obstacle concentration ϕ is decoupled from ϕ_P; for instance, we can place multiple ($\phi > \phi_P$) or less than one ($\phi < \phi_P$) obstacles per well. When $\phi = \phi_P$, $\lambda \to 0$ results in completely periodic systems. When $\lambda > 1/\sqrt{\phi_P}$, the obstacles cannot easily be traced back to their generating well, and at twice this spacing we obtain significant clustering as in random systems. An ensemble size $\Omega = 1000$ is used for these data sets.

D. Periodic Tiling of Random Subsystems (PTRS)

Here we build $L \times L$ tiles and randomly place N obstacles in each one. This results in an overall concentration $\phi = {}^{N}\!/_{L^2}$ and concentration fluctuations that are limited to distances $\approx L$. These tile systems are somewhat similar to the HPW method with $\lambda \approx L$. Once the tiles get larger than the crossover length for a random system at the given concentration, $L \gg r^*(\phi)$, L is irrelevant and we are in the $F_z = 1$ (random) limit. Note that arbitrarily small values of F_z cannot be generated for a fixed obstacle concentration ϕ using this method (e.g., at $\phi = {}^{1}\!/_{9}$, the smallest tiles correspond to $[N = 1, L = 3]$ and $F_z \approx 0.7$; larger tiles yield larger values of F_z).

E. No-Cluster Random (NoCR)

Obstacles in random systems can form clusters which lead to percolation issues. A short range "depletion" constraint can limit such features. We thus introduce the probability Q of accepting an obstacle placed adjacent to ≥ 1 obstacles. When $Q \to 0$, every obstacle is surrounded by a strict exclusion zone which eliminates clustering and reduces disorder. The $\phi \to 0$ limit for $Q \to 0$ appears random as clusters are already unlikely to form, while $\phi \to 1/4$ must be periodic since this is the maximum concentration allowed when we impose the $Q=0$ condition. The opposite, $Q=1$, accepts all configurations and is simply a random system. Due to computational limits, we use a smaller ensemble size of $\Omega=1000$ for this method.

III. RESULTS

We first investigate the transition from random behaviour to periodic behaviour by looking at the general qualitative features of D, r^*, α, and β as functions of the obstacle concentration ϕ when we vary the disorder parameter F_z. We then examine the "order-disorder transition" as F_z is varied at a fixed obstacle concentration ϕ. The steady-state diffusion coefficient $D(\phi)$ is an exact calculation per obstacle configuration[16] and averaged over an ensemble of obstacle configurations. In practice[1], we calculate the crossover length $r^* = \sqrt{4Dt^*}$ using the steady-state diffusion coefficient D and the crossover time t^* (the time at which the anomalous regime and steady-state regime yield the same displacement, mathematically $4Dt^* = 4D_\alpha {t^*}^\alpha$). We evaluate the excess diffusivity β at time t^* by taking the difference between the observed mean squared displacement and the predicted mean squared displacement ($= 4Dt^*$) had we been in the steady state.

A. Effects of Obstacle Concentration Keeping the Disorder Parameter Fixed

The diffusion coefficient D: As Fig. 2 shows, the curvature of the function $D(\phi)$ goes smoothly from negative when $F_z \to 1$ (the random limit) to positive when disorder decreases. The curvature is zero at $F_z \approx \frac{2}{3}$. Indeed, the data agrees nicely with the first two terms of the series expansion[17] (valid for all systems with 1×1 obstacles and walkers on a square lattice) $D(\phi) = 1-(\pi-1)\phi$ when $F_z \approx \frac{2}{3}$ (a similar critical value is found for the AR2P and HPW methods; data not shown).

The anomalous exponent: $\alpha(F_z)$ is difficult to investigate due to the difference between

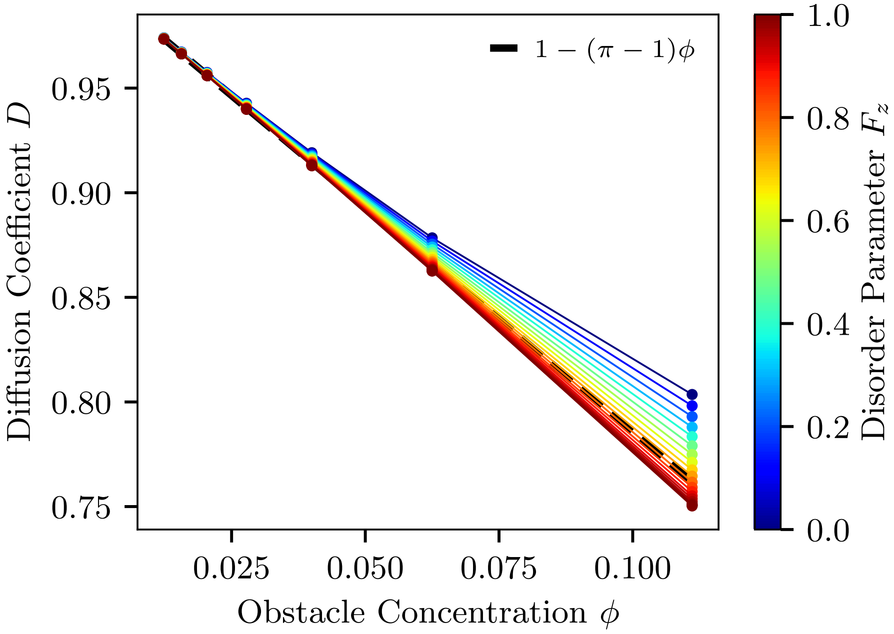


FIG. 2: Steady-state diffusion coefficient D *vs* obstacle concentration ϕ for several values of F_z as obtained when using the PUD method. The curves from top to bottom correspond to replacing $0 \rightarrow 60\%$ of the periodic obstacles randomly by steps of 5% (the bottom curve corresponds a fully random system or 100%).

$\alpha(F_z = 0)$ and $\alpha(F_z = 1)$ being very small ($< 0.05\%$) when $\phi \leq \frac{1}{9}$. As a result, small statistical (ensemble) fluctuations strongly affect the value of α and mask any effect disorder might have. This can be seen at low ϕ in Fig. 3 as all of the curves are essentially overlapping. The last few data points at higher concentrations show a change of curvature around $F_z \approx \frac{2}{3}$.

The crossover length r^*: Remarkably, for intermediate values $\frac{1}{2} \leq F_z \leq \frac{2}{3}$ in Fig. 4, r^* is a weak function of the obstacle concentration ϕ. This region thus separates fuzzy systems that are periodic-like (r^* decreases with ϕ) from those that are random-like (r^* increases with ϕ).

Excess diffusivity: Prior to discussing the results of β we first introduce a proxy value $\beta(t = t^*)$ as β by definition implies the $t \rightarrow \infty$ limit. We simply evaluate β at $t = t^*$ as long enough times are not accessible to properly estimate the $t \rightarrow \infty$ limit. We see a smooth qualitative transition present in Fig. 5. The bottom curve represents $F_z = 0$ (periodic systems) and we note the negative slope between the last two points at $\phi = \frac{1}{16}$ and $\phi = \frac{1}{9}$. The sign of this slope changes from negative to positive at $F_z \approx \frac{1}{3}$ (a similar value is found for the AR2P and PUD methods – data not shown). A second key transition is observed at $F_z \approx \frac{2}{3}$ where an inflection point is no longer present for

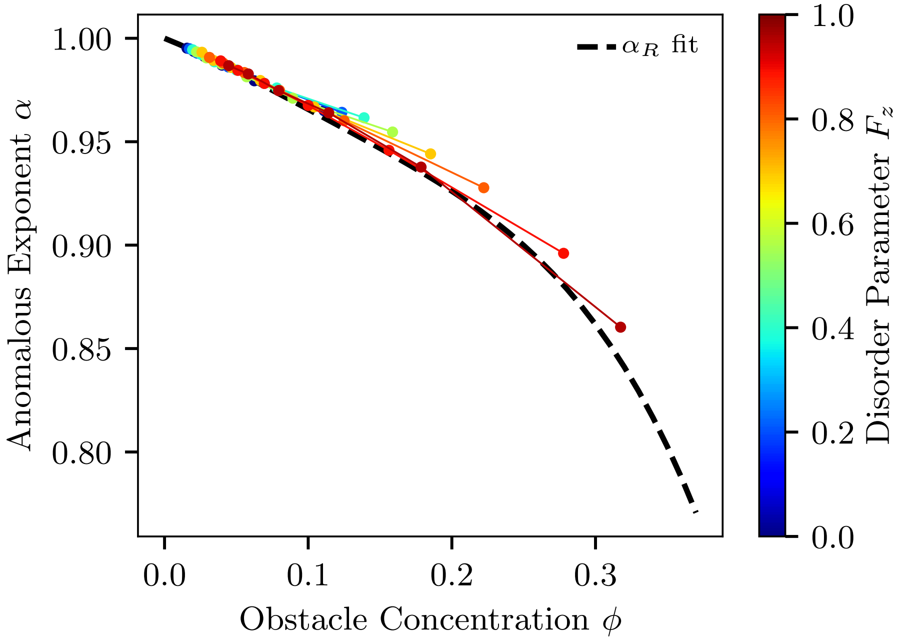


FIG. 3: Anomalous exponent α *vs* obstacle concentration ϕ when using the AR2P method. The curves from top to bottom correspond to the percentage of obstacles that are randomly placed within the system, on the interval of 0% to 60% by steps of 10%; the bottom curve is for 65%. The dashed line shows a fit of α for random obstacle configurations.

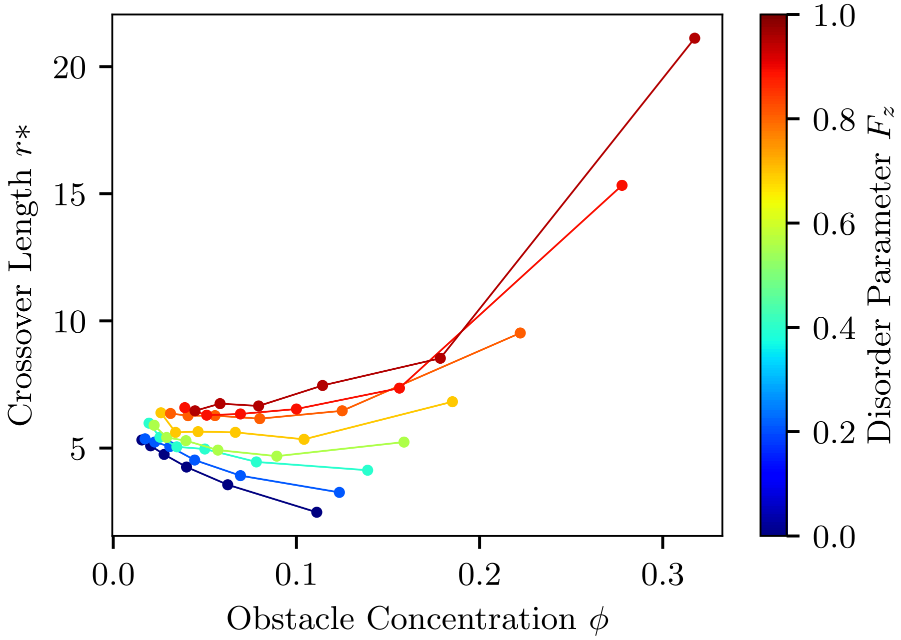


FIG. 4: Crossover length r^* *vs* obstacle concentration ϕ. See the caption of Fig. 3 for the details (note that the curves are ranked from bottom to top here).

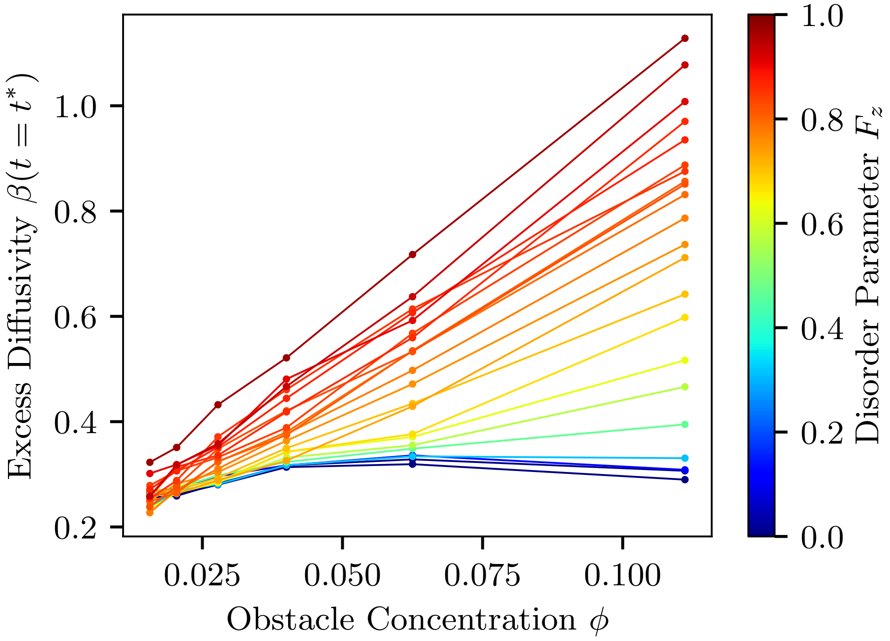


FIG. 5: Excess diffusivity β *vs* obstacle concentration ϕ when using the HPW method with a single obstacle per well. The curves from bottom to top correspond to length scales $\lambda \in \{0.1, 0.2, 0.3, ..., 1.8, 1.9, 2.0, 2.5, 4.0\}$.

$F_z > \frac{2}{3}$, i.e. the excess diffusivity becomes a monotonic function of the obstacle concentration.

We are able to link the crossover length r^* with the excess diffusivity β by first introducing the instantaneous rate of diffusion $\rho = \partial\langle r^2(t)\rangle/\partial t$. The time at which the instantaneous rate of diffusion is equal for both the steady-state and anomalous regimes is given by

$$t_c = \alpha^{\frac{1}{1-\alpha}}\, t^* \,, \tag{4}$$

where $t^* = (D_\alpha/D)^{1/(1-\alpha)}$ is the crossover time. This critical time inherently means that for $t > t_c$ the excess diffusivity β has essentially reached its asymptotic value $\beta(t \to \infty)$. We can estimate $\beta(t = t_c)$ using the approximation

$$\beta^2(t = t_c) \approx 4D_\alpha t_c^\alpha - 4Dt_c \,. \tag{5}$$

Using the value of t_c in eq. 4 we obtain

$$\beta^2(t = t_c) \approx 4Dt^* \left(\alpha^{\frac{1}{1-\alpha}}\right)^\alpha - \alpha^{\frac{1}{1-\alpha}} 4Dt^* \,. \tag{6}$$

We can then use ${r^*}^2 = 4Dt^*$ to further simplify eq. 6

$$\beta^2(t = t_c) \approx (1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} {r^*}^2 \,. \tag{7}$$

We thus predict that β should be directly proportional to r^*. The prefactor is a non-trivial function of the anomalous exponent α:

$$A(\alpha) = (1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} \ . \tag{8}$$

Using $A(\alpha)$ we are able to collapse β, for all values of F_z, ϕ, and method of generating disorder (PUD, HPW, AR2P), onto a single line $\beta = 1.218(6) \times r^* \sqrt{A(\alpha)}$ in Fig. 6.

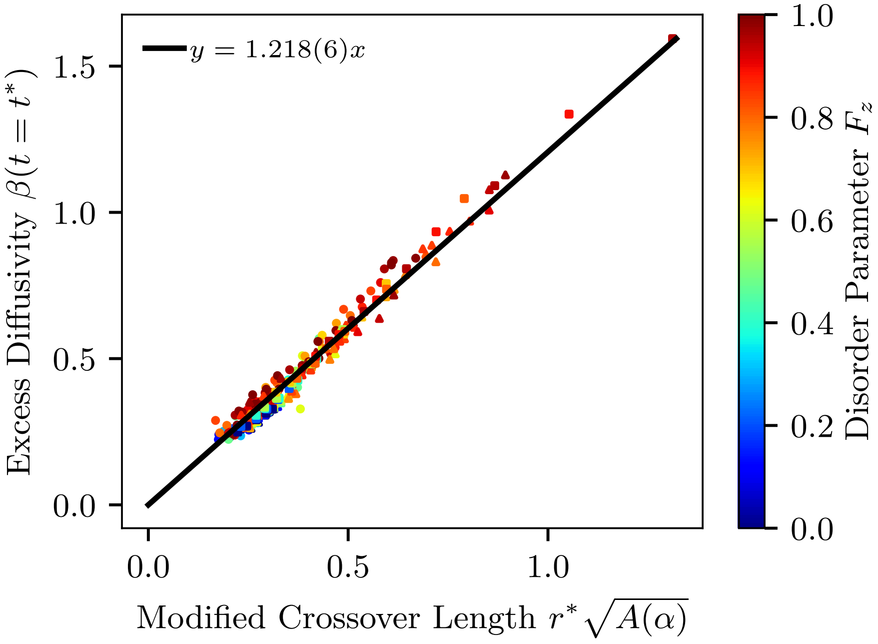


FIG. 6: Excess Diffusivity β *vs* the Modified Crossover Length $r^* \sqrt{A(\alpha)}$. The data for the PUD (●), AR2P (■) and HPW (▲) methods are shown in scatter form together with the combined linear fit.

In summary, we observe qualitative transitions at a critical value of $F_z \approx \frac{2}{3}$ for D, α and r^*, and for β we find two values: $F_z \approx \frac{1}{3}$ and $F_z \approx \frac{2}{3}$. However, the linear relation between β and $r^* \sqrt{A(\alpha)}$ is not affected by the degree of disorder.

B. Transition from Order to Disorder

We now take a single slice of the Figures in Section III A in order to investigate the "order-to-disorder transition" as we vary F_z for an obstacle concentration of $\phi = \frac{1}{9}$. There are two exceptions to this: 1) when studying $\alpha(F_z)$ we use $\phi = \frac{1}{4}$ because the largest differences between random and periodic systems occur at high concentrations, and 2) a result in Fig. 7 (b) led us to explore $\phi = \frac{1}{16}$

for comparison. Each figure contains data corresponding to the different methods used to generate disorder, as well as horizontal lines denoting the values found in the periodic ($F_z = 0$) and random ($F_z = 1$) limits. All Figures in this section, with the exception of Fig. 7, have estimated error-bars calculated through bootstrapping; to increase clarity in the data, we show the median error-bar of each data set on a single point.

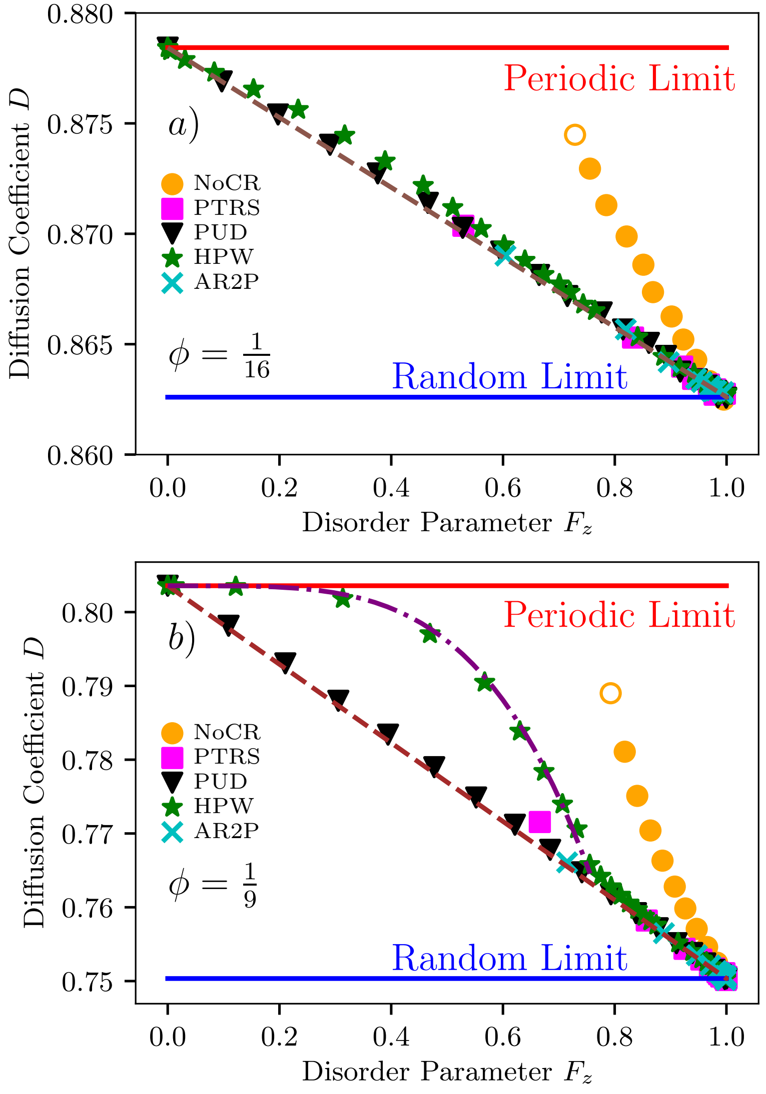


FIG. 7: Steady-state diffusion coefficient D *vs* disorder parameter F_z for obstacle concentrations (a) $\phi = \frac{1}{16}$ and (b) $\phi = \frac{1}{9}$. The dashed lines show the linear interpolations $D(F_z) = (1 - F_z)\,D(0) + F_z\,D(1)$. In (b), the dash-dot line shows the empirical interpolating function $D(F_z) = D(0) - [D(0) - D(\frac{3}{4})] \times [F_z/(\frac{3}{4})]^4$ between $F_z = 0$ and $F_z = \frac{3}{4}$. Without the large bubbling for the HPW data in (a) the empirical interpolating function does not capture the behaviour of the data.

The steady-state diffusion coefficient D: The most remarkable result in Figs. 7 (a) and (b) is the PUD data agreement with a linear interpolation between the Periodic and Random limits independent of obstacle concentration. This is an unexpected result as F_z is not a linear function of the fraction of obstacles that are replaced randomly (data not shown). This demonstrates that F_z is indeed a useful measure of the degree of disorder.

Overall, the data for $D(F_z)$ (and, indeed, for all properties) are sandwiched between the NoCR and PUD data. The fact that the NoCR case is widely different demonstrates that clustering greatly impacts diffusivity. Also, we clearly have a key transitional point at $F_z \approx \frac{3}{4}$, where all of the methods (except NoCR) begin to converge with the linear interpolation.

The HPW data is located between these two limiting cases because clusters are rare when the length scale λ is small. Note that HPW becomes similar to PUD if $\lambda > 1/\sqrt{\phi_P}$, equivalent to $F_z > \frac{3}{4}$ here. The HPW data can be represented by an empirical quartic form in the range $F_z \in [0, \frac{3}{4}]$, as shown; however, this "bubble" is significant only for $\phi_P = \frac{1}{9}$ due to the close proximity between wells; the data is closer to the linear interpolation when $\phi_P \leq \frac{1}{16}$.

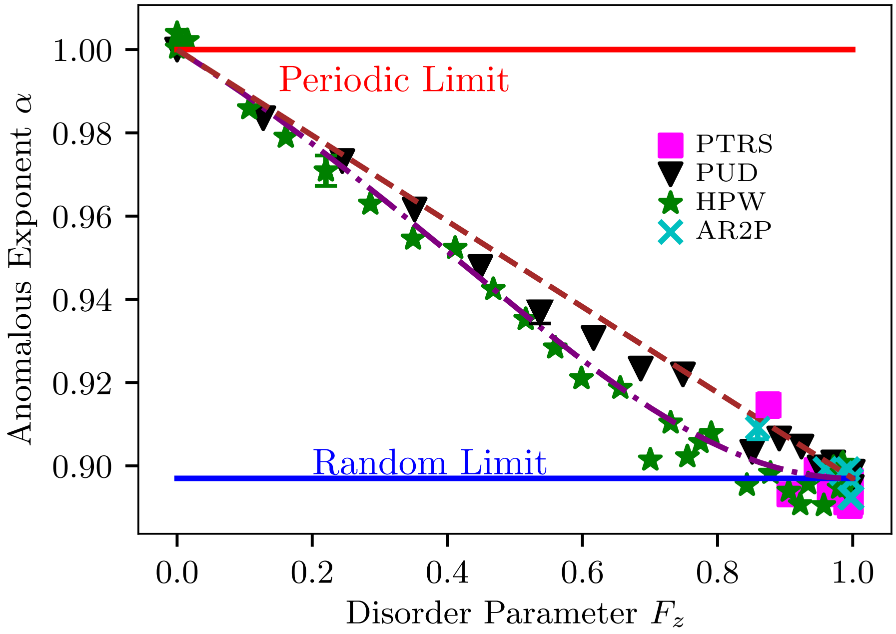


FIG. 8: Anomalous exponent α *vs* disorder parameter F_z for an obstacle concentration $\phi = \frac{1}{4}$. The dashed line shows the linear interpolation $\alpha(F_z) = (1-F_z)\ \alpha(0) + F_z\ \alpha(1)$. The dot-dash line shows the empirical function $\alpha(1) + (\alpha(0) - \alpha(1)) * 2/(1 + (1-F_z)^{-2})$.

The anomalous exponent α: Since α is a weak function of disorder at low concentrations, we use $\phi = \frac{1}{4}$ in this subsection. Unfortunately, this means that we lose access to the NoCR method and impose severe limitations for the PTRS and AR2P methods (also, $\alpha(F_z = 0)$ is ill-defined due to this periodic system being made of corridors of width 1). As for $D(F_z)$, the PUD data in Fig. 8 is very close to the linear interpolation. However the HPW data now remains below the linear interpolation over the whole range of F_z values.

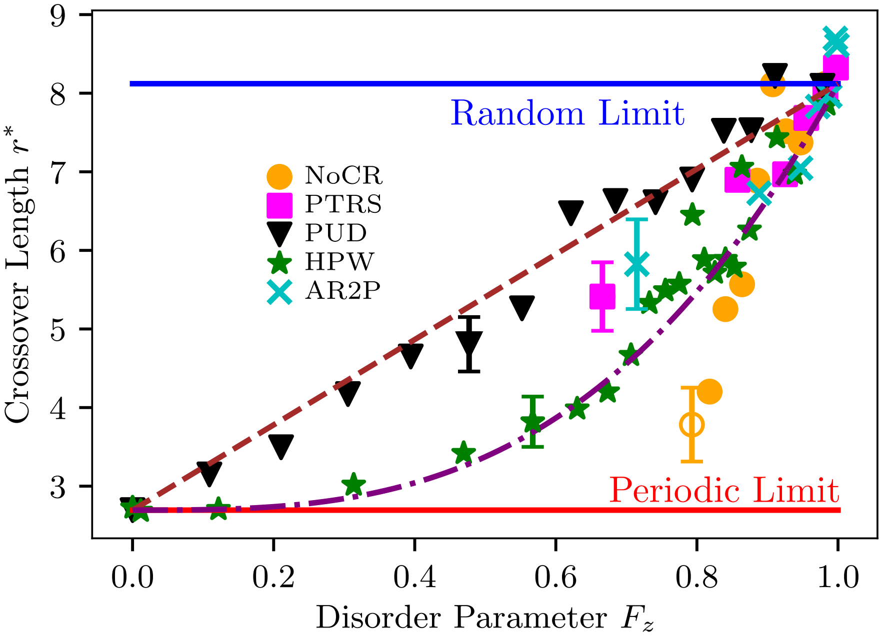


FIG. 9: Crossover length r^* *vs* disorder parameter F_z for an obstacle concentration $\phi = \frac{1}{9}$. The dashed line shows the linear interpolation $r^*(F_z) = (1-F_z)\ r^*(0) + F_z\ r^*(1)$. The dot-dash line shows the empirical interpolating function $r^*(F_z) = r^*(0) + [r^*(1) - r^*(0)] \times F_z^3 \approx r^*(0) \times (1 + [F_z/(\frac{4}{5})]^3)$, where $\frac{4}{5} \approx \left(\frac{r^*(0)}{r^*(1) - r^*(0)}\right)^{\frac{1}{3}}$ is the nearest rational fraction mathematically calculated using the limit values of $r^*(1)$ and $r^*(0)$.

The crossover length r^*: Figure 9 follows suit with Fig. 7 (b) (and (a) however it is less clear) in that the data from all methods are sandwiched between the PUD and NoCR limits. Similar to Fig. 7 (b) we see an apparent departure from the Periodic Limit for the HPW data near $F_z \approx \frac{1}{3}$; an empirical interpolation function (see caption) gives a critical value of disorder $F_z \approx \frac{3}{4}$. Once more, the PUD data remain very close to the linear interpolation.

Excess diffusivity: Figures 9 and 10 are remarkably similar. In other words, the comments about $r^*(F_z)$ above apply to $\beta(F_z)$ as well. This confirms the close relation between excess diffusivity and correlation lengths. In the case of the HPW data, the interpolating function (see caption) gives an apparent critical value of disorder $F_z \approx \frac{2}{3}$ here.

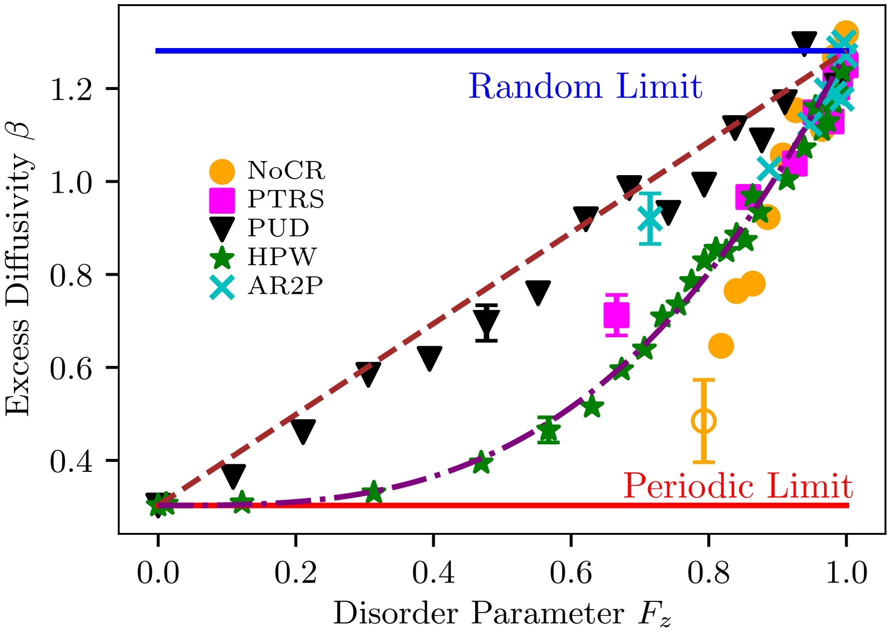


FIG. 10: Excess diffusivity β *vs* disorder parameter F_z for an obstacle concentration $\phi = \frac{1}{9}$. The dashed line shows the linear interpolation $\beta(F_z) = (1-F_z)\ \beta(0) + F_z\ \beta(1)$. The dot-dash line shows the empirical interpolating function $\beta(F_z) = \beta(0) + [\beta(1) - \beta(0)] \times F_z^3 \approx \beta(0) \times (1 + [F_z/(\frac{2}{3})]^3)$, where $\frac{2}{3} \approx \left(\frac{\beta^*(0)}{\beta^*(1)-\beta^*(0)}\right)^{\frac{1}{3}}$ is the nearest rational fraction mathematically calculated using the limit values of $r^*(1)$ and $r^*(0)$.

IV. DISCUSSION AND CONCLUSION

Diffusion in the presence of periodically or randomly distributed obstacles is fundamentally different, as discussed in Section III A. Perhaps the most striking qualitative difference is the concentration dependence of the crossover length $r^*(\phi)$ – see Fig. 4. Our goal was to investigate how a system transitions from quasi-periodic to quasi-random behavior as we increase the amount of disorder in a controlled (tunable) way. We thus designed five different ways to introduce (and control) disorder in a periodic system and proposed a simple measure of disorder, the disorder (or fuzziness) parameter F_z.

The anomalous exponent α is not significantly impacted by disorder at low concentrations – see Fig. 3. It is only when $\phi > \frac{1}{9}$ that we clearly see the impact of "fuzziness" in the change of curvature; the critical disorder parameter here is $F_z^{\alpha} \approx \frac{2}{3}$. We also observe a change in curvature for the diffusion coefficient at $F_z^D \approx \frac{2}{3}$ in Fig. 2. For the excess diffusivity $\beta(\phi)$, we find two transitions; one that is closer to $F_z^{\beta} \approx \frac{1}{3}$ in Fig. 5, as well as a second behavioural change at a critical value more in line with the other parameters $F_z^{\beta} \approx \frac{2}{3}$. The crossover length r^* decreases (increases) in a periodic (random) system when ϕ increases. Remarkably, we observe that $r^*(\phi)$ is nearly constant at a critical "fuzziness" $F_z^{r*} \approx \frac{1}{2}$, (see Fig. 4). In short, the transition point appears to be in the range $\frac{1}{3} \leq F_z \leq \frac{2}{3}$, depending on the physical parameter we examine and the type of disorder we use.

In Section III B, we showed that the "Pick-Up and Drop" (PUD) method yields results that essentially fall on a linear interpolation between the periodic and random results when the concentration is kept constant. This result applies to all of the different parameters studied here (namely, D , β , r^* and α), indicating that our definition of F_z captures the basic elements of the problem in the Fuzzy regime, as F_z is not a linear function of the fraction of randomly replaced obstacles.

At the other end of the spectrum we find the NoCR method (the only method that does not use any underlying periodic motif). Indeed, a recurring theme in Section III B is the different nature of the results obtained with the NoCR approach. When $Q = 0$, NoCR eliminates the presence of aggregates (similar to periodic systems) while keeping some degree of randomness (which decreases as the concentration of obstacles increases); in other words, it could be called "randomly periodic". Clearly, the main effect of NoCR, which is to limit clustering, points to the significant importance of clusters.

As mentioned above, the fact that the NoCR systems are fundamentally different from the other ones points to the key role played by large obstacle aggregates. Future work along the lines of this project should thus include investigating methods for generating Fuzzy systems that could directly tune the size of the aggregates. Of specific interest is the bridge between the HPW data and the NoCR data in Fig. 7. In fact, the NoCR method lends itself to easily making hybrid methods (e.g., combining the HPW with a NoCR constraint to see if data between these two data sets can be obtained in Fig. 7). We are also interested in the applications to drug release, e.g. investigating the release curves as functions of F_z to see if one can change the release pattern by varying F_z while keeping either ϕ or D fixed (work in progress).

REFERENCES

[1]M. J. Saxton, "Anomalous diffusion due to obstacles: a monte carlo study," Biophys. J. **66**, 394–401 (1994).

[2] S. Havlin and D. Ben-Avraham, "Diffusion in disordered media," Adv. Phys. **36**, 695–798 (1987).

[3] S. Havlin and A. Bunde, "Fractals and disordered systems," (1991).

[4] D. Stauffer and A. Aharony, *Introduction to percolation theory* (CRC press, 2018).

[5] M. Palombo, A. Gabrielli, V. Servedio, G. Ruocco, and S. Capuani, "Structural disorder and anomalous diffusion in random packing of spheres," Sci. Rep. **3**, 1–7 (2013).

[6] G. W. Slater and H. L. Guo, Electrophoresis **17**, 977 (1996).

[7] P. Shorten and J. Sneyd, "A mathematical analysis of obstructed diffusion within skeletal muscle," Biophys. J. **96**, 4764–4778 (2009).

[8] P. Donovan, Y. Chehreghanianzabi, M. Rathinam, and S. P. Zustiak, "Homogenization theory for the prediction of obstructed solute diffusivity in macromolecular solutions," PloS one **11**, e0146093 (2016).

[9] H. Lodish and S. L. Zipursky, "Molecular cell biology," Biochem Mol Biol Educ **29**, 126–133 (2001).

[10] E. A. Codling, M. J. Plank, and S. Benhamou, "Random walk models in biology," J R Soc Interface **5**, 813–834 (2008).

[11] R. Phillips, J. Kondev, J. Theriot, H. G. Garcia, and N. Orme, *Physical biology of the cell* (Garland Science, 2012).

[12] A. J. Ellery, R. E. Baker, S. W. McCue, and M. J. Simpson, "Modeling transport through an environment crowded by a mixture of obstacles of different shapes and sizes," (2016).

[13] A. Wedemeier, H. Merlitz, C.-X. Wu, and J. Langowski, J. Chem. Phys. **131**, 064905 (2009).

[14] E. Vilaseca, A. Isvoran, S. Madurga, I. Pastor, J. L. Garcés, and F. Mas, PCCP **13**, 7396 (2011).

[15] N. Ilow and G. W. Slater, "Estimating the steady state diffusion coefficient using data from the transient anomalous regime," (2021), arXiv:2108.04944 [physics.bio-ph].

[16] J.-F. Mercier, G. W. Slater, and H. L. Guo, "Numerically exact diffusion coefficients for lattice systems with periodic boundary conditions. i. theory," J. Chem. Phys. **110**, 6050–6056 (1999).

[17] G. W. Slater and H. L. Guo, "An exactly solvable ogston model of gel electrophoresis: I. the role of the symmetry and randomness of the gel structure," Electrophoresis **17**, 977–988 (1996).

CHAPTER 4

ESTIMATING THE STEADY STATE DIFFUSION COEFFICIENT USING NUMERICAL DATA FROM THE TRANSIENT ANOMALOUS REGIME

Estimating the Steady State Diffusion Coefficient Using Numerical Data from the Transient Anomalous Regime

Nicholas Ilow and Gary W. Slater*
University of Ottawa, Department of Physics, Ottawa, Ontario, K1N 6N5, Canada

(Dated: January 10, 2022)

When particles/molecules diffuse in systems that contain obstacles, the steady-state regime (during which the mean-square displacement scales linearly with time, $\langle r^2 \rangle \sim t$) is preceded by a transient regime. It is common to characterize this transient regime using the concept of anomalous (sub)diffusion with the scaling law $\langle r^2 \rangle \sim t^{\alpha}$, where the corresponding exponent $\alpha < 1$. We propose a new method to estimate the critical time t^* that marks the transition between these two regimes. The method uses short-time data from the transient regime to estimate t^*, which can then be used to estimate the steady-state diffusion coefficient D. In other words, we propose a procedure that makes it possible to estimate the steady state diffusion coefficient without reaching the steady-state. We test the procedure with various two-dimensional lattice systems.

I. INTRODUCTION

Diffusion in the presence of obstacles, often called obstructed diffusion, is present in a wide range of physical, chemical and biological systems [1–5]. Generally speaking, we can expect two different time regimes. At short times, the diffusing particles explore the spatial constraints and start colliding with the obstacles [6, 7] (this is sometimes preceded by a free diffusion regime if the obstacles are far from one another [8, 9]). The steady-state, which is reached at long times, is characterized by the fact that the mean squared displacement (MSD) of a tracer particle is linear in time, with some excess contribution (to be denoted β^2) which is due to the fact that diffusion is faster at early times [8] (the collisions slow down the diffusion process):

$$\left\langle r^2(t) \right\rangle = \beta^2 + 4Dt \ , \tag{1}$$

where D is the diffusion coefficient. Therefore, in order to access the diffusion coefficient, the experimental or simulation time must be much larger than the crossover time t^* between the transient and steady-state regimes[10].

Equation 1 is not valid during the transient regime. Instead, the MSD is often fitted using the concept that it will follow a power law with an "anomalous" exponent α [11, 12]

$$\left\langle r^2(t) \right\rangle = 4D_\alpha t^\alpha \ , \tag{2}$$

where in our case $\alpha < 1$. Consequently, this regime is also called the anomalous diffusion regime. In practical cases, defining the time range over which eq. 2 might be valid is often arbitrary [13–17]. In Sections III A and III B, we propose methods to determine the center t_I and the width Σ_t of the anomalous regime, respectively.

Using eqs. 1 (with $\beta = 0$) and 2, we can estimate the crossover time t^*. The time at which both equations predict the same displacement is simply

$$t^* = \left(D_\alpha / D\right)^{1/(1-\alpha)} , \tag{3}$$

while the corresponding crossover distance r^* is given by

$$r^{*2} = 4Dt^* = 4D_\alpha t^{*\alpha} = \left(D_\alpha / D^\alpha\right)^{1/(1-\alpha)} . \tag{4}$$

Note that obtaining t^* and r^* requires knowledge of the steady-state diffusion coefficient D.

The main result of this paper can be summarized as follows: the width Σ_t of the anomalous regime, as we define it in Section III B, is an excellent approximation for the crossover time t^*, even though it does not require any knowledge of the diffusion coefficient D. This has an important and direct application since we can rearrange eq. 3 as

$$D = D_\alpha {t^*}^{\alpha-1} \ . \tag{5}$$

This equation implies that if we could find a proxy for t^* using data from the anomalous regime, we would be able to estimate the steady state diffusion coefficient D without reaching the steady state. As we shall demonstrate, this is precisely what our definition of Σ_t allows us to do.

Experimentally diffusion problems can be explored using a range of techniques including Fluorescence Recovery After Photobleaching [18, 19], and Single Particle Tracking [20, 21]. However, most methods do not easily allow experimentalists to probe both regimes [22, 23] because of the wide range of time scales involved. Our paper thus offers a potential approach to estimating the steady-state diffusion coefficient D using short-time data only.

In this manuscript we introduce (Sections III A–III B) and test (III C) our new concepts using a simple model of obstructed diffusion with random obstacle configurations on a two dimensional square lattice (Section II) since this allows us to obtain high-precision numerical data for both the anomalous and steady-state regimes.

II. METHODOLOGY

We consider random walks on a square lattice with a mesh size a. The standard Lattice Monte Carlo (LMC) algorithm [24, 25] includes random jumps of length a along one of the four Cartesian directions, each with a probability $p = \frac{1}{4}$. The second moment of the distribution of displacements after N

---

* gary.slater@uottawa.ca

steps is then given by $\left\langle r^2\right\rangle_G = Na^2$; if the duration of these MC steps is τ, this can be rewritten as $\left\langle r^2\right\rangle_G = 4Dt$, with $D = a^2/4\tau$ the diffusion coefficient and $t = N\tau$ the time. Since the second moment is identical to the one predicted by the solution of the diffusion equation (which is a Gaussian distribution, hence the subscript G), LMC models are often used to simulate diffusion problems. However the 2D LMC algorithm actually gives the following fourth moment

$$\left\langle r^4(t)\right\rangle = 32D^2t^2 + (4-32p)Dta^2 \ , \tag{6}$$

while the solution of the diffusion equation gives $\left\langle r^4(t)\right\rangle_G = 32D^2t^2$. The last term in eq. 6 is negligible at long times, but it does affect the mean time between collisions with obstacles and hence impacts the transient regime including the value of the anomalous exponent α (data not shown). To eliminate this correction term and improve the algorithm, we use $p = \frac{1}{8}$ and a probability $p_o = 1-4p = \frac{1}{2}$ of staying put during a time step.

Typically, obtaining the steady-state diffusion constant D would require long LMC simulations. However calculating D for a random-walker in a particular system of randomly distributed obstacles (which act as reflecting boundaries and occupy a fraction ϕ of the surface area) with periodic boundary conditions is made possible using the numerical methods outlined in [26]. This approach first involves computing the steady state concentration profile under a weak bias ε. This is done by solving an $N\times N$ matrix, where N is the number of lattice sites accessible to the particle, and each row is determined by a rate equation with biased jumping probabilities. One can then calculate the mean velocity using the computed concentration profile and the local velocities resulting from the bias and the presence of reflecting obstacles. Finally, the mean velocity $v(\varepsilon)$ is used to compute the diffusion constant D via the Nernst-Einstein relation:

$$\frac{D(\phi)}{D(0)} = \lim_{\varepsilon\to 0} \frac{v(\varepsilon,\phi)}{v(\varepsilon,0)} \ . \tag{7}$$

We use this approach to obtain very high precision estimates of D below. For these calculations, we use a lattice of size $L = 128$ and an ensemble size of 2000. A lattice spacing of $a = 1$ is chosen, with $\tau = \frac{1}{8}$ corresponding to the required timestep for $p = \frac{1}{8}$, such that $D(0) = 1$.

We use a Markov Chain Monte Carlo (MCMC) method to obtain high-precision short-time data. Since the obstacles are reflecting boundary conditions, the concentration evolves (starting with a unit concentration on a single site at the center of the lattice) as follows:

$$\begin{aligned} C_{x,y}(t+1) = {} & [C_{x-1,y}(t) + C_{x,y}(t)O_{x+1,y}]p_{+x} \\ & + [C_{x+1,y}(t) + C_{x,y}(t)O_{x-1,y}]p_{-x} \\ & + [C_{x,y-1}(t) + C_{x,y}(t)O_{x,y+1}]p_{+y} \\ & + [C_{x,y+1}(t) + C_{x,y}(t)O_{x,y-1}]p_{-y} \\ & + [C_{x,y}(t)]p_o, \end{aligned} \tag{8}$$

where $C_{x,y}(t)$ is the concentration at lattice site (x,y) at time (or iteration) t, and $O_{x,y}$ is a binary value (0 or 1) describing the presence of an obstacle at (x,y). Note that $p_{\pm x} = p_{\pm y} = p = \frac{1}{8}$. The MCMC calculations are completed on a lattice of size 512×512, and we handle disorder by averaging over an ensemble of 2000 different obstacle configurations.

A feature of random systems is the possible presence of closed areas, which we call "lakes". A lake is an area in which unoccupied sites exist, but these sites are inaccessible to a tracer particle initially located outside. We must make sure that a Markov chain calculation does not start in a lake since $\left\langle r^2(t)\right\rangle$ would then quickly plateau. When we calculate D using exact matrix calculations and eq. 7, lakes lead to block-diagonal matrices, and the only non-zero value of D is found for the block that corresponds to the connected pathway through the network. To avoid such issues, we simply fill all unoccupied sites within lakes with phantom obstacles.

III. RESULTS

A. Data analysis

Central to our methodology is plotting $\log\langle r(t)^2\rangle/4t$ *vs* $\log(t)$. As Fig. 1 shows, the transient and steady-state regimes are clearly visible when the data are plotted in this way. Since the location and boundaries of the transient regime are often defined arbitrarily [13–17], the value of the anomalous exponent α is generally method-dependent. However, an inflection point is present for all cases in this type of log-log plot. We thus propose that the inflection point provides a robust way of defining the center of the anomalous regime. The anomalous exponent can then be extracted from the slope of the tangent at the inflection point (which is $\alpha-1$ given eq. 2) while D_α is simply given by the ratio $r_I^2/4t_I^\alpha$ (the subscript I refers to evaluation at the inflection point). We investigate the width of the anomalous regime in Sec. III B. Figure 1 further defines the crossover time t^* as the intersection of the tangent and the horizontal line marking the steady state diffusion coefficient D. The inset details the numerically calculated inflection saddle points, thus determining the fitting region for estimating the interpolated position of the inflection point and both α and D_α.

We present the results of t^*, $r^* = \sqrt{4Dt^*}$ and α, for a 1×1 walker in random systems with 1×1 obstacles for varying obstacle concentrations ϕ in Fig. 2. At larger obstacle concentrations the transition length scale r^* grows to values such that our system size ($L = 512$) is no longer large enough to avoid finite-size effects (r^* is related to the crossover length or cluster size in percolation theory [27]). The transition time t^* behaves similarly. The exponent α, on the other hand, decreases continuously with ϕ to reach a value of about $\alpha = 0.733$ at $\phi = \phi^*$. Finally, the blue data points in Fig. 3 show how the diffusion coefficient D decays with ϕ; in particular, we see that $D \to 0$ at the percolation threshold $\phi = \phi^* \approx 0.40725$ (this actually defines ϕ^*).

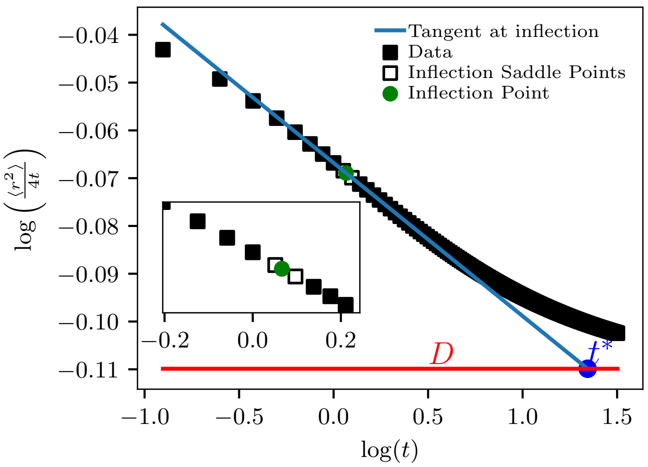


FIG. 1: Log-log plot of the instantaneous diffusivity ratio $\langle r^2 \rangle/4t$ *vs.* time t for a 1×1 random-walker, with 1×1 obstacles at an obstacle concentration $\phi = 10\%$. The steady-state diffusion coefficient D (horizontal red line) was calculated using the matrix technique described in Section II, and the slope of the line tangent to the inflection point is $\alpha - 1$. Where these two lines intersect defines the crossover time t^*.

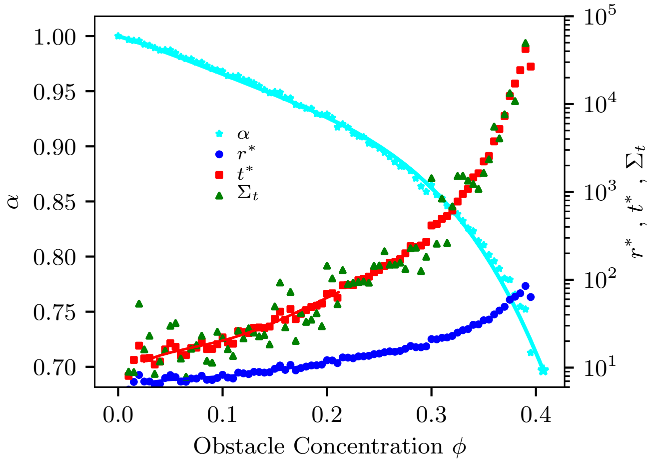


FIG. 2: Key system parameters *vs.* obstacle concentration ϕ; both the obstacles and the particles are of size 1×1. The crossover length r^* and crossover time t^* diverge as ϕ approaches the percolation threshold[29] $\phi^* = 0.4072539492079$. The anomalous exponent α decreases (from unity) towards the percolation threshold value[10] $\alpha(\phi \to \phi^*) = \frac{2}{2.73}$. Our estimate of the width of the anomalous regime, Σ_t, provides an excellent approximation for t^*.

B. The width of the anomalous regime

Obviously, the transient regime can only be said to satisfy eq. 2 over a small region centered around the inflection point in Fig. 1. Since the slope at this point is $\alpha - 1$ while the second derivative is zero, we can use the third derivative to measure the width of the region where the second derivative remains negligible (i.e., the width of the region where a straight line fit might be valid). We thus propose to define the width Σ_α of the transient/anomalous regime (in this specific log-log space) using the expression

$$\Sigma_\alpha^3 = \left.\frac{f(x)}{\partial^3 f(x)/\partial x^3}\right|_{x=log(t_I)} , \qquad (9)$$

where $f = \log(\langle r(t)^2\rangle/4t)$, $x = \log(t)$, and t_I corresponds to the inflection point. The third derivative here represents the rate of change of the second derivative, in this case measuring how quickly it becomes non-zero. The resulting Cartesian temporal width, Σ_t, is then

$$\Sigma_t = 10^{[log(t_I)+\Sigma_\alpha]} - 10^{[log(t_I)-\Sigma_\alpha]}. \qquad (10)$$

In principle, we should expect the transient (or anomalous) regime to transition into the steady-state regime when the time t exceeds $\approx \Sigma_t$. Furthermore, we expect this width to diverge at the percolation threshold ϕ^* because there is no steady-state at that critical point [28].

Figure 2 shows that the width Σ_t defined in eq. 10 does indeed behave as expected: it increases with ϕ and diverges at percolation. Surprisingly, Σ_t is found to be essentially equal to t^* over the whole range of concentrations. The fact that it relies on a third derivative of discrete data explains the presence of noise. Unlike t^*, the width Σ_t does not require any knowledge of the steady-state diffusion coefficient D; in other words, it can be obtained using short-time data only.

Since $t^* \approx \Sigma_t$, we can rewrite eq. 5 as

$$D \approx D_\alpha \Sigma_t^{\alpha-1} . \qquad (11)$$

We stress again the fact that the three parameters on the *rhs* of this expression can be obtained using transient data only. Therefore, eq. 11 implies that it is possible to obtain an estimate of the steady-state diffusion coefficient D that uses only parameters extracted from the transient regime (there is no need to reach the steady-state). We test this approach to estimating D in the next two sections for several sizes of random-walkers and obstacles.

C. Estimating the diffusion coefficient D using Σ_t

In order to test the accuracy of eq. 11, we compare its prediction to exact values obtained using the method described in Section II, and we do this for six different systems of random walkers and randomly distributed obstacles in Figs. 3 (for obstacles of different sizes) and 4 (for random walkers of different sizes).

As Fig. 3 shows, the approximation holds very well for 1×1 particles and different obstacle sizes across all obstacle concentrations. Figure 4 shows equally good results when larger particles move between small obstacles. In both cases, the results become more noisy and less reliable near the relevant percolation threshold ϕ^* because of finite size effects. Indeed, the last few points of the exact diffusion constant data set in Figs. 3 and 4 show a change in curvature; this is the result of the crossover length r^* increasing quickly near perco-

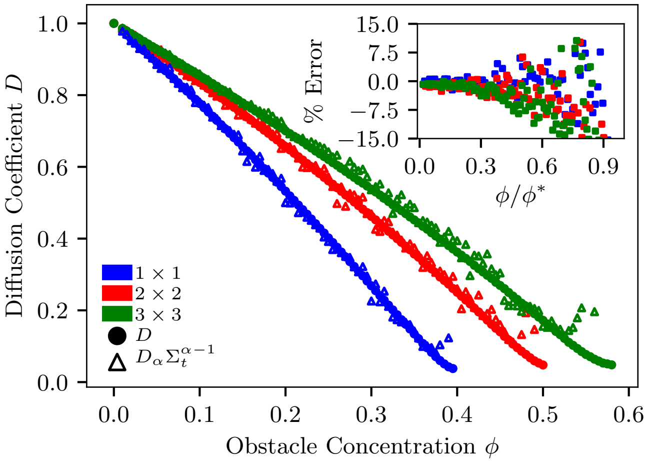


FIG. 3: Diffusion coefficient of a 1×1 particle *vs* obstacle concentration ϕ (fraction of the surface area covered by obstacles). The solid circles correspond to the high-precision values of D obtained using the matrix method; the triangles give the approximation $D_\alpha \Sigma_t^{\alpha-1}$ as defined by eq. 11. Results are shown for three different obstacle sizes. Inset: relative error $\left[\frac{D-D_\alpha\Sigma_t^{\alpha-1}}{D}\right]$ (in %) *vs* obstacle concentration divided by the percolation threshold ϕ^*.

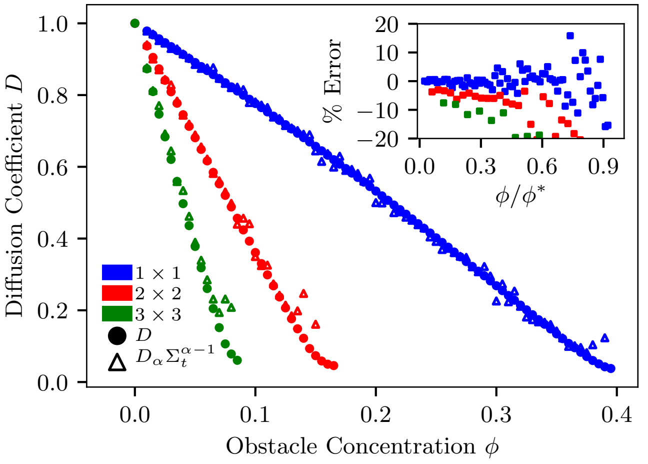


FIG. 4: Diffusion coefficient *vs* concentration of 1×1 obstacles. The solid circles correspond to the high-precision values of D obtained using the matrix method; the triangles give the approximation $D_\alpha \Sigma_t^{\alpha-1}$ as defined by eq. 11. Results are shown for three different random walker sizes. Inset: relative error $\left[\frac{D-D_\alpha\Sigma_t^{\alpha-1}}{D}\right]$ (in %) *vs* obstacle concentration divided by the percolation threshold ϕ^*.

lation (see Fig. 2), and finite size effects becoming prominent (as the condition $L \gg r^*$ is no longer valid).

To estimate the percolation thresholds for our systems, we performed a linear fit on the final 10 data points prior to the change of curvature observed due to finite size effects. This linear fit is extrapolated to $D \to 0$ to estimate ϕ^*. The results are given in Table I. The insets in Figs. 3 and 4 show the relative error in our estimates of D as a function of the scaled concentration ϕ/ϕ^*. These errors increase with the concentration and become of the order 15% as we approach the percolation thresholds here; as usual, these errors decrease with the ensemble size (data not shown). Note that the three curves in the main parts of Figs. 3 and 4 nearly collapse on a universal curve if D is plotted as a function of ϕ/ϕ^*, as one would expect (shown in Appendix A).

IV. CONCLUSION

In this paper, we first proposed a non-arbitrary way to characterize the transient/anomalous regime in the case of obstructed diffusion problems. In particular, we introduced a method to locate the center of this regime and estimate the value of the related anomalous exponent α. Furthermore, we suggested a way to measure the width Σ_t of the regime during which the MSD could potentially be fit with $\langle r^2(t) \rangle \sim t^\alpha$.

Using simple two-dimensional lattice models of obstructed diffusion, we found that the width Σ_t of the anomalous regime can act remarkably well as a proxy for the crossover time t^* marking the transition between the anomalous and steady-state regimes. For instance, both diverge identically near the percolation threshold. This allows us to write eq. 11 which yields an estimate of the steady state diffusion coefficient D using only short time data. Our simulations have shown that this method for estimating the steady state diffusion coefficient is robust for a variety of random obstacle variants (i.e., larger obstacles, and larger walkers).

One drawback to this approach is the need to evaluate a third-derivative to compute Σ_t and hence D. This may limit the usefulness of our findings when the data are noisy (we saw examples of this when close to percolation thresholds). Clearly, this needs to be explored further, e.g. using data coming from Molecular Dynamics simulations or experimental data.

Experimentally (and with computer simulations), our novel data analysis method can allow one to access the steady-state diffusion coefficient where otherwise it would be inaccessible due to the inability to reach late time data. Interestingly, eq. 4 can also be rewritten as ${r^*}^2 = 4D_\alpha\Sigma_t^\alpha$; this means that the correlation length ${r^*}^2$ can also be estimated using short-time data.

ACKNOWLEDGMENTS

NI thanks the University of Ottawa for an admission scholarship. GWS acknowledges the support of both the University of Ottawa and the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN/046434-2013.

	1×1 (see ref. [29])	2×2	3×3
Obstacles	0.40725394920790(2)	0.51(1)	0.58(1)
Walkers	0.40725394920790(2)	0.161(4)	0.083(1)

TABLE I: Percolation threshold ϕ^* for walkers and obstacles of different sizes, where the counterpart remains at a size of 1×1.

Appendix A: Collapsed diffusion coefficients using percolation thresholds

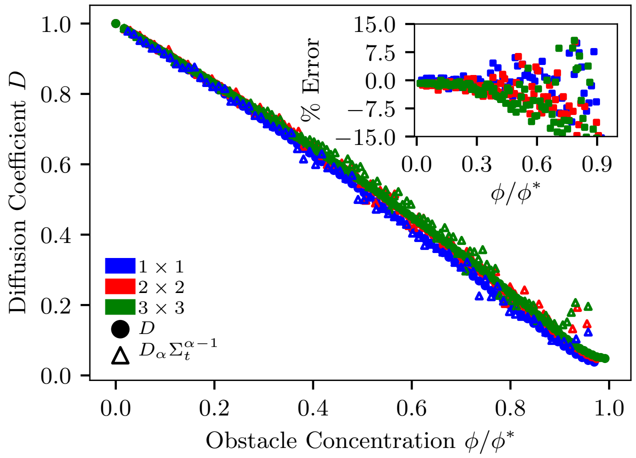


FIG. 5: The steady-state diffusion coefficient D and its estimate $D_\alpha \Sigma_t^{\alpha-1}$ as a function of the scaled obstacle concentration ϕ/ϕ^* for 1×1 walkers and various obstacle sizes.

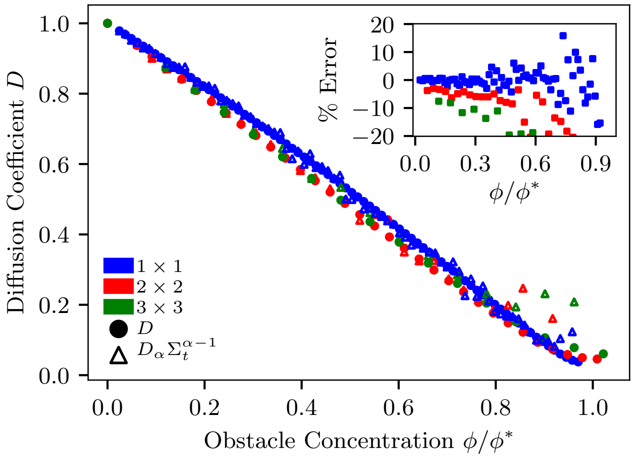


FIG. 6: The steady-state diffusion coefficient D and its estimate $D_\alpha \Sigma_t^{\alpha-1}$ as a function of the scaled obstacle concentration ϕ/ϕ^* for 1×1 obstacles and various walker sizes.

---

[1] T. V. Ratto and M. L. Longo, Obstructed diffusion in phase-separated supported lipid bilayers: a combined atomic force microscopy and fluorescence recovery after photobleaching approach, Biophys. J. **83**, 3380 (2002).

[2] B. J. Rogers and M. J. Wirth, Obstructed diffusion in silica colloidal crystals, J. Phys. Chem. **117**, 6244 (2013).

[3] A. V. Weigel, S. Ragi, M. L. Reid, E. K. Chong, M. M. Tamkun, and D. Krapf, Obstructed diffusion propagator analysis for single-particle tracking, Phys. Rev. E **85**, 041924 (2012).

[4] P. Shorten and J. Sneyd, A mathematical analysis of obstructed diffusion within skeletal muscle, Biophys. J. **96**, 4764 (2009).

[5] S. J. Anderson, J. Garamella, S. Adalbert, R. J. McGorty, and R. M. Robertson-Anderson, Subtle changes in crosslinking drive diverse anomalous transport characteristics in actin–microtubule networks, Soft Matter **17**, 4375 (2021).

[6] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai, Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking, PCCP **16**, 24128 (2014).

[7] D. Krapf, Mechanisms underlying anomalous diffusion in the plasma membrane, Curr Top Membr **75**, 167 (2015).

[8] F. Höfling and T. Franosch, Anomalous transport in the crowded world of biological cells, Rep. Prog. Phys. **76**, 046602 (2013).

[9] M. D. Stolle and C. Fradin, Anomalous diffusion in inverted variable-lengthscale fluorescence correlation spectroscopy, Biophys. J. **116**, 791 (2019).

[10] M. J. Saxton, Anomalous diffusion due to obstacles: a monte carlo study, Biophys. J. **66**, 394 (1994).

[11] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach, Phys. Rep. **339**, 1 (2000).

[12] I. M. Sokolov, Models of anomalous diffusion in crowded environments, Soft Matter **8**, 9043 (2012).

[13] J. E. Goose and M. S. Sansom, Reduced lateral mobility of lipids and proteins in crowded membranes, PLoS Comput. Biol. **9**, e1003033 (2013).

[14] N. Alcázar-Cano and R. Delgado-Buscalioni, A general phenomenological relation for the subdiffusive exponent of anomalous diffusion in disordered media, Soft Matter **14**, 9937 (2018).

[15] A. J. Ellery, R. E. Baker, and M. J. Simpson, An analytical method for disentangling the roles of adhesion and crowding for random walk models on a crowded lattice, Phys. Biol. **13**, 05LT02 (2016).

[16] A. Wedemeier, H. Merlitz, and J. Langowski, Anomalous diffusion in the presence of mobile obstacles, EPL **88**, 38004 (2009).

[17] E. Vilaseca, A. Isvoran, S. Madurga, I. Pastor, J. L. Garcés, and F. Mas, New insights into diffusion in 3d crowded media by monte carlo simulations: effect of size, mobility and spatial distribution of obstacles, PCCP **13**, 7396 (2011).

[18] M. Arrio-Dupont, G. Foucault, M. Vacher, P. F. Devaux, and S. Cribier, Translational diffusion of globular proteins in the cytoplasm of cultured muscle cells, Biophys. J. **78**, 901 (2000).

[19] O. Seksek, J. Biwersi, and A. Verkman, Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus, J. Cell Biol. **138**, 131 (1997).

[20] K. Murase, T. Fujiwara, Y. Umemura, K. Suzuki, R. Iino, H. Yamashita, M. Saito, H. Murakoshi, K. Ritchie, and A. Kusumi, Ultrafine membrane compartments for molecular diffusion as revealed by single molecule techniques, Biophys. J. **86**, 4075 (2004).

[21] A. Kusumi, H. Ike, C. Nakada, K. Murase, and T. Fujiwara, Single-molecule tracking of membrane molecules: plasma membrane compartmentalization and dynamic assembly of raft-philic signaling molecules, Semin Immunol. **17**, 3 (2005).

[22] N. Monnier, S.-M. Guo, M. Mori, J. He, P. Lénárt, and M. Bathe, Bayesian approach to msd-based analysis of particle motion in live cells, Biophys. J. **103**, 616 (2012).

[23] M. Woringer, I. Izeddin, C. Favard, and H. Berry, Anomalous subdiffusion in living cells: Bridging the gap between experiments and realistic models through collaborative challenges, Front. Phys. **8**, 134 (2020).

[24] G. H. Weiss, *Aspects and applications of the random walk* (Elsevier Science & Technology, 1994).

[25] B. D. Hughes, Random walks and random environments, Bull. Am. Math. Soc. **35**, 347 (1998).

[26] J.-F. Mercier, G. W. Slater, and H. L. Guo, Numerically exact diffusion coefficients for lattice systems with periodic boundary conditions. i. theory, J. Chem. Phys **110**, 6050 (1999).

[27] D. Stauffer and A. Aharony, *Introduction to percolation theory* (CRC press, 2018).

[28] P. King and M. Masihi, Percolation in porous media, in *Encyclopedia of Complexity and Systems Science*, edited by R. A. Meyers (Springer New York, New York, NY, 2009) pp. 6565–6579.

[29] J. L. Jacobsen, Critical points of potts and o (n) models from eigenvalue identities in periodic temperley–lieb algebras, J. Phys. A Math. Theor. **48**, 454003 (2015).

CHAPTER 5

AN EMPIRICAL METHOD TO CHARACTERIZE DISPLACEMENT DISTRIBUTION FUNCTIONS FOR ANOMALOUS AND TRANSIENT DIFFUSION

An empirical method to characterize displacement distribution functions for anomalous and transient diffusion

Le Qiao, Nicholas Ilow, Maxime Ignacio, Gary W. Slater
Department of Physics, University of Ottawa,
Ottawa, Ontario K1N 6N5, Canada.

(Dated: January 4, 2022)

Abstract

We propose a practical empirical fitting function to characterize non-Gaussian displacement distribution functions (DispD) often observed for heterogeneous diffusion problems. We first test this fitting function with the problem of a colloidal particle diffusing between two walls using Langevin Dynamics (LD) simulations of a raspberry particle coupled to a lattice Boltzmann (LB) fluid. We also test the function with a simple model of anomalous diffusion on a square lattice with obstacles. In both cases, the fitting parameters provide more information than just the Kurtosis (which is often the method used in such cases), including a length scale characterizing the tails of the DispD begin. In all cases, the fitting parameters smoothly converge towards the Gaussian values when the systems become less anomalous.

I. INTRODUCTION

Diffusion in inhomogeneous media is ubiquitous in Nature and can be observed in a wide range of systems including surface diffusion of atoms in an elastic field[1–3]; diffusion of colloidal particles in a confined fluid[4–8] or polymer network[9,10]; polymer translocation through a nanopore membrane[11–14]; protein diffusion in crowded cell environments[11,15–17]. Experimental observations have shown that some of these systems can lead to non-trivial dynamical properties which require special attention.

For instance, some systems give rise to anomalous diffusion where the mean-square displacement (MSD) increases like t^γ, with $\gamma < 1$. This is often related to diffusion in disordered systems, in which case there is a crossover distance beyond which diffusion becomes Brownian ($\gamma = 1$). Another particularly interesting case is the possible existence of *anomalous yet Brownian* diffusion which is characterized by a linear MSD ($\gamma = 1$) coexisting with a non-Gaussian Displacement Distribution (DispD)[18–23]. The physical origin of anomalous yet Brownian diffusion remains a very active field of research, in particular because it may differ between systems[24]; a common feature though is the fact that the diffusion coefficient varies during the process (e.g., in space).

In both classes of problems, the nature of the DispD is central to our understanding of the physics. The Kurtosis $K(t) = \mu_4(t)/\mu_2(t)^2$ of the DispD is often used to characterize the deviations from Gaussian (normal diffusion) dynamics[4,10,20,25] (μ_i is the i^{th} central moment of the distribution). However, the DispD obviously contains more information than what the Kurtosis provides, including the shape of the tails and potentially the length scale(s) that separate the various dynamical regimes.

In this short article, we propose a new empirical function that can be used to fit a DispD that has different regimes for short and large distances; in particular, it can capture both the Gaussian and non-Gaussian components of a DispD thus allowing us to locate the transition between the two as the external control parameters are changed. To test this flexible interpolating function, we simulate two different systems corresponding to the two classes of problems mentioned above. First, we use a coupled Langevin Dynamics–Lattice Boltzmann (LD-LB) method to simulate the diffusion of a particle in a liquid between two flat walls, an example of anomalous yet Brownian motion[21,26]. We then study diffusion on a lattice with obstructed sites, a case where short-time diffusion is known to be anomalous[27,28].

II. A PRACTICAL FITTING FUNCTION

To fit both the central and tail parts of "anomalous" DispD distributions at time t, we propose to use the 3-parameter functional form

$$P(r,t) \sim \exp\left[1-\left[1+\left(\tfrac{r}{r_o}\right)^{2-\alpha}\right]^{1-\beta}\right], \tag{1}$$

where r is the displacement and t is the time. The time-dependent fitting parameters are the length scale $r_o(t)$ and the two exponents, which we expect to be in the ranges $0 \leq \alpha(t) < 2$ and $\beta(t) < 1$. A one-regime DispD corresponds to $\beta = 0$, including the Gaussian distribution for $\alpha = 0$, the exponential function for $\alpha = 1$, and the stretched exponential distribution for $1 < \alpha < 2$.

We have two regimes (which are equivalent when $\beta = 0$), namely

$$P(r,t) \sim \begin{cases} \exp\left[-(1-\beta)\,(r/r_o)^{2-\alpha}\right] & \text{for } r \ll r_o \\ \exp\left[-\,(r/r_o)^{(2-\alpha)(1-\beta)}\right] & \text{for } r \gg r_o. \end{cases} \tag{2}$$

The combination $[\beta = \frac{1}{2}; \alpha = 0]$ is a Gaussian with exponential tails, while $1 - \beta = \frac{2}{2-\alpha}$ is a distribution with Gaussian tails (a special case being $[\beta = -1; \alpha = 1]$, an exponential distribution with Gaussian tails). Note that r_o marks not only the transition between these two limits, but also the length scale characterizing the decay of the tails of the DispD.

III. EXAMPLE I: WALL-HINDERED DIFFUSION

Our first example is a simulation of the wall-hindered diffusion of a spherical particle of radius R between two walls separated by a distance h. In bulk solution, the particle's diffusion coefficient is given by Stokes' law $D_0 = k_B T/6\pi\eta R$, with η the viscosity of the fluid. Hydrodynamic interactions (HI) make the diffusion coefficient space-dependent and anisotropic near surfaces. For a particle at a distance z (see Fig. 1a) from a single flat wall[29], the diffusivities parallel and perpendicular to the wall are (with $\Gamma = \frac{R}{R+z}$)

$$D_{\parallel}(z)/D_0 \approx 1 - \tfrac{9}{16}\Gamma + \tfrac{1}{8}\Gamma^3 - \tfrac{45}{256}\Gamma^4 - \tfrac{1}{16}\Gamma^5 + ... \tag{3}$$

$$D_{\perp}(z)/D_0 \approx (6 - 10\Gamma + 4\Gamma^2)/(6 - 3\Gamma - \Gamma^2), \tag{4}$$

As recently shown by Matse *et al.*[30], the z-dependence of $D_\perp$ leads to anomalous yet Brownian motion (linear time-dependence of the MSD but non-Gaussian DispD).

Since this problem is one-dimensional (along z) while $\alpha = 0$ (the central part of the distribution is Gaussian[30]), the fitting function will read

$$P(z,t) = \frac{1-A}{\sqrt{2\pi z_o^2}} \, \exp\left[1 - \left[1 + \frac{z^2}{2z_o^2}\right]^{1-\beta}\right] , \tag{5}$$

where A is a normalization factor. Note that for a Gaussian distribution we simply have $\beta = A = 0$, which leads to $\langle z^2(t)\rangle = \int_{-\infty}^{+\infty} z^2 P(z,t)\, dz = z_o(t)^2 = 2Dt$. The even moments of the distribution are given by

$$\langle z^{2i}\rangle = z_o^{2i} \times \frac{Q(2i,\beta)}{Q(0,\beta)} \tag{6}$$

where $Q(2i,\beta) = \int_0^\infty y^{2i} \exp\left[1 - [1 + y^2/2]^{1-\beta}\right] \mathrm{d}y$ and i is an integer. In general, both z_o and the exponent β can be time-dependent. The corresponding Kurtosis is given by

$$K(t,\beta) = \frac{\langle z^4(t,\beta)\rangle}{\langle z^2(t,\beta)\rangle^2} = \frac{Q(4,\beta)\; Q(0,\beta)}{Q^2(2,\beta)} \, . \tag{7}$$

Although $Q(2i,\beta)$ has no closed form, the following second-order approximations are useful for a nearly Gaussian DispD (*i.e.*, when $|\beta| \ll 1$):

$$\frac{\langle z^2\rangle}{z_o^2} \approx 1 + 1.815\beta + 3.121\beta^2 + ... \tag{8}$$

$$K - 3 \approx 1.0315\beta + 1.588\beta^2 + ... \, . \tag{9}$$

Interestingly, the excess Kurtosis $K-3$ is (to first order) approximately equal to the exponent β that determines the fatness of the tails of the DispD. Note that since β is expected to be a function of time while $\langle z^2\rangle \sim t$, the critical length scale z_o cannot increase like $\sim \sqrt{t}$. Finally, we note that for a Gaussian distribution with perfect exponential tails, our interpolating function predicts that $K(\beta = \frac{1}{2}) \approx 4.857$, which appears to agree with the limiting value reported in ref[30].

A. Raspberry colloidal particle diffusion

We use Langevin Dynamics as implemented in the ESPResSo package[31]. The HI are included by coupling the particle's velocity to a lattice Boltzmann fluid. We employ the raspberry particle model[32–34] shown in Fig 1b. The particle comprises $N = 454$ beads of size σ, for a total radius of $R = 3\,\sigma$ and a volume $V_o = \frac{4}{3}\pi R^3 \approx 113\,\sigma^3$. We freeze all beads relative to the center of mass using virtual momentum-transferring rigid bonds. The particle and the solvent share the same density, $\rho_s = m_o/\sigma^3$, where m_o is the mass of a one bead.

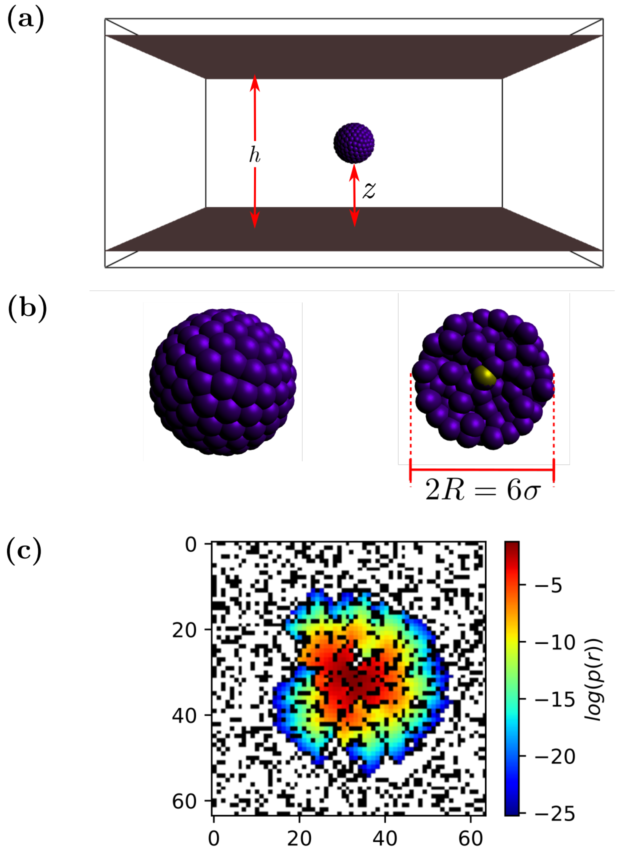


FIG. 1. The systems used to test the empirical fitting function. (a) Two repulsive boundaries are separated by a distance h, while z is the distance between the raspberry particle surface and the wall. (b) A raspberry particle of radius $R = 3\,\sigma$ comprising 454 coupled beads. The golden bead is at the center of mass. (c) A two dimensional $64{\times}64$ square lattice with $\phi = 30\%$ of the lattice sites occupied by obstacles (in black). The concentration profile, starting from the site at the center of the system, is shown after 30 time steps.

The implicit solvent is modeled by a GPU implementation of the three-dimensional 19 velocity LB method (D3Q19)[35]. The velocity $\boldsymbol{v}$ of a bead is coupled to the fluid via a drag force[35] $\boldsymbol{f}_{\gamma,i} = -\gamma(\boldsymbol{v}_i - \boldsymbol{u}_i)$, where $\boldsymbol{u}_i$ is the velocity of the fluid at the position where the bead resides (with $i = x, y, z$), and the coupling friction γ has to be to tuned to insure that the particle's hydrodynamic radius $R_H = R$.

The repulsive Weeks-Chandler-Andersen (W) potential[36] $U_{\mathrm{W}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \epsilon$ models the steric interactions between the raspberry beads and a wall when $r < r_c = 2^{1/6}\,\sigma$, while $U_W(r > r_c) = 0$ beyond the cutoff distance r_c. We use ϵ for our unit of energy, σ for length and $t_o = \sigma\sqrt{m_o/\epsilon}$ for time. The temperature is chosen to be $k_BT = \epsilon$ while the kinematic viscosity of the fluid is $\eta = 12\,\sigma^2/t_o$ and the coupling per bead is $\gamma = 15\,m_o/t_o$. These choices give a mean hydrodynamic radius $R_H \approx 2.98\,\sigma$ and a diffusion coefficient $D_o = 0.00148\,\sigma^2/t_o$ in free solution. We use an integration time step $\delta t = 0.005\,t_o$.

B. Displacements distributions

The DispD$_\perp$ $P(z,t)$ is computed for different time intervals and locations between two non-slip walls (Fig. 1b) separated by a distance $h = 30\,\sigma$. We start the particle in the center ($z = 13.5\,\sigma$) and let it diffuse freely until it reaches $z = \sigma/2$. The simulation times are long enough to generate 2000 uncorrelated sub-trajectories; we can then use any position in these trajectories as an effective initial position.

When a particle diffuses away from a given position for a brief period of time, its diffusivity is essentially constant during the resulting trajectory: our results (Appendix A) then show that it undergoes normal diffusion (MSD $\sim t$) but with a local anisotropic diffusivity $D(z)$, in agreement with eqs. 3-4. However, when the DispD$_\perp$ is averaged over initial positions located in a high $\nabla D_\perp(z)$ region (the diffusing diffusivity regime[37]), it includes both Gaussian and non-Gaussian components. Figure 2a shows examples for starting points located in the region $\frac{z}{R} \in [0.5, 2]$: the dashed Gaussian lines demonstrate the presence of fat tails, in agreement with ref.[30], while the solid lines show that eq. 5 provides an excellent fit for all values of the displacement Δz and all times. Nevertheless, the MSD is still increasing linearly with time – see Fig. 2b. The DispD remains Gaussian (within the limits of precision we can achieve) in the parallel direction (not shown) because this diffusivity gradient is weaker.

As previously mentioned, a common way to characterize such an "anomalous DispD" is to compute its Kurtosis. Figure 2c shows that the excess Kurtosis (both of the fitted function, $K_f - 3$, and of the raw data, $K_d - 3$) vanishes at short times (we then probe local regions with uniform diffusivity). We note that $K_f > K_d$ because the tails of the raw data distributions are heavily truncated here (at distances comparable to the length scale z_o).

Our fitting function includes additional information. For instance, we see that the exponent β mirrors the behavior of the Kurtosis. In fact, we note that $K_f(t) - 3 \approx \beta(t)$ when β is small in Fig. 2c, in agreement with eq. 9. The fit also provides information about the length scale z_o beyond which the Gaussian part (short distances) changes to fat tails (large distances): as shown in Fig. 2b, our results indicate that z_o^2 increases roughly as $\sim t^{0.84}$ over the time range used here. The fitting parameters thus indicate that the MSD increases linearly with time because the sublinear increase of z_o^2 is compensated by an increase of the tail anomalous exponent β over the time periods studied here.

IV. EXAMPLE II: ANOMALOUS DIFFUSION IN RANDOM SYSTEMS

We now examine the usefulness of the general form given by eq. 1 for a case of obstructed diffusion. We use the simplest model: a random walk on a two-dimensional square lattice with a fraction ϕ of the sites being randomly occupied by immobile obstacles (see Fig. 1c). In short, diffusion is expected to be normal (*i.e.*, the MSD grows linearly with time and the DispD is Gaussian) for short times (before the particle starts colliding with the obstacles; in practice, this regime only exists at very low obstacle concentration) and long times (the steady-state, achieved for times larger than the crossover time t^* and distances larger than the system's crossover length r^*). For intermediate times, the MSD grows roughly as t^γ, where the anomalous exponent $\gamma < 1$. See refs.[27,38] for previous studies of this system. Obviously, given the transient (or anomalous) regime mentioned above, we expect non-Gaussian distribution functions $P(r)$ (where $r^2 = x^2 + y^2$) unless $t \gg t^*$.

A. Random-Walk on a 2D Lattice

We randomly place a concentration ϕ of 1×1 obstacles on a 1600×1600 square lattice, and average over an ensemble of 500 different obstacle configurations for a value of ϕ. In order to obtain the distribution function $p(r, t)$ at different times t we use a standard Markov Chain Monte Carlo propagation algorithm: we initially place a unit concentration on the lattice site at the center of the system, and we propagate the concentration throughout the system using the master equation. Jumping probabilities $p_{\pm x} = p_{\pm y} = p = \frac{1}{8}$ are chosen,

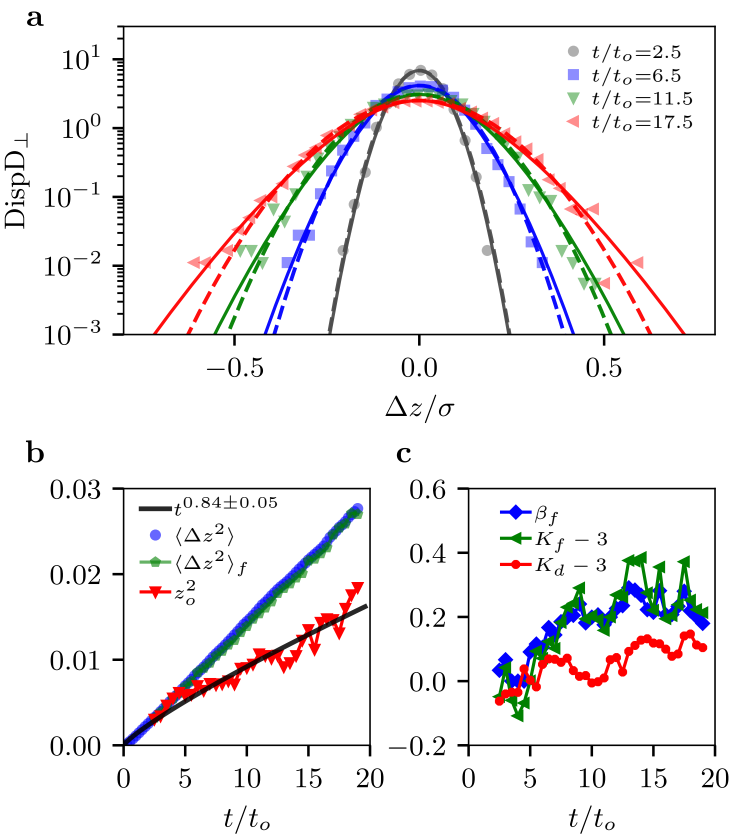


FIG. 2. Simulation data and fitting parameters for a spherical particle diffusing between two walls. (a) The $\mathrm{DispD}_\perp$ averaged over different starting positions $z/R \in [0.5, 2]$ at four different times. The simulation data are binned with a bin size $0.03\,\sigma$. The dashed lines are Gaussian fits while the solid lines show the interpolating function given by eq. 5. (b) Length scales as a function of time: $\langle \Delta z^2 \rangle$ is the perpendicular MSD calculated from the raw data, while the other two are obtained from the fits. The black curve shows a power law fit for z_o^2. (c) The excess Kurtosis $K_f - 3$ and the exponent β_f (as obtained from the fits) as a function of time; note that $K_d - 3$ is calculated directly from the raw data.

leaving a probability of not jumping of $p_0 = \frac{1}{2}$ at each time step. The free diffusion coefficient is set to $D = 1$, and the lattice spacing is set to $a = 1$. The time duration of a Markov step is then simply $\tau = \frac{a^2 p}{D} = \frac{1}{8}$. The resulting distribution functions (a few examples are shown in Fig. 3), at specific times, are then fitted with both a Gaussian and the fitting function

$$P(r) = \frac{2(1-A)}{\sqrt{\pi r_o^2}}\ \exp\left[1 - \left[1 + \left(\frac{r}{r_o}\right)^{2-\alpha}\right]^{1-\beta}\right] \tag{10}$$

(the Gaussian distribution corresponds to $A = \alpha = \beta = 0$). Once the fitting parameters are obtained, it is possible to compute the even moments of the radial displacement

$$\langle r^{2i} \rangle = r_o^{2i} \times \frac{G(2i, \alpha, \beta)}{G(0, \alpha, \beta)} \tag{11}$$

where $G(2i, \alpha, \beta) = \int_0^\infty y^{2i+1} \exp\left[1 - [1 + y^{2-\alpha}]^{1-\beta}\right] \mathrm{d}y$. In principle, the only explicit time-dependence is in the length scale $r_o(t)$, with $r_o^2(t) = \langle r^2(t) \rangle = 4Dt$ for normal diffusion in two dimensions; however, the exponents α and β can also vary with time, as we shall see. The 2D Kurtosis is then given by

$$K_2(\alpha, \beta) = \frac{G(4, \alpha, \beta)\ G(0, \alpha, \beta)}{G^2(2, \alpha, \beta)}\ . \tag{12}$$

Although $G(2i, \alpha, \beta)$ has no closed form, the following approximations are useful for a nearly Gaussian DispD (*i.e.*, when $\alpha \ll 1$ and $|\beta| \ll 1$):

$$\frac{\langle r^2 \rangle}{r_o^2} \approx 1 + 0.7114\,\alpha + 2\,\beta + ... \tag{13}$$

$$K_2 \approx 2 + \tfrac{1}{2}\,\alpha + 0.5963\,\beta + ...\ . \tag{14}$$

Interestingly, the excess Kurtosis $K_2 - 2$ is (to first order) approximately equal to the mean value of the two exponents, $(\alpha + \beta)/2$. It is also possible to compute the moments of the distribution itself (treating the problem as a one-dimensional distribution). We then obtain the following 1D Kurtosis:

$$K_1 = \frac{G(3, \alpha, \beta)\ G(-1, \alpha, \beta)}{G^2(1, \alpha, \beta)} \approx 3 + \alpha + 1.0315\,\beta + ...\ . \tag{15}$$

We will be using K_2 in our analysis below.

B. Displacement distributions *vs* time

Our fitting function does not work well at short times in Fig. 3 because of the coarse lattice discretization effects for short displacements (although it is actually slightly better than a Gaussian fit). Once we approach the crossover time t^* (this time marks the transition between the transient and steady-state regimes; see ref.[39] for details regarding the method used to obtain t^* from the data) the tails are captured very well by our fit while the Gaussian fits are clearly inadequate. In fact, the Gaussian fits slowly converge with the data when $t \gg t^*$, as expected.

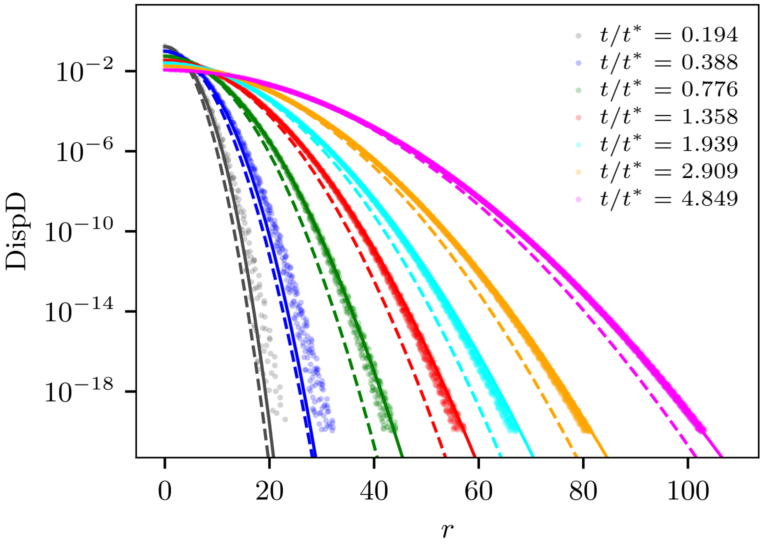


FIG. 3. The DispD at six different times for the two-dimensional obstructed diffusion problem. The simulation data are not binned (every possible distance on a square lattice is considered, hence the apparent noise in the otherwise exact data). The dashed lines are Gaussian fits while the solid lines show the interpolating function given by eq. 10. The obstacle concentration is $\phi = \frac{1}{9}$. Note that the transient/anomalous exponent is $\gamma = 0.963$ for this system, while the crossover length and time are $r^* = 6.95$ and $t^* = 16.11$ respectively.

We examine how the fitting parameters evolve as a function of inverse scaled time t^*/t in Fig. 4. As discussed above, we expect the distribution to become more Gaussian (which means the exponent characterizing the tail of the distribution, $(1-\beta)(2-\alpha)$, converges towards 2) as time increases, and fully Gaussian distributions (*i.e.*, $\alpha = \beta = 0$) for times $t \gg t^*$. Indeed, these parameters, as well as the Kurtosis, converge towards their Gaussian limits when $t \gg t^*$ in Fig. 4. This agrees with Fig. 3 where it is clear that the fits are becoming more Gaussian as t/t^* increases. We also note that the excess Kurtosis $K_d - 2$ is essentially given by the mean $(\alpha+\beta)/2$ over the whole time interval.

However, the distribution function is not expected to be Gaussian over distances $r < r^*$ (or times $t < t^*$) since diffusion is anomalous over these length scales (*i.e.*, we then have $\langle r^2(t)\rangle \sim t^\gamma$, with $\gamma = 0.961$ here). In Fig. 4 we see two transitions in the time dependence of the fitting exponent β, namely a change in the sign of the slope at $t^*/t \approx 0.5$, and a change in sign at $t \approx t^*$. The value of α, on the other hand, remains positive and decreases

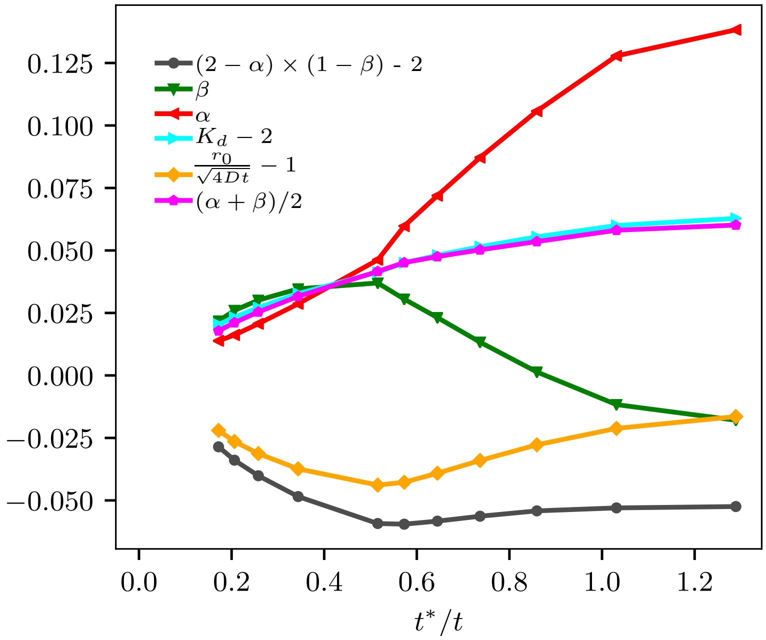


FIG. 4. Fitting exponents α and β, from eq. 10, *vs* inverse rescaled time t^*/t. We also show that the Kurtosis of the distribution, K_d, and the product $(2-\alpha)(1-\beta)$ both converge to the Gaussian value of 2 at long time. The average of our two exponents α and β surprisingly return $K_d - 2$. Finally, the fitting length scale r_o clearly approaches the diffusion length $\sqrt{4Dt}$ in the asymptotic limit.

very rapidly with time. We also note that the product $(2-\alpha)(1-\beta) - 2 \approx -0.05$ is roughly constant for $t < t^*$, implying that the tail of the distribution decays roughly as $\sim \exp[-(r/r_o)^{1.95}]$ during the anomalous/transient regime.

Of great interest is the length scale $r_o(t)$ that characterizes the tail of the distribution (see eq. 2). In Fig. 4 we renormalize the value of $r_o(t)$ by the long time limit $r_0(t) = \sqrt{4Dt}$. The ratio $r_o(t)/\sqrt{4Dt}$ should thus converge towards unity when time increases: this is precisely what the diamond (orange) data points show. At short times, however, we expect anomalous diffusion with $\langle r^2(t)\rangle \sim t^\gamma$, hence the non-monotonic behavior. Finally we find that the characteristic length scale $r_o(t)$ is equal to the crossover length r^* at $t \approx 1.06\, t^*$, showing the connection between the anomalous diffusion regime and the tails of the DispD.

V. CONCLUSION

In this article, we propose a new interpolating function that can conveniently characterize displacement distributions that contain both Gaussian and non Gaussian exponential components. The key advantage of our function is that it can describe distributions that have core and tail components with different behaviours, and yet it includes only two additional fitting parameters (compared to the Gaussian fit).

We tested our interpolating function using two simple examples. In the first test, the physics of the problem is such that the distribution has to be Gaussian at short distances, but may have non-Gaussian tails. In the second test, the distribution is more general and the fits provided detailed information about its time evolution. We are currently studying how the fit parameters (especially the length scale r_o and tail exponent $(2-\alpha)(1-\beta)$) are connected to fundamental elements of the physics of these two problems such as the local diffusivity gradient (in the case of the first example) and the anomalous exponent γ (the second example). It is interesting to note that a recent paper by Miotto et al.[18] discusses the time-dependence of the length scale λ related to the tail of the DispD for a problem of "Brownian yet Non-Gaussian Dynamics"; their main conclusion is that λ does not scale diffusively with time, just like z_o and r_o in our two examples.

ACKNOWLEDGMENTS

GWS acknowledges the support of both the University of Ottawa and the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN/046434-2013. LQ is supported by the Chinese Scholarship Council and the University of Ottawa. NI is partially supported by the University of Ottawa.

Appendix A: Diffusion coefficient at different heights

Figure 5 shows the diffusion coefficients $D_\perp$ and $D_\parallel$ measured at different distances z from the wall. Our data are in a good agreement with theory (eqs. 3 and 4), thus validating our simulation approach.

The inset shows that we indeed obtain a Gaussian DispD$_\perp$ for $t = 12.5\, t_o$, with a variance

$\sim D_{\perp} t_o$, for different initial positions z/R.

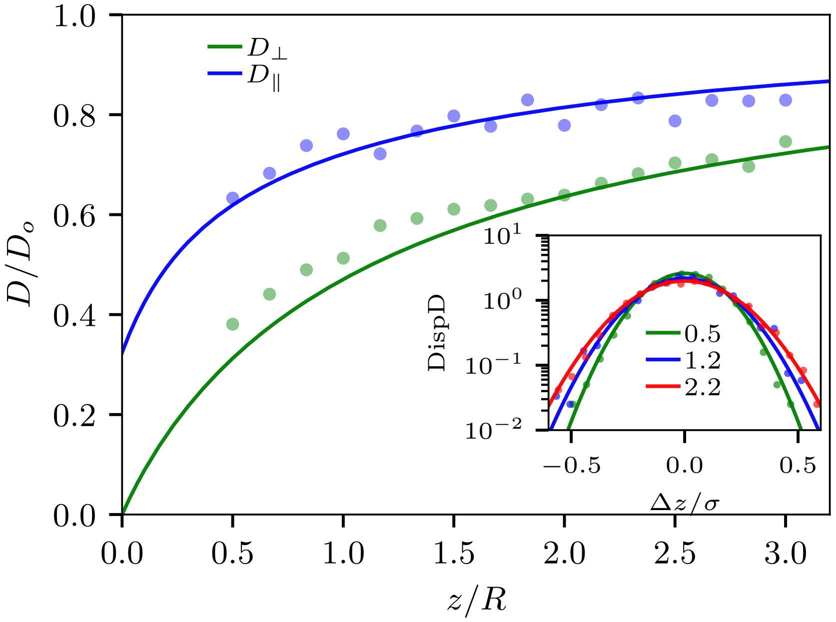


FIG. 5. Reduced diffusion coefficients parallel ($D_{\parallel}/D_o$) and perpendicular ($D_{\perp}/D_o$) to the wall *vs* the vertical distance z/R from the closest wall, where D_o is the free solution diffusion coefficient. The data points are obtained by fitting the time dependence of the MSD between $t = 0$ and $t = 20\, t_o$. The solid lines are from eqs. 3 and 4. Inset: Gaussian fit of the vertical DispD$_{\perp}$ for three different initial positions z/R=0.5, 1.2 and 2.2.

---

[1] R. F. Sabiryanov, M. I. Larsson, K. J. Cho, W. D. Nix, and B. M. Clemens, Phys. Rev. B **67**, 125412 (2003).

[2] X. Xu, J.-N. Aqua, and T. Frisch, J. Phys. Condens. Matter **24**, 045002 (2012).

[3] M. Ignacio, Y. Saito, P. Smereka, and O. Pierre-Louis, Phys. Rev. Lett. **112**, 146102 (2014).

[4] C. Mejía-Monasterio, S. Nechaev, G. Oshanin, and O. Vasilyev, New J. Phys. **22**, 033024 (2020).

[5] P. Sharma, S. Ghosh, and S. Bhattacharya, Appl. Phys. Lett. **97**, 104101 (2010).

[6] A. Banerjee and K. D. Kihm, Phys. Rev. E **72**, 042101 (2005).

[7] M. A. Bevan and D. C. Prieve, J. Chem. Phys. **113**, 1228 (2000).

[8] E. R. Dufresne, T. M. Squires, M. P. Brenner, and D. G. Grier, Phys. Rev. Lett. **85**, 3317 (2000).

[9] C. Xue, X. Shi, Y. Tian, X. Zheng, and G. Hu, Nano Lett. **20**, 3895 (2020).

[10] R. K. Singh, J. Mahato, A. Chowdhury, A. Sain, and A. Nandi, J. Chem. Phys. **152**, 024903 (2020).

[11] A. Sabri, X. Xu, D. Krapf, and M. Weiss, Phys. Rev. Lett. **125**, 058101 (2020).

[12] J. L. A. Dubbeldam, A. Milchev, V. G. Rostiashvili, and T. A. Vilgis, EPL **79**, 18002 (2007).

[13] R. Metzler and J. Klafter, Biophys. J. **85**, 2776 (2003).

[14] Y. Kantor and M. Kardar, Phys. Rev. E **76**, 061121 (2007).

[15] G. Guigas and M. Weiss, Biophys. J. **94**, 90 (2008).

[16] H. Sanabria, Y. Kubota, and M. N. Waxham, Biophys. J. **92**, 313 (2007).

[17] D. S. Banks and C. Fradin, Biophys. J. **89**, 2960 (2005).

[18] J. M. Miotto, S. Pigolotti, A. V. Chechkin, and S. Roldán-Vargas, Phys. Rev. X **11**, 031002 (2021).

[19] F. Zhou, H. Wang, and Z. Zhang, Langmuir **36**, 11866 (2020).

[20] W. Wang, F. Seno, I. M. Sokolov, A. V. Chechkin, and R. Metzler, New J. Phys. **22**, 083041 (2020).

[21] B. Wang, S. M. Anthony, S. C. Bae, and S. Granick, PNAS **106**, 15160 (2009).

[22] B. Wang, J. Kuo, S. C. Bae, and S. Granick, Nat. Mater. **11**, 481 (2012).

[23] A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov, Phys. Rev. X **7**, 021002 (2017).

[24] R. Metzler, Eur. Phys. J. Spec. Top. **229**, 711 (2020).

[25] T. Nagai, S. Tsurumaki, R. Urano, K. Fujimoto, W. Shinoda, and S. Okazaki, J. Chem. Theory Comput. **16**, 7239 (2020).

[26] S. Hapca, J. W. Crawford, and I. M. Young, J. R. Soc. Interface. **6**, 111 (2009).

[27] M. J. Saxton, Biophys. J. **66**, 394 (1994).

[28] R. Metzler and J. Klafter, Phys. Rep. **339**, 1 (2000).

[29] H Faxen, Ark. Mat., Astron. Fys **18**, 29 (1924).

[30] M. Matse, M. V. Chubynsky, and J. Bechhoefer, Phys. Rev. E **96** (2017), 10.1103/PhysRevE.96.042604, arXiv:1706.02039.

[31] F. Weik, R. Weeber, K. Szuttor, K. Breitsprecher, J. de Graaf, M. Kuron, J. Landsgesell, H. Menke, D. Sean, and C. Holm, Eur. Phys. J. Spec. Top. **227**, 1789 (2019).

[32] L. P. Fischer, T. Peter, C. Holm, and J. de Graaf, J. Chem. Phys. **143**, 084107 (2015), arXiv:1503.02671.

[33] J. de Graaf, T. Peter, L. P. Fischer, and C. Holm, J. Chem. Phys. **143**, 084108 (2015), arXiv:1503.02681.

[34] P. Kreissl, C. Holm, and R. Weeber, Soft Matter **17**, 174 (2021).

[35] B. Dünweg, U. D. Schiller, and A. J. C. Ladd, Phys. Rev. E **76**, 036704 (2007).

[36] J. D. Weeks, D. Chandler, and H. C. Andersen, J. Chem. Phys. **54**, 5237 (1971).

[37] M. V. Chubynsky and G. W. Slater, Phys. Rev. Lett. **113** (2014), 10.1103/PhysRevLett.113.098302.

[38] G. W. Slater and H. L. Guo, Electrophoresis **17**, 1407 (1996).

[39] N. Ilow and G. W. Slater, "Estimating the steady state diffusion coefficient using data from the transient anomalous regime," (2021), arXiv:2108.04944 [physics.bio-ph].

CHAPTER 6

CONCLUSION

We investigated the dynamics of a single tracer particle diffusing on a two-dimensional Monte Carlo lattice model in the presence of immobile obstacles. Due to the fact that purely random or periodic obstacle configurations are unlikely in nature, we focused our research on understanding how configurations between random and periodic, which we refer to as "Fuzzy", affect the dynamics of a tracer particle.

We began by revising the Monte Carlo algorithm, and specifically investigating the impact of p, the hopping probability of the tracer particle to an adjacent lattice site. We found that the most common choice ($p = 1/4$) was not optimal due to the fourth moment of the displacement distribution disagreeing with the solution of the diffusion equation for a free particle. We studied the fourth moment as a function of p and found agreement when $p = 1/8$. Consequently the time-step in our simulations is reduced; this means we lose access to very long time data. However the improved time resolution and access to earlier time data is beneficial for our work due to their roles when investigating short time dynamics, such as particle diffusion during the anomalous regime.

There are many ways to generate disorder in systems with immobile obstacles. We chose five rather simple yet different methods. However the characteristic disorder parameters of each method are not easily comparable. To address this problem we introduced a disorder parameter F_z which works for all methods. To our surprise, when studying the Pick Up and Drop method at a fixed concentration we found that all output parameters (the crossover length r^* , the anomalous exponent α, and the steady-state diffusion coefficient D) yield linear interpolations between their periodic and random limits as a function of F_z. We were unable to find a clear reason as to why this occurs; further work investigating the connection between F_z and the fraction of randomly replaced obstacles may shed light on this intriguing relationship. In general we found F_z to be a non-linear function of the characteristic disorder parameters of each method, suggesting a non-trivial transition between the periodic and random limit cases. When studying key parameters as functions of the obstacle concentration ϕ whilst keeping F_z, constant we found that there was a critical value of F_z which signalled the transition from Period to Random behaviour ($F_z \approx 2/3$). Finally we found that clustering is a key component in the dynamics of our tracer particle, as the behaviour for the No Cluster method deviates greatly from all other methods.

The anomalous transient regime is most often described by a single parameter, the anomalous exponent α, which is often extracted using an arbitrary time frame. We proposed to use the inflection point as an objective way to pinpoint the center of the anomalous regime (an idea previously introduced and studied by Neo Nguiya Passi [21]), from which we extract the anomalous exponent. We then built upon our understanding of the transient regime by introducing a second parameter, the anomalous regime width Σ_t. Initially our goal was to better characterize the transient regime, and to assess the region surrounding the inflection point that could be well approximated by a straight line (as suggested by the theory of anomalous diffusion). However, we found that the approximation of the width acts as a very good estimate of the crossover time t^*. Using the crossover time t^* we were able to relate the steady-state diffusion coefficient D to the anomalous coefficient D_α and anomalous exponent α. For our goal of approximating the steady-state diffusion coefficient, this relationship would not be useful without our approximation of t^* using Σ_t due to the need for long time data to extract t^*. However the anomalous regime width Σ_t being approximated using data around the inflection point allowed us to estimate the steady-state diffusion coefficient solely using parameters that are extracted from data near the inflection point, namely α, Σ_t, and D_α.

Understanding heavy tailed distributions and anomalous diffusion has widespread applications in economics [23], online gambling [36], and aging systems [20]. We investigated the displacement distribution functions on log-log plots, similar to Plerou et al's [23] analysis of an ink particles cumulative probability distribution against the number of random walk steps within a time step Δt. Our findings agree with those of Plerou et al's; e.g., at large distances the distributions exhibit heavy tails. We confirmed the approach to Gaussian behaviour for long times by studying the saturation of the fitting parameters to their respective Gaussian limits as $t \to \infty$. However the main benefit of our work on fat tailed distributions is the ability to characterize both the short and large distance parts of the displacement distributions, independent of time, with one fitting function and only two additional fitting parameters; otherwise the different regions must be fit separately [23, 35, 36].

In summary, we studied anomalous diffusion in the context of a two-dimensional lattice populated with immobile obstacles (acting as reflecting boundary conditions) configured in different ways (random, periodic, "Fuzzy"). We found a critical value of disorder beyond which parameters became independent of the method by which disorder was generated. Certain parameters showed key transitions from their periodic to random limit behaviour (e.g. r^* became a weak function of ϕ), and non-trivial concentration dependent effects arose. We presented a fitting function to capture both the short range behaviour and fat tail effects of anomalous diffusion simultaneously. We proposed that this function can be used to conveniently fit mixed cumulative probability distributions in the future. Our most significant result was the walker/obstacle size-independent estimate of the steady-state diffusion coefficient using the anomalous regime width as a proxy for the crossover time. This proxy allowed us to bypass the need for late time data to extract a crossover time; thus we hypothesized it may be useful to experimentalists to study the steady-state of their system when access to late time data is unavailable. However we must first study the effects of noise on our estimate, as experimentalists will not have the same precision we do in our simulations.

REFERENCES

[1] Mie Andersen, Chiara Panosetti, and Karsten Reuter. A practical guide to surface kinetic monte carlo simulations. *Front. Chem.*, 7:202, 2019.

[2] Louis Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure*, 3e série, 17:21–86, 1900.

[3] Robert Brown. XXVII. a brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine*, 4(21):161–173, 1828.

[4] Timothy A Davis. Algorithm 832: Umfpack v4. 3—an unsymmetric-pattern multifrontal method. *ACM Trans. Math. Softw.*, 30(2):196–199, 2004.

[5] Aryeh Dvoretzky and Paul Erdös. Some problems on random walk in space. In *Proc. Second Berkeley Symp. Math. Statist. Probab*, pages 353–367, 1951.

[6] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4, 1905.

[7] Paul Erdős and S James Taylor. Some problems concerning the structure of random walk paths. *Acta Mathematica Academiae Scientiarum Hungarica*, 11(1-2):137–162, 1963.

[8] Herbert Freundlich and Deodata Krüger. Anomalous diffusion in true solution. *J. Chem. Soc. Faraday Trans.*, 31:906–913, 1935.

[9] Ido Golding and Edward C Cox. Physical nature of bacterial cytoplasm. *Physical Review Letters*, 96(9):098102, 2006.

[10] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, and Nathaniel J. Smith et al. Array programming with NumPy. *Nature*, 585(7825):357–362, 2020.

[11] Radu Herbei, Rajib Paul, and L Mark Berliner. Applying diffusion-based markov chain monte carlo. *PloS one*, 12(3):e0173453, 2017.

[12] Nicolas E Humphries, Nuno Queiroz, Jennifer RM Dyer, Nicolas G Pade, Michael K Musyl, Kurt M Schaefer, Daniel W Fuller, Juerg M Brunnschweiler, Thomas K Doyle, Jonathan DR Houghton, et al. Environmental context explains lévy and brownian movement patterns of marine predators. *Nature*, 465(7301):1066–1069, 2010.

[13] Aaron J Ihde. The Karlsruhe congress: a centennial retrospective. *J. Chem. Educ.*, 38(2):83, 1961.

[14] Jesper Lykke Jacobsen. Critical points of Potts and O(N) models from eigenvalue identities in periodic temperley–lieb algebras. *J. Phys. A Math. Theor.*, 48(45):454003, 2015.

[15] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nat. Methods*, 5(8):695–702, 2008.

[16] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001.

[17] Vicenç Méndez, Daniel Campos, and Frederic Bartumeus. *Stochastic foundations in movement ecology*. Springer, 2016.

[18] Jean-François Mercier, Gary W Slater, and Hong L Guo. Numerically exact diffusion coefficients for lattice systems with periodic boundary conditions. I. theory. *J. Chem. Phys.*, 110(12):6050–6056, 1999.

[19] Ralf Metzler, Jae-Hyung Jeon, Andrey G Cherstvy, and Eli Barkai. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *PCCP*, 16(44):24128–24164, 2014.

[20] Philipp G Meyer, Vidushi Adlakha, Holger Kantz, and Kevin E Bassler. Anomalous diffusion and the moses effect in an aging deterministic model. *New J. Phys.*, 20(11):113033, 2018.

[21] Neo Nguiya Passi. Modelling diffusion through environments that contain immobile obstacles: The short-time transient regime, anomalous diffusion and crowding. Université d'Ottawa/University of Ottawa, Master's Thesis, 2019.

[22] Peter J Olver. *Introduction to partial differential equations*. Springer, 2014.

[23] Vasiliki Plerou, Parameswaran Gopikrishnan, Luís A Nunes Amaral, Xavier Gabaix, and H Eugene Stanley. Economic fluctuations and anomalous diffusion. *Phys. Rev. E*, 62(3):R3023, 2000.

[24] Michael J Saxton. Anomalous diffusion due to obstacles: a monte carlo study. *Biophys. J.*, 66(2):394–401, 1994.

[25] Michael J Saxton. Single-particle tracking: connecting the dots. *Nat. Methods*, 5(8):671–672, 2008.

[26] Hao Shen, Lawrence J Tauzin, Rashad Baiyasi, Wenxiao Wang, Nicholas Moringo, Bo Shuang, and Christy F Landes. Single particle tracking: from theory to biophysical applications. *Chem. Rev.*, 117(11):7331–7376, 2017.

[27] Kurt E Shuler and Udayan Mohanty. Random walk properties from lattice bond enumeration: Anisotropic diffusion in lattices with periodic and randomly distributed scatterers. *PNAS*, 78(11):6576–6578, 1981.

[28] David W Sims, Emily J Southall, Nicolas E Humphries, Graeme C Hays, Corey JA Bradshaw, Jonathan W Pitchford, Alex James, Mohammed Z Ahmed, Andrew S Brierley, Mark A Hindell, et al. Scaling laws of marine predator search behaviour. *Nature*, 451(7182):1098–1102, 2008.

[29] Gary W Slater and Hong L Guo. An exactly solvable ogston model of gel electrophoresis: I. the role of the symmetry and randomness of the gel structure. *Electrophoresis*, 17(6):977–988, 1996.

[30] Gary W Slater and Hong L Guo. An exactly solvable ogston model of gel electrophoresis ii. sieving through periodic gels. *Electrophoresis*, 17(9):1407–1415, 1996.

[31] Dietrich Stauffer and Ammon Aharony. *Introduction to percolation theory*. CRC Press, 2018.

[32] Jean-Yves Tinevez, Nick Perry, Johannes Schindelin, Genevieve M Hoopes, Gregory D Reynolds, Emmanuel Laplantine, Sebastian Y Bednarek, Spencer L Shorte, and Kevin W Eliceiri. Trackmate: An open and extensible platform for single-particle tracking. *Methods*, 115:80–90, 2017.

[33] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.

[34] Alex von Diezmann, Yoav Shechtman, and WE Moerner. Three-dimensional localization of single molecules for super-resolution imaging and single-particle tracking. *Chem. Rev.*, 117(11):7244–7275, 2017.

[35] Xiangwen Wang and Michel Pleimling. Behavior analysis of virtual-item gambling. *Phys. Rev. E*, 98(1):012126, 2018.

[36] Xiangwen Wang and Michel Pleimling. Online gambling of pure chance: wager distribution, risk attitude, and anomalous diffusion. *Sci. Rep.*, 9(1):1–17, 2019.

[37] Eric R Weeks, TH Solomon, Jeffrey S Urbach, and Harry L Swinney. Observation of anomalous diffusion and Lévy flights. In *Lévy flights and related topics in physics*, pages 51–71. Springer, 1995.

[38] Rand R Wilcox. *Fundamentals of modern statistical methods: Substantially improving power and accuracy*. Springer, 2010.