Technische Universität München

THESIS

To obtain the grade of

Doctor of the Université Pierre et Marie Curie and of the Technische Universität München

Speciality : **Physics**

Presented by

Vincent TEJEDOR

Thesis jointly supervised by:
Dr. Raphaël VOITURIEZ,
Pr. Dr. Ralf METZLER,
Dr. Olivier BÉNICHOU

prepared in **Laboratoire de Physique Théorique de la Matière Condense, CNRS/UPMC and Physics Department T30g, TUM** and **Doctoral School ED389 - P2MC**

Random walks and first-passage properties

Trajectory analysis and search optimization

Thesis defended on **July 3, 2012**,

Dr. Pierre LEVITZ ..... President
PECSA, UPMC (Paris).

Pr. Jean-François DUFRÊCHE ..... Referee
Institut de Chimie Séparative de Marcoule (Bagnols-sur-Cèze, France).

Pr. Tapio ALA-NISSILÄ ..... Referee
Department of Applied Physics, Aalto University (Finland).

Pr. Dr. Friedrich SIMMEL ..... Examiner
Physic Department E14, TUM (München)

Dr. Raphaël VOITURIEZ ..... Thesis Advisor
LPTMC, UPMC (Paris).

Pr. Dr. Ralf METZLER ..... Thesis Advisor
Physic Department T30g, TUM (München)

Dr. Olivier BÉNICHOU ..... Guest
LPTMC, UPMC (Paris).

TECHNISCHE UNIVERSITÄT MÜNCHEN

UPMC PARISUNIVERSITAS

Acknowledgments

This manuscript is highly non Markovian: it is not a solitary work, but the result of a succession of interactions with different peoples, a slow maturation of a subject, namely the first-passage properties of random walkers, during several years. Instead of just thanking the people that helped me to find some targets in the maze of the random world, I would like to explain how this manuscript came to life.

In 2003, I discovered what the scientific research was, and through Hakim, I began to like questions without answers. During the classes of statistical physics, Pr. Jean-François Allemand showed me the fun of ensemble average and the subtleties of the single particles experiments. After six month in Osaka, where Pr. Toshio Yanagida made me understand how hard it was to perform a clean experiment, I chose to focus on more theoretical works. Pr. Emmanuel Maisonhaute unveiled for me some aspects of theoretical electrochemistry, but due to some random problem, he was unable to propose me a Master thesis. He addressed me to one of his friend, Dr. Olivier Bénichou, for the Master internship. Through Olivier, I met Sylvain, Claude, Alexis, Laura, Bob, and the other PhD students of LPTMC, as well as Raphaël. During those six months of Master thesis, I understood that theoretical problems could be even harder to solve than those of single particles experiments. I think that I arrived too young this first time at the LPTMC: I was surrounded by people too bright for me, and I lost faith in myself, being sure that I could never match them.

In 2007, I decided to become an engineer, and integrated the Corps des Mines. I met new comrades, and went during one year to a "real" company, GRTgaz, for the first time in my life. My tutor, Vincent, and my colleagues tried to convince me that big companies were a good place to work. It did not worked: too many meetings, too few intellectual challenges. After this internship, I decided to go back in research. I had to go abroad, and Olivier tipped me with Pr. Dr. Ralf Metzler. There, with Leila, Radina, Jae-Hong and Michaella, I discovered the anomalous diffusion, the fractional calculus, and the life in München. After two months in Lyngby, where Pr. Tomas Bohr welcomed me, I came back in Paris. During one more year, I followed engineer classes, studied with Jessica the french research system, and did somehow some researches for my PhD.

In 2010, I got my engineering degree, and started a Masters in Law. I was old enough to stomach the fact that each of my advisor was thrice time better than me, but I had lost faith in the research system. I could not figure how I would ever get a position in theoretical physics, even after several post-doc jobs. It was hopeless, I decided to prepare my reconversion. New formation, new comrades, I found out new intellectual challenges based on a different ground with Pr. Michel Verpeaux: interactions between science, constitutional and environmental law. New students came in the LPTMC, Thibault, Thomas, Jean-François and Clément, and I continued slowly my PhD work, using the ideas I had during discussions with friends, like Arnaud, Antonin, Grégoire or Delphine.

At last in 2011, I was drawn in a start-up with Benjamin and Jean-Charles, Expliseat SAS, that aim to design aircraft seats. My wife Fanny gave birth to a wonderful child, Olivier, and my research activities suffered from all those events. A fourth year was necessary to obtain a work sufficient for a PhD defense.

After finishing this manuscript, I realize that this whole work is based on a succession of random opportunities, that I followed without thinking where it would lead me. When I was alone in München, I tried really hard to solve some problems, and I found out that my efficiency was not really correlated to the time I spent in front of an equation: I had better results when I was splitting my time between various activities, and more ideas when I was alternating research, formation, discussion with friends, manual activities and sport. Maybe that is the real result of this PhD: to find an hidden target, one has to do anything but searching it, and then to interact with various people while keeping an eye – and the mind – open.

Contents

Introduction

With the global financial crisis that lasts since 2007, everybody has become familiar with mathematicians that use powerful tools to predict the behavior of random variables. Those variables can be the share prices, the exchange rates, the probability that an average American reimburses a loan or that an European country fails to reimburse his public debt. We can indeed consider that the exchange rate between Euro and US Dollar for instance is a random variable, that we note x, namely a quantity that can take (almost) any real value, $x \in [0, \infty[$. This random variable depends on time, and makes "jumps" each day between the value at t and the value at $t + dt$. We can plot this random variable as a function of time, and try to predict its value in the future. Lets assume that we buy 12 \$ with 10 € before a trip to the US, and that the exchange rate is 0.8 \$ for 1 € at our return. How long will we have to wait before the exchange rate passes above 1.2 \$ for 1€, for the first time? We will call this time the first-passage time, starting from 0.8 \$/€ to 1.2 \$/€. To compute the average value of this random variable, the mean first-passage time (MFPT) we will need to characterize the random process, and then to use the equation satisfied by this process (diffusion equation for instance) with the correct initial and boundary conditions.

First-passage properties in general, and MFPT in particular, are widely used in the context of diffusion-limited processes [168], either in chemistry [169], in biology [183], in electrical blackout spreading [53], in epidemiology [129], for foraging animals [24, 220]... They play a crucial role in those real situations, the random process modeling a transport process in disordered media[71, 102], a neuron firing dynamics[213], the spreading of diseases[129] or of computer viruses[155], or a target search process[22]. For each of those examples, a random walker trigger an event when reaching for the first time a given target, and it can be advantageous to minimize (transport process) or to maximize (viruses spreading) this first passage time. When several targets are available, for instance when a given protein can react with several different target in a living cell, it becomes important to know which one is reached first.

We will focus along this manuscript on proteins evolving in living cells, but the results we will obtain can be directly transposed to the above mentioned situations. In this biological context, modern experimental set-ups allow to follow single particles trajectories. After a general introduction to the random world for non-specialist in the first chapter, we will answer in the second and in the third chapter to two main questions. First, using a set of single particles trajectories, what can we say about the underlying mechanism? Once this mechanism is known, how can we optimize the MFPT?

Since they are the key parameter to quantify transport processes, the first-passage properties have been extensively studied during the last decades, and a lot of results were already written in textbooks before the beginning of this thesis. Concerning the mean first-passage time, a formal result for Markovian processes (processes without memory) have been found years ago [6, 148]. Other first-passage observables can also be expressed in formal ways, like splitting probabilities [65] or occupation times [67]. For a long time, these results were explicit for unbounded environments, or in (pseudo)-unidimensional confined media. The formal expression have been made usable – for scale-invariant processes in a general confined environment – thanks to an approximation of pseudo-Green functions by Sylvain Condamin [66], during his thesis under the direction

of Olivier Bénichou and Raphaël Voituriez. We will recall these results as well as the pseudo-Green function approximation in the first chapter of this manuscript. First-passage properties of non scale-invariant processes remain so far elusive quantities, and we will study them during this manuscript. We will also focus on random walk on discrete networks, in order to get more insight on the important parameters of first-passage properties that have not been identified as such yet, like target connectivity or network topology.

Real processes are not always purely Brownian: in the last few years, non-Brownian behaviors have been observed in an increasing number of systems [172, 173], ranging from physics [122, 187] or geophysics [186] to biology [99, 212]. In particular, living cells provide striking examples for systems where this behavior has been repeatedly observed experimentally, either in the cytoplasm [54, 99, 212, 232], the nucleus [160, 223] or the plasmic membrane [94, 127, 199]. We will thus consider non-Brownian processes, like continuous time random walk (CTRW) [115, 172] or fractional Brownian motion (fBm) [137]. Such processes can lead to anomalous diffusion, namely to a mean square displacement $\langle r^2 \rangle$ scaling like t^α, where $\alpha \neq 1$ is the anomalous diffusion exponent. If $\alpha < 1$, the process is called sub-diffusive, if $\alpha > 1$, super-diffusive. The mathematical description of anomalous diffusion [41], popularized by Ralf Metzler [172], will be introduced in the first chapter.

For generic random processes, we can rephrase the two questions that will structure this manuscript as follows:

- How can we discriminate between several (sub)-diffusion mechanism using a finite data set of single trajectories?
- Knowing the diffusion mechanism, how can we calculate the search time, and then optimize it, or similarly how can we calculate and minimize a given MFPT?

The first question will be the core of the second chapter of this manuscript. Many direct observation of diffusing proteins in living cells exhibit anomalous diffusion. However, the microscopical origin of sub-diffusion in cells remains debated, even if believed to be due to crowding effects in a wide sense as indicated by *in vitro experiments* [8, 9, 98, 228].

Several models can cause sub-diffusion, and each model rely on a different kind of interaction between the random walker and its environment. We will focus during the second chapter on the three prominent models mentioned above: diffusion on fractals, CTRW and fBm. Diffusion on fractals model assumes that the space available is fractal, namely that dead ends and bottlenecks exist at all scales. CTRW model assumes that the random walker interacts with its environment, and can be trapped somehow for a very long time. At last fBm model assumes that the trajectory possesses long-term spatial correlation, a "memory" of the past steps. This can be related to repeated interactions with a very wide object, like a long polymer (DNA) or a biological membrane. We can gain a better understanding on the environment surrounding the random walker if we can discriminate between those three models. Some tools have been proposed in the literature, like the normalized variance [74] or the ergodicity breaking parameter [104], but those observables are not adapted to small data sets. We will develop tests that allow to discriminate between those three models with realistic data sets, and taking into account the experimental constraints. We will also focus on alternative models, similar to diffusion on fractals, CTRW or fBm, but that are not "infinite", and that could explain better our observations of lipid granules diffusion within fission yeast cells. Indeed, the diffusion on fractals model assumes that the process is infinitely scale-invariant, even in non physical scales (femtometric or kilometric scales

do not make sense for a living cell), CTRW assumes that the waiting time can be infinitely long (a waiting time of a century has very little sense for a protein), and fBm assumes that the trajectory is infinitely correlated in space. Our alternative models will be finite versions of those models.

At last, in the third chapter, we will assume that the diffusive mechanism is known, and we will see how to calculate and if possible to minimize the search time, namely the first-passage time to a given target. This problem goes far beyond the simple trading problem of exchange rates: we could look at a transcription factor looking for its anchor site on the DNA [30, 83], or at animals looking for food [15, 218]. Lévy flights [81, 219] and intermittent strategies [23, 27] have been proposed as optimal search processes to find an hidden target, under some assumptions. Intermittent strategies are adapted to a case where the target is hard to find: the random walker alternates fast relocalization, where he moves fast and cannot find the target, and searching phases, where he focus only on finding the target. So far, optimization of search strategies with "easy" target, namely where the target is found as soon as the random walker pass in a neighboring area, have been done with Lévy processes in infinite space, with an infinite number of target. We will transpose this problem into a confined volume.

Search strategies assume that the random walker moves without knowing where the target is. If we want to optimize a purely diffusive process, knowing where the target is before starting the random walk, we can also change the surrounding space in order to increase the probability to find the target. We will first investigate the case of a moving target, already solved formally in 1D [106]. This problem is very cumbersome from the analytical point of view, but meaningful as soon as we consider bimolecular reactions, or a predator hunting a moving prey. We will try to obtain a tractable expression of the first encounter time between the two random walkers.

If the protein motion within a living cell can be considered as a continuous process, in other example, like social networks [234], viruses spreading [155] or electrical networks [53], the random walk occurs on discrete networks. Those networks can either be seen as a discretization of a continuous problem, or directly as a discrete problem. For regular networks such as Euclidian lattices, we can transpose directly the results we have for continuous processes. For "complex" networks [93], that share some common features such as small-world property [13, 225], scale free property [4, 3] or even fractal scalings [205] the topology can be a become a crucial parameter of the search efficiency. The link between search efficiency and network topology have for long been controversial: the presence of loops [12], the scale-free feature [241] or the fractal features [236] have been investigated for its potential impact. General results linking the Global Mean First-Passage Time (GMFPT) and the network topology already exists [6, 66], we will see how these results help to quantify the impact of network topology on search efficiency. We will also focus directly on the Mean First-Passage Time (MFPT), and see how it is affected by the target connectivity or the modification of a part of the network. The general approximation of pseudo-Green functions developed in reference [66] will be used extensively during this last chapter.

1 A guided visit into the random world

Contents

Introduction

The purpose of this guided visit is to introduce some concepts that will be used along this thesis. The reader already familiar with random processes and first-passage observables can save some time by just skipping this visit. The unfamiliar reader will also save some time by reading this part before the two following. We will introduce all the theoretical background of this manuscript, the two first sections being mainly usual textbook knowledge, while in the last one, we introduce some new recent results on first-passage properties.

This visit will be organized in three steps. It will begin with an introduction to random walks and diffusion, with a focus on the mathematical description and on the mathematical tools. Analytical approaches of continuous space random processes will be quickly presented, as well as a link between diffusion on discrete networks and continuous space diffusion. The usual mathematical transform (Fourier, Laplace and Z-transform) will be presented in the last sub-section, with in particular the convention chosen for this manuscript.

The second part of the visit will be a focus on first-passage properties: what is a first-passage process, and what kind of observables can one compute. For this section, we will mainly focus on Brownian motion. We will introduce the renewal equation, that is valid for any kind of Markovian random walk, and see how Green functions are related with first-passage observables. The last sub-section will present results that are not (so far) in textbooks: splitting probabilities, pseudo-Green functions for regular lattices and occupation times. Those results were part of my Master degrees, and will be used again in this manuscript.

The last part of the visit will be an opening on non-Brownian diffusion: anomalous diffusion, strange spaces with non-integer dimension, or even non-Euclidian spaces, processes with memory... The first sub-section will introduce anomalous diffusion, when the mean square displacement $\langle r^2 \rangle$ is no more proportional to the time t. Four processes are introduced, and show how a given diffusion mechanism can lead to anomalous diffusion. The second sub-section describe networks that lead to anomalous diffusion: all networks that will be used in this manuscript are described here. The last sub-section shows how to approximate pseudo-Green functions H for a scale-invariant process (either continuous or discrete). The results presented in this last sub-section will be widely used during this manuscript, and some of them are original research work performed in my Master thesis framework.

1.1 Diffusion and random walks

We will first define the different concepts used in this manuscript. The first subsection will be dedicated to basic notions, like randomness, graphs, discrete and continuous random walk. We will define what random means for us, and what we call a random walk, either in discrete or in continuous space. The second sub-section will introduce a link between discrete space and continuous space diffusion: we present it on Euclidian lattices, and we will use it on complex networks as well as for fractional diffusion. The last sub-section will present the mathematical tools used all across the manuscript: Fourier and Laplace transforms, average definition as well as a link between discrete and continuous random walk. This sub-section content could be found in any (good) textbook related to diffusion, we simply picked up what would be used along this manuscript, and defined the conventions used for the mathematical transforms.

1.1.1 Basic notions

Randomness

Random will in this manuscript mean non deterministic or non predictable: if I throw a die, I know that I will obtain a number between 1 and 6. If I know exactly the initial position, the initial velocity, the force applied, and so on, I could determine in advance the final result (if we except any quantum effect). But in the practice, I just know that for a normal die, I have an equal probability to obtain any integer between 1 and 6.

In such experiment, we only have a probabilistic information on the result: we don't know what will be the result of a given experiment, but we know that if we repeat the same experiment N times, with $N \gg 1$, we will obtain in average $p_i N$ times the result i, with p_i the occurrence probability of result i, with some fluctuations that decrease (relatively) as N increases. The law of the large number ensures that the average will be $p_i N$, and the central limit theorem gives in addition the magnitude of fluctuations.

In this manuscript, "random process" will then be used for a probabilistic process, and characterized by a probability distribution. The resulting event is a "random variable", and can be either discrete or continuous. A discrete random variable may assume either a finite number of values, or and infinite sequence of values, a continuous random variable may assume any value in an interval or collection of intervals.

For a discrete random variable, we define the probability mass $\text{Prob}(X = i)$ that the random variable X takes the value i, such that

$$\sum_i \text{Prob}(X = i) = 1. \tag{1.1}$$

The outcomes of a rolling die is an example. The random variable X can take six different values $\{1, 2, 3, 4, 5, 6\}$, and the probability mass is:

$$\text{Prob}(X = i) = \begin{cases} \frac{1}{6}, & \text{if } i = 1, 2, 3, 4, 5, 6 \\ 0, & \text{otherwise.} \end{cases} \tag{1.2}$$

For a continuous random variable, we will use a probability density function $f(x)$ such that $\text{Prob}(a \leq X \leq b)$, the probability that the random variable X takes a value between a and b,

satisfies

$$\text{Prob}(a \leq X \leq b) = \int_a^b f(x)dx \tag{1.3}$$

$f(x)$ is normalized: for a real random variable, $f(x)$ satisfies

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{1.4}$$

Graph

The main random process of interest in this manuscript are the random walks. Before defining properly what we call here a random walk, we will define the spatial environment in which the random walk occurs.

We will first focus on discrete spaces, that we will model using graphs. A graph is a set of nodes (or points) connected through edges (or links). We can define several quantities using this simple definition:

- "neighbors" nodes are two nodes connected by an edge,
- the "connectivity" is the number of neighbors of a given node,
- a "cycle" is a given number of edges forming a closed chain (the first node is also the last one),
- a "tree" is a graph with no cycle,
- a "directed graph" is a graph where it exists at least two nodes i and j such that the edge $i \to j$ is not symmetric to the edge $j \to i$ (edges are here the probability to jump from i to j: the probability to go from i to j is not equal to the probability to go from j to i),
- a "finite" graph is a graph containing a finite number of nodes,
- an "infinite" graph is in the contrary a graph made with an infinite set of nodes,
- the "chemical distance" between two nodes is the number of edges linking those two nodes following the shortest possible path.

Some simple examples of graph can now be introduced.

A complete graph is a graph where every pair of node is connected. The complete graph of degree n, noted K_n (from the German *komplett*) has n nodes, $n(n-1)/2$ edges. Every node has a connectivity $n-1$. The chemical distance between two nodes is always 1 (all the nodes are connected), and $\forall n > 2$, graphs K_n are finite and contain cycles. Figure 1.1 shows, from left to right, K_3, K_4 and K_5.

A lattice graph is a graph who corresponds to a usual grid, like square grid or triangular grid. For instance a usual $2D$ Euclidian lattice of size X is a square grid of size $X \times X$. For this graph, interior nodes have a connectivity 4, border nodes a connectivity 3, and angle nodes a connectivity 2. Fig. 1.2 shows a 2D Euclidian lattice of size $X = 4$.

At last, the Bethe lattice or Cayley tree is a graph where the connectivity of every node is the same, and without any cycle. This tree is by definition infinite. Figure 1.3 shows such graph for a connectivity $z = 3$.

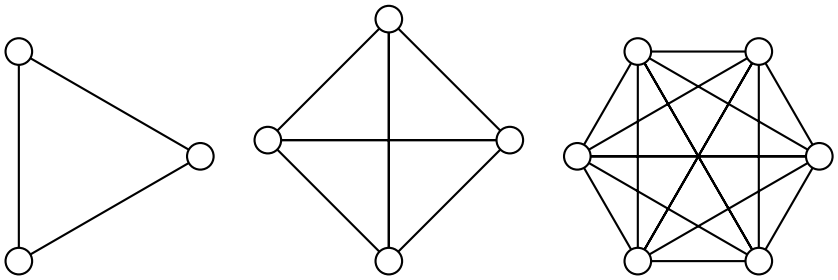

Figure 1.1: Complete graphs K_3, K_4 and K_5.

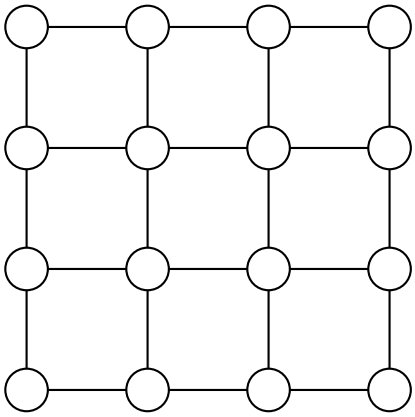

Figure 1.2: 2D Euclidian lattice with $X = 4$.

Physicists use indifferently the words "graph" and "network". Some networks are called "complex", when they have topological properties very different from usual Euclidian lattices. Among the "complex" features, we can introduce:

- small-world networks, where the network diameter D, namely the longer chemical distance between two nodes, grows like $D \propto \ln(N)$, where N is the number of nodes in the network ;
- scale-free networks, where the connectivity distribution follows a power-law (at least asymptotically): the probability that a node have k neighbors is $P(k) \propto k^{-\gamma}$;
- scale-invariant networks, where the same properties are conserved, at least in average, at all scales: the number of nodes $N(r)$ within a shell of size r is $N(r) \propto r^{d_f}$, where d_f is the fractal dimension. Among those networks, we can introduce deterministic fractal networks that are self-similar, namely where the same pattern is observed at all scales.

As example of real complex networks, we can introduce the yeast protein interaction network (PIN) [101], which is scale-free and exhibit a small-world structure [217]. In this network, every node is a protein present in yeast cell. When two proteins interact, namely when two proteins stick together in a solution, a link between the two corresponding node is added. The resulting network is shown in Figure 1.4: the more the node is connected, the bigger is the sphere representing this node. We observe that the majority of nodes have a small connectivity, and that the few proteins that interact with many other structure the network.

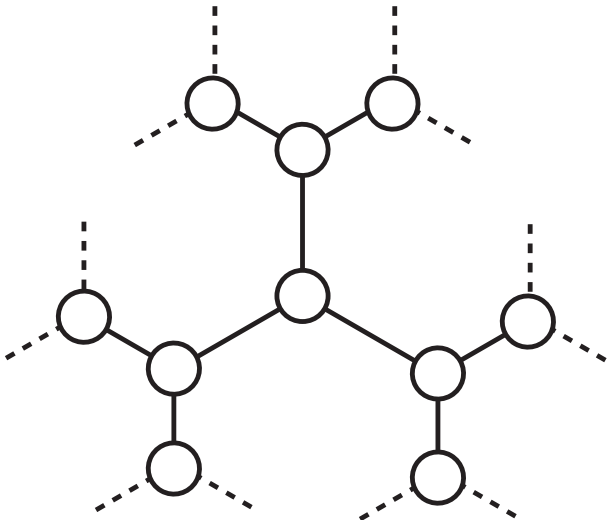

Figure 1.3: Bethe lattice of connectivity $z = 3$.

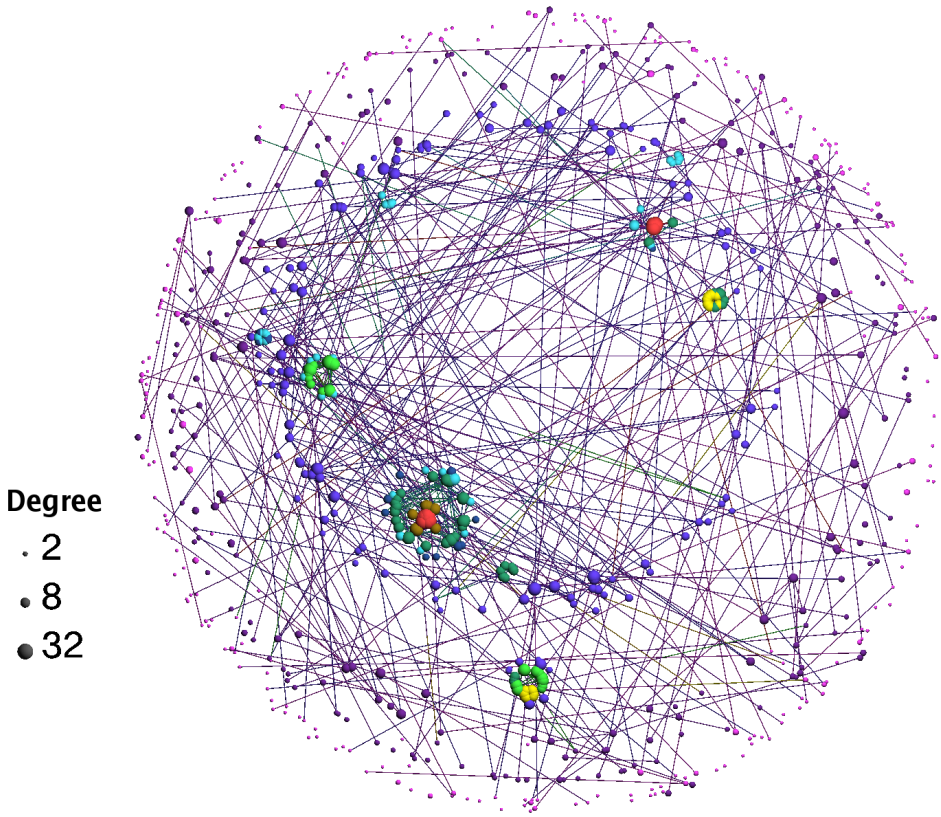

Figure 1.4: The yeast PIN (Picture generated by LaNet-vi software, http://lanet-vi.soic.indiana.edu/), obtained from the filtered yeast interactome.

Very different real complex networks have been studied, like the World Wide Web network [4], social networks and epidemic spreading [155], urban transit system [230] or electric power

transmission grids [53]. A lot of them have been found to be scale-invariant [204].

We will now introduce some theoretical models that illustrate one of the properties defined previously.

A model of scale-free networks is given by the Barabási-Albert model[3]: we start with a small connected network of m_0 nodes, and we add one by one the other nodes. Each new node is connected with m existing nodes, but not with a flat probability. The new node have a higher probability to make a link with an already highly connected node. The probability to make a link with node i is:

$$P_i = \frac{k_i}{\sum_j k_j}, \tag{1.5}$$

where k_i is the connectivity of node i. Such network has a connectivity distribution scaling like $P(k) \propto k^{-3}$. Figure 1.5 shows an example of such network, with $m = 2$, and 64 nodes.

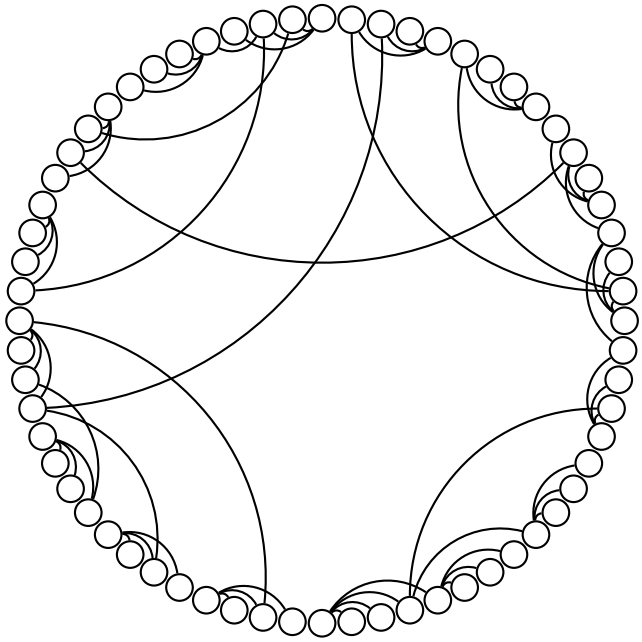

Figure 1.5: Drawing of a Barabási-Albert graph (m=2, 64 nodes).

Self-similar networks will be illustrated with a deterministic fractal network: if scale-invariant networks have the same property *in average* at all scales, self-similar networks like deterministic fractals have *exactly* the same pattern at an infinite number of length scale. We can for instance introduce the Sierpinski gasket, defined as a triangle where at each step, we remove an inner triangle in any full triangle. If we consider the network of the triangle edges, we obtain a fractal network. As shown in Figure 1.6, this network has $3(3^g + 1)/2$ nodes at generation g ($g = 0$ is the original triangle). Every node have a connectivity $k = 4$ except the three initial nodes that have a connectivity $k = 2$.

We can introduce the fractal dimension: within a ball of radius r, we have $N(r)$ nodes. $N(r)$ is, for a fractal, a power-law: $N(r) \propto r^{d_f}$, where d_f is the fractal dimension. We can compute this dimension for the Sierpinski gasket: when we double the ball radius, we encompass a network of

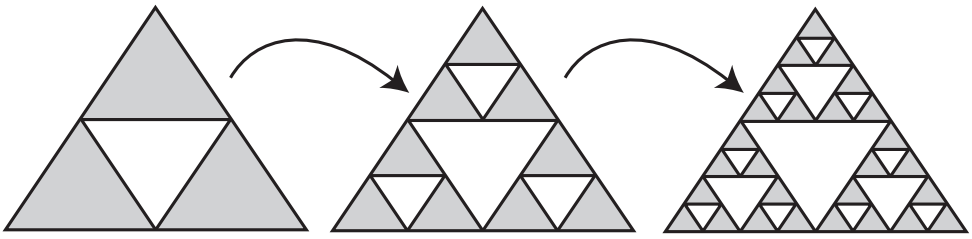

Figure 1.6: Scheme of the Sierpinski gasket first generations.

the next generation. This leads to:

$$\frac{N(2r)}{N(r)} = \frac{3^{g+2}+3}{3^{g+1}+3} \underset{r\to\infty}{\to} 3 \Leftrightarrow 2^{d_f} = 3 \Leftrightarrow d_f = \frac{\ln(3)}{\ln(2)} \simeq 1.58\ldots \tag{1.6}$$

Discrete space random walk

Now that we have defined a random process and a discrete space, we can define what a discrete space random walk is. The random walker perform a succession of jumps: at each step, the random walker chooses a random node between the neighboring nodes, and jump toward the chosen one.

The jump process is a random process and the walker position the resulting random variable: we can define the probability mass for a single step $\omega_{ij} = \omega_{i\to j}$, namely the probability that being at node i, the random walker choose the node j. Normalization gives

$$\sum_j \omega_{ij} = 1 \tag{1.7}$$

We will first assume that the random walker has no memory: every time he is on site i, the probability ω_{ij} to jump toward site j does not depend on the previous steps. This is a property called "Markov property": the probability to be in $\mathbf{r}_{n+1}$ only depends on the position $\mathbf{r}_n$. This property can be written as follow

$$\text{Prob}(\mathbf{r}_{n+1} = \mathbf{x}|\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) = \text{Prob}(\mathbf{r}_{n+1} = \mathbf{x}|\mathbf{r}_n) \tag{1.8}$$

The probability $\text{Prob}(\mathbf{r}_{n+1} = \mathbf{x}|\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ to be at position $\mathbf{x}$ after $n+1$ steps, knowing the n previous positions of the random walker $\mathbf{r}_1, \ldots, \mathbf{r}_n$ is equal to the probability $\text{Prob}(\mathbf{r}_{n+1} = \mathbf{x}|\mathbf{r}_n)$ to be at position $\mathbf{x}$ after $n+1$ steps, knowing only the previous position $\mathbf{r}_n$. The positions $\mathbf{r}_1, \ldots, \mathbf{r}_{n-1}$ do not have any influence on the $n \to n+1$ jump.

We will consider, at least in the beginning of the manuscript, only Markovian random walk. Non-Markovian random walks allows to take in account memory effects, but are far more cumbersome to deal with.

A discrete random walk is then characterized by a succession of random variables, the position of the random walker after n steps, $n \in \mathbb{N}$. For a given graph, we will often consider nearest neighbor random walk: we set $\omega_{ij} = 0$ if i and j are not neighbor, and $\omega_{ij} = 1/k_i$ else, where k_i is the connectivity of node i.

In the case of Markovian processes, to characterize entirely the random walk, we will use the propagator $P(\mathbf{r}, n|\mathbf{r}_S)$, defined as the probability to find the random walker at the position $\mathbf{r}$ after n steps, knowing that the initial position is $\mathbf{r}_S$. This propagator is, for a given n, the simple probability mass of the random walker position. Normalization gives in particular:

$$\forall n \in \mathbb{N},\ \sum_i P(\mathbf{r}_i, n|\mathbf{r}_S) = 1 \tag{1.9}$$

To compute this propagator, we can write a forward equation

$$P(\mathbf{r}, n|\mathbf{r}_S) = \sum_{\mathbf{r}'} \omega_{\mathbf{r}'\mathbf{r}} P(\mathbf{r}', n-1|\mathbf{r}_S) \tag{1.10}$$

This equation links the probability to be at $\mathbf{r}$ at time n and at $\mathbf{r}'$ at time $n-1$: the only way to be in $\mathbf{r}$ at n is to be in a site $\mathbf{r}'$ linked to $\mathbf{r}$ one step before, and to perform a step $\mathbf{r}' \to \mathbf{r}$. The probability to be in $\mathbf{r}'$ after $n-1$ step is $P(\mathbf{r}', n-1|\mathbf{r}_S)$, and the probability to jump from $\mathbf{r}'$ to $\mathbf{r}$ is $\omega_{\mathbf{r}'\mathbf{r}}$. If we sum this over all possible $\mathbf{r}'$, we find $P(\mathbf{r}, n|\mathbf{r}_S)$.

For a finite graph of size N, we can transform equation (1.10) into a matrix equality, using the transition matrix $\Omega = [\omega_{ij}]$ ($N \times N$ matrix)

$$\begin{pmatrix} P(\mathbf{r}_1, n|\mathbf{r}_S) \\ \vdots \\ P(\mathbf{r}_N, n|\mathbf{r}_S) \end{pmatrix} = \Omega. \begin{pmatrix} P(\mathbf{r}_1, n-1|\mathbf{r}_S) \\ \vdots \\ P(\mathbf{r}_N, n-1|\mathbf{r}_S) \end{pmatrix} \tag{1.11}$$

A simple recurrence then gives

$$\begin{pmatrix} P(\mathbf{r}_1, n|\mathbf{r}_S) \\ \vdots \\ P(\mathbf{r}_N, n|\mathbf{r}_S) \end{pmatrix} = \Omega^n. \begin{pmatrix} P(\mathbf{r}_1, 0|\mathbf{r}_S) \\ \vdots \\ P(\mathbf{r}_N, 0|\mathbf{r}_S) \end{pmatrix} \tag{1.12}$$

This equation is called the Chapman-Kolmogorov equation for Markov chains. If we know the initial position $\mathbf{r}_S$, and the transition matrix Ω, we can compute the whole propagator, at least formally.

We have defined discrete random walks, we will now show how to extend this defintion to obtain continuous random walks (continuous in space and in time).

Continuous random walks

Several ways exist to define a continuous random walk. A simple manner is to use the problem described in 1905 by Karl Pearson, in a study of mosquito's population migration in a forest. Pearson asked [156]:

> *"A man starts from a point O and walks l yards in a straight line; he then turns through any angle whatever and walks another l yards in a second straight line. He repeats this process n times. I require the probability, that after these n stretches he is at distance between r and $r + dr$ from his starting point."*

Lord Rayleigh gave the answer a week after [166]. If n is great enough, the probability is

$$dP(r) \sim \frac{2}{nl^2} e^{-\frac{r^2}{nl^2}} r dr \tag{1.13}$$

This random walk takes place in a $2D$ Euclidian space. At each step, the mosquito chooses randomly a direction, and then goes straight for a given distance l. The answer given by Lord Rayleigh is the probability to be in position r after n steps. Knowing the time needed to perform a step, we could extract a propagator, namely the probability to be at a distance r at time t.

Pearson walks are thus not really easy to deal with, since they are not Markovian: the joint process (position, speed) is Markovian, but not the process (position) alone. We indeed cannot predict the position at $t + dt$ knowing only the position at t: during a step, we also need the speed to predict the future position.

Another way to define a continuous Markov process is to take the continuous limit of an infinite discrete lattice, with a classical rule $\omega_{ij} = 0$ is i and j are neighbors, and $\omega_{ij} = 1/k_i$ else.

We can try to obtain as previously an equation linking $p(\mathbf{x}, t + \tau|\mathbf{x}_S)$ and $p(\mathbf{x}, t|\mathbf{x}_S)$, where $p(\mathbf{x}, t|\mathbf{x}_S)$ is the probability to be in position $\mathbf{x}$ at time t starting from $\mathbf{x}_S$ at time $t = 0$. If we consider a random walk on a 2D infinite Euclidian lattice, where the distance between two sites is ϵ, and the time between two steps is τ, we can write a forward equation like equation (1.10)

$$p(\mathbf{x}, t + \tau|\mathbf{x}_S) = \frac{1}{4}\left(p(\mathbf{x} - \epsilon\mathbf{e}_x, t|\mathbf{x}_S) + p(\mathbf{x} + \epsilon\mathbf{e}_x, t|\mathbf{x}_S) + p(\mathbf{x} - \epsilon\mathbf{e}_y, t|\mathbf{x}_S) + p(\mathbf{x} + \epsilon\mathbf{e}_y, t|\mathbf{x}_S)\right) \tag{1.14}$$

This equation can be understood with Figure 1.7 as follows: to be in position $\mathbf{x}$ at time $t+\tau$, the random walker has to be at time t in one of the four neighboring sites $\mathbf{x} - \epsilon\mathbf{e}_x$, $\mathbf{x} + \epsilon\mathbf{e}_x$, $\mathbf{x} - \epsilon\mathbf{e}_y$ or $\mathbf{x} + \epsilon\mathbf{e}_y$, and to perform a jump toward $\mathbf{x}$. The probability to do the right jump is $1/k_i = 1/4$ for a 2D Euclidian lattice.

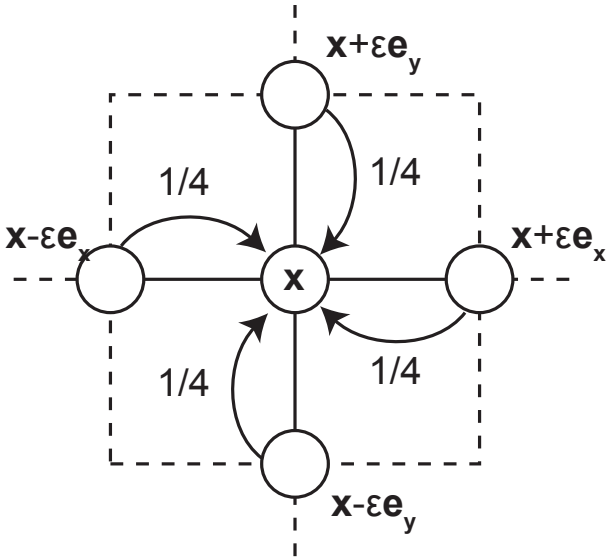

Figure 1.7: Scheme of the forward Kolmogorov equation.

Here, τ and ϵ are real numbers. We have in the limit $\tau \to 0$ and $\epsilon \to 0$, using Tailor's theorem:

$$p(\mathbf{x}, t|\mathbf{x}_S) + \tau\frac{\partial p}{\partial t}(\mathbf{x}, t|\mathbf{x}_S) = p(\mathbf{x}, t|\mathbf{x}_S) + \frac{\epsilon^2}{4}\left(\frac{\partial^2 p}{\partial x^2}(\mathbf{x}, t|\mathbf{x}_S) + \frac{\partial^2 p}{\partial y^2}(\mathbf{x}, t|\mathbf{x}_S)\right) \tag{1.15}$$

If we assume that $\tau \sim \epsilon^2$, we can write:

$$\frac{\partial p}{\partial t}(\mathbf{x}, t|\mathbf{x}_S) = \frac{\epsilon^2}{4\tau}\left(\frac{\partial^2 p}{\partial x^2}(\mathbf{x}, t|\mathbf{x}_S) + \frac{\partial^2 p}{\partial y^2}(\mathbf{x}, t|\mathbf{x}_S)\right) \tag{1.16}$$

If we generalize this result for a dimension d, we obtain

$$\frac{\partial p}{\partial t}(\mathbf{x}, t|\mathbf{x}_S) = \frac{D}{2d}\sum_{i=1}^{d}\frac{\partial^2 p}{\partial x_i^2}(\mathbf{x}, t|\mathbf{x}_S) \tag{1.17}$$

where $D = \epsilon^2/\tau$ is the diffusion coefficient. We have here a forward Fokker-Planck equation to compute the propagator, knowing the initial conditions.

As for discrete random walk, the random walker position at time t is a continuous random variable whose probability density function is the propagator. Normalization gives

$$\forall t \in \mathbb{R}^+, \int_{-\infty}^{\infty} p(\mathbf{x}, t|\mathbf{x}_S)d\mathbf{x} = 1 \tag{1.18}$$

We will now see how we can link this miscroscopical description of random walk with some macroscopic observables of diffusion.

1.1.2 From random walk to macroscopic diffusion

So far, we looked at a single random walker evolving on a graph or on an Euclidian space: a pertinent observable of this evolution is the propagator. If instead of looking at a single random walker we focus on a very large number of independent random walkers, this propagator becomes the particle concentration, a macroscopic observable.

The macroscopic behavior of a large number of identical and independent random walkers will be called "diffusion".

To describe diffusion processes, one often use a distribution function $u(\mathbf{r}, t)$ giving the value of an observable at position $\mathbf{r}$ and at time t, satisfying the so-called heat equation:

$$\frac{\partial u}{\partial t} - \alpha\Delta u = 0 \tag{1.19}$$

where α is a positive constant and Δ the Laplace operator or Laplacian. This operator is defined in Cartesian coordinates as:

$$\Delta u = \sum_{i=1}^{d}\frac{\partial^2 u}{\partial x_i^2} \tag{1.20}$$

We note that this equation is exactly the one obtained for a single walker with Kolmogorov forward equation (1.17), with $\alpha = D/(2d)$.

The steady-state, or stationary state, can be found solving:

$$\alpha\Delta u = 0 \tag{1.21}$$

This steady-state will prove to be useful to compute first-passage observables.

Adolf Fick derived in the year 1855[88] the fundamental laws of diffusion. If we note $c(\mathbf{r}, t)$ the particle concentration at position $\mathbf{r}$ and at time t, and $D \geq 0$ the diffusion coefficient, we have:

$$\frac{\partial c}{\partial t} = \frac{D}{2d}\Delta c \tag{1.22}$$

where Δ is the Laplace operator. We retrieve a classical heat equation, where $u = c$ and $\alpha = D/(2d)$.

The link between single particles diffusion, observed by Robert Brown [42] or Jean Perrin [157], and the heat equation has been made in 1905 by Albert Einstein[82] . He showed that the diffusion coefficient can be deduced from the mean square displacement $\langle r^2 \rangle$ of a single particle:

$$D = \frac{\langle r^2 \rangle}{t} \tag{1.23}$$

Here, D is a macroscopic observable, and $\langle r^2 \rangle$ a microscopic one.

1.1.3 Mathematical tools

The solution of the heat equation (1.19) has been found in 1822 by Joseph Fourier[89]. A generalization of his approach involves two kind of mathematical transforms: Fourier transform and Laplace transform.

Fourier transform

Fourier transform decomposes a time-dependent function into its constituent frequencies:

$$\mathcal{F}(f)(\omega) = \hat{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{-\imath 2\pi x\omega} dx \tag{1.24}$$

Knowing the Fourier transform, we can retrieve the initial function using the inverse Fourier transform:

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\omega) e^{\imath 2\pi\omega x} d\omega \tag{1.25}$$

A differential equation becomes a simple polynomial equation in Fourier space:

$$\mathcal{F}\left(\frac{d^n f}{dx^n}\right)(\omega) = (\imath 2\pi\omega)^n \hat{f}(\omega) \tag{1.26}$$

Characteristic functions

A classical application of Fourier transform is the characteristic function. If we have a continuous random variable X, the Fourier transform of the probability density function f_X (with a slightly different convention) is called the characteristic function ϕ_X:

$$\phi_X(k) = \int_{\infty}^{\infty} e^{ikx} f_X(x) dx \tag{1.27}$$

Characteristic function defines completely the probability density function.

Fourier transform of a convolution is the multiplication of Fourier transforms:

$$\mathcal{F}\left(\int_0^x f(x')g(x-x')dx'\right) = \widehat{(f \star g)} = \hat{f}.\hat{g} \tag{1.28}$$

Using this property, we can obtain easily the characteristic function of a sum of n identically distributed random variables:

$$\phi_{X_1+X_2+\ldots+X_n}(k) = \phi_{X_1}(k)\phi_{X_2}(k)\ldots\phi_{X_n}(k) = (\phi_X(k))^n \tag{1.29}$$

This property will be widely used when considering discrete random walks, when each jump is a random variable whose probability density function is known.

Fourier series

Fourier transform is adapted to a continuous variable defined on $\mathbb{R}$. If we have boundary conditions and a variable confined on a given segment of size T, we can extend the function f on $\mathbb{R}$ with a T-periodic function. The Fourier transform can here be reduced to Fourier series. If we use exponential Fourier series, we have:

$$\forall n \in \mathbb{Z},\ \hat{f}(n) = \frac{1}{T}\int_{-T/2}^{T/2} f(x)e^{-\imath 2\pi\frac{n}{T}x}dx \tag{1.30}$$

As previously, we can retrieve the initial function through resummation:

$$f(x) = \sum_{n=-\infty}^{\infty} f(n)e^{\imath 2\pi\frac{n}{T}x} \tag{1.31}$$

Fourier series still transform a derivative equation into a polynomial one:

$$\widehat{\frac{d^k f}{dt^k}}(n) = \widehat{f^{(k)}}(n) = \left(\frac{2\imath\pi n}{T}\right)^k \widehat{f}(n) \tag{1.32}$$

Discrete Fourier transform

At last, for a variable defined on a finite discrete set, typically on a graph, we can adapt the Fourier transform using discrete Fourier transform. Let consider a sequence $x_0, \ldots, x_{N-1}$. The discrete Fourier transform is:

$$\forall k \in [0, N-1],\ X_k = \sum_{n=0}^{N-1} x_n e^{-\imath 2\pi\frac{k}{N}n} \tag{1.33}$$

This formula is similar to the one for Fourier series, except that the integral is replaced by a finite sum. The inverse discrete Fourier transform is given by:

$$\forall n \in [0, N-1],\ x_n = \frac{1}{N}\sum_{n=0}^{N-1} X_k e^{\imath 2\pi\frac{k}{N}n} \tag{1.34}$$

Derivative have no meaning for discrete sequences, but this discrete Fourier transform will still be a way to simplify an analog of heat equation for random walks on graphs.

Laplace transform

Laplace transform is an analog of Fourier transform, where instead of transforming a function into its frequencies, Laplace transform resolves it into its moments:

$$\mathcal{L}(f)(s) = \widetilde{f}(s) = \int_0^\infty e^{-st} f(t) dt \tag{1.35}$$

The inverse Laplace transform is given by a complex integral, usually cumbersome to compute:

$$f(t) = \frac{1}{2\pi\imath} \lim_{T\to\infty} \int_{\gamma-\imath T}^{\gamma+\imath T} e^{st} \widetilde{f}(s) ds \tag{1.36}$$

Laplace transform simplify once again differentiation:

$$\mathcal{L}\left(\frac{d^n f}{dt^n}\right)(s) = \widetilde{f^{(n)}}(s) = s^n \widetilde{f}(s) - s^{n-1} f(0) - \ldots - f^{(n-1)}(0) \tag{1.37}$$

This formula now involves the values of the real function f and its derivative in $t = 0$.

We also note that Laplace transform of convolution is simply the product of Laplace transforms:

$$\mathcal{L}\left(\int_0^x f(x')g(x-x')dx'\right) = \widetilde{(f \star g)} = \widetilde{f}.\widetilde{g} \tag{1.38}$$

Z-transform

For a discrete sequence, the discrete Laplace transform is called Z-transform:

$$\widetilde{x}(z) = \sum_{n=0}^{\infty} x[n] z^{-n} \tag{1.39}$$

z is here a complex number, chosen in the region of convergence of the Z-transform.

The inverse Z-transform is given by:

$$x[n] = \frac{1}{2\pi\imath} \oint_C \widetilde{x}(z) z^{n-1} dz, \tag{1.40}$$

where C is a counterclockwise closed path encircling the origin, and chosen so that the integral exists. This contour C has to encircle all the poles of $\widetilde{x}(z)$.

Average

When using probabilities, one often performs an "average" of a given quantity. In the framework of random walks, we will use two kind of averages: time average and ensemble average. A time average, noted $\langle \ldots \rangle_t$, is defined for a single random walk as the average along a trajectory. It is defined as:

$$\langle A \rangle_t = \int_0^\infty A(t) dt \tag{1.41}$$

For instance the time-averaged mean square displacement is:

$$\langle r^2 \rangle_t(\tau) = \int_0^\infty \|\mathbf{r}(t+\tau) - \mathbf{r}(t)\|^2 dt, \tag{1.42}$$

where $\mathbf{r}(t)$ is the position of a given random walker at time t.

An ensemble average, noted $\langle \ldots \rangle$, is defined for a large set of random walks. The averaged quantity is weighted by the probability density function f of the corresponding random variable:

$$\langle A \rangle = \int A df \tag{1.43}$$

For instance the mean first-passage time (MFPT) is weighted by the first-passage density FPT(t):

$$\langle \mathbf{T} \rangle = \int_0^\infty t d\mathrm{FPT} = \int_0^\infty t \mathrm{FPT}(t) dt \tag{1.44}$$

General heat equation solution

We will here show how Fourier and Laplace transforms can be used to solve a general heat equation, for instance in three dimensions:

$$\frac{\partial f}{\partial t} = \frac{D}{6}\Delta f = \frac{D}{6}\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}\right) \tag{1.45}$$

If we Fourier transform (on each direction) and Laplace transform this equation, we obtain:

$$s\widehat{\widetilde{f}} - \widehat{f}(t=0) = \frac{D}{6}\left((\imath 2\pi\omega_x)^2 + (\imath 2\pi\omega_y)^2 + (\imath 2\pi\omega_z)^2\right)\widehat{\widetilde{f}} \tag{1.46}$$

We assume that the initial condition is that all particles are in $\mathbf{r} = (0,0,0)$ at time $t = 0$. $f(t=0)$ is thus a Dirac, and $\widehat{f}(t=0) = 1$. This leads to:

$$\widehat{\widetilde{f}} = \frac{1}{s + \frac{D}{6}\left((2\pi\omega_x)^2 + (2\pi\omega_y)^2 + (2\pi\omega_z)^2\right)} \tag{1.47}$$

An inverse Laplace transform gives:

$$\tilde{f} = e^{-\frac{D}{6}\left((2\pi\omega_x)^2 + (2\pi\omega_y)^2 + (2\pi\omega_z)^2\right)t} \tag{1.48}$$

At last an inverse Fourier transform leads to the following result:

$$f(x,y,z,t) = \frac{1}{(2\pi\frac{D}{3}t)^{3/2}} e^{-\frac{x^2+y^2+z^2}{\frac{2D}{3}t}} \tag{1.49}$$

This solution works for a particle starting in $(0,0,0)$ at time $t = 0$, and evolving in a $3D$ media without boundary. We easily retrieve Einstein's formula:

$$\begin{aligned}
\langle r^2(t) \rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x^2+y^2+z^2) f(x,y,z,t) dx dy dz && (1.50)\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x^2+y^2+z^2) \frac{1}{(2\pi\frac{D}{3}t)^{3/2}} e^{-\frac{x^2+y^2+z^2}{\frac{2D}{3}t}} dx dy dz && (1.51)\\
&= \int_0^\infty \frac{r^2}{(2\pi\frac{D}{3}t)^{3/2}} e^{-\frac{r^2}{\frac{2D}{3}t}} 4\pi r^2 dr && (1.52)\\
&= \frac{8Dt}{3\sqrt{\pi}} \int_0^\infty u^4 e^{-u^2} du && (1.53)\\
&= Dt && (1.54)
\end{aligned}$$

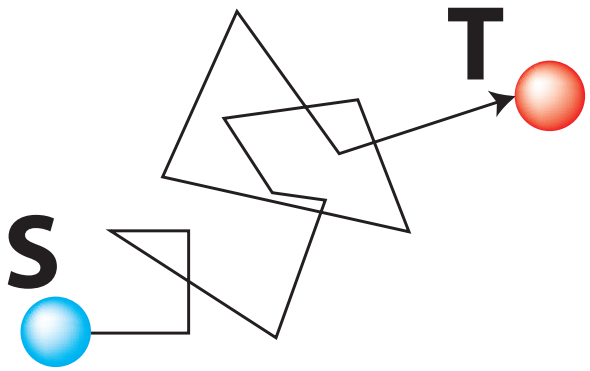

Figure 1.8: Scheme of a generic first-passage problem.

QUICK SUMMARY

- We have defined what diffusion and random walks are ;
- we have introduced networks and some mathematical tools ;
- we have shown a link between random walks and diffusion.

1.2 First-passage properties

We will now focus on a more specific subject concerning diffusion and random walks: first-passage properties. The first sub-section will define the first-passage observables that we will use all along this manuscript. The second sub-section will introduce the renewal equation, only valid for Markovian processes, and shows how to relate the mean first-passage time (MFPT) with the pseudo-Green functions. This link is central for this manuscript: it allows a quantitative estimation of MFPT as soon as the pseudo-Green function can be calculated or approximated. The last sub-section extend this relationship to other first-passage time observables. The first two sub-sections presents textbook results while the last one introduce some original results that have been obtained during my Master thesis.

1.2.1 The basic of first-passage

We will consider a random walk starting from $\mathbf{r}_S$. We name first-passage time the first time the random walker hits a target site $\mathbf{r}_T$. This time is called the first-passage time (FPT). As shown in Figure 1.8, this quantity is defined for a single random walker. This quantity is fundamental in the study of transport limited reactions [169, 185, 133], as it gives the reaction time in the limit of perfect reaction.

This quantity is also useful in target search problems [198, 69, 22, 24, 84, 121], and other physical systems [66, 192, 168], and will be used throughout this manuscript.

We can extract several quantities with this first definition. The more complete quantity is the first-passage density $\mathrm{FPT}(t)$ giving for a single particle the probability that the first-passage time is t. This density is the probability density function of the first-passage time, satisfying in

particular:

$$\int_0^\infty \mathrm{FPT}(t)dt = 1 \tag{1.55}$$

We can also consider the moments of this random variable, namely:

$$\langle \mathbf{T}^n \rangle = \int_0^\infty t^n \mathrm{FPT}(t)dt \tag{1.56}$$

The first moment ($n = 1$) is called the mean first-passage time (MFPT), and is for some cases a way to characterize the whole first-passage density.

Those moments are not always defined: the integral of equation (1.56) does not always converge. If for instance $\mathrm{FPT} \propto t^{-1-\alpha}$ when $t \to \infty$, with $\alpha \in]0, 1[$, the first-passage density can be correctly normalized, but all integer moments are infinite.

For a discrete time random walk, this first-passage density becomes the probability mass of a discrete random variable, the first-passage time. Normalization becomes:

$$\sum_{t=0}^{\infty} \mathrm{FPT}(t) = 1 \tag{1.57}$$

The moments are also defined using infinite sums:

$$\langle \mathbf{T}^n \rangle = \sum_{0}^{\infty} t^n \mathrm{FPT}(t) \tag{1.58}$$

If the starting point and the target point are the same, the first-passage time is called first-return time.

Discrete and continuous first-passage

So far, we talked about "sites" as a region of space that the random walker could hit. This notion is not exactly the same for a random walk on a graph, and for a random walk in continuous space.

Figure 1.9 shows a random walk on an usual $2D$ Euclidian lattice. A site is in this case a network node, and one can define easily the first-passage as the first time the random walker hits the target node.

For a random walk on a continuous space, the target site is defined as a subpart of the embedding space. If the target site is a point, in dimension $d \geq 2$, the probability that a random walker hits the target is smaller than 1: the first-passage time is in this case not normalized. In order to be almost sure to hit the target eventually, one has to define an extended (or non point-like) target. We will often use the approximation that the target extension is very small, but one has to remember that for a continuous space random walk, the target is never, except for unidimensional systems, a point.

Recurrence and compacity

We can here introduce two related notions: compact and non-compact exploration, and transient or recurrent sites.

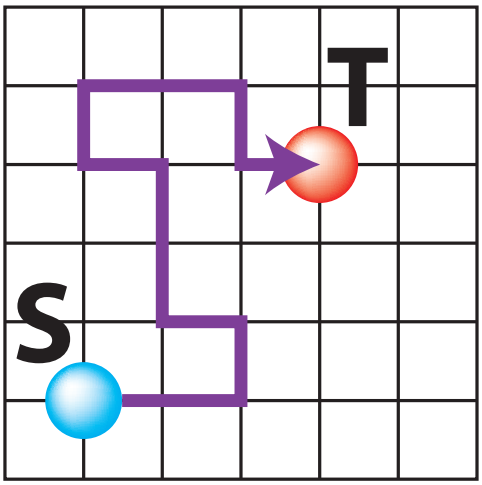

Figure 1.9: Scheme of a discrete first-passage problem.

To define a transient or a recurrent site, we have to compute the probability that, starting from a given point, a random walker in an infinite media returns eventually to this starting point. If this return probability is 1, then almost surely, the random walker will come back to the starting point. Such point is called "recurrent". If not, the point is called "transient". Pólya theorem [161] states that the $1D$ and $2D$ Euclidian spaces are recurrent (everywhere), and the Euclidian spaces of dimension $d \geq 3$ are transient (everywhere). One has to take care that some networks can contain both transient and recurrent sites, like the one described in reference [2].

Since this property is site-dependent, we can define a closely related notion that is media-dependent: the compacity of the exploration[72]. Basically, if all sites are recurrent, then the exploration will be called "compact", and if all sites are transient, the exploration will be called "non-compact". For homogeneous networks, namely where all sites have the same recurrence property (either recurrent or transient), compact is a synonym of recurrent, and non-compact of transient.

We can understand this notion as follows: for a compact exploration, once a site is visited, the near vicinity will almost certainly be visited. For a non-compact exploration, one site can be visited while the neighboring one has a probability strictly smaller than 1 to be visited.

We can propose a link with a more usual definition of compacity by looking at the trace of a random walker in continuous space. This trace, namely the ensemble of all the sites visited between $t = 0$ and t has a fractal dimension d_t equal to the embedding space dimension for a compact exploration, and a d_t smaller that the embedding space dimension for a non-compact one. Namely, a compact exploration visit all the sites in a given area, while a non-compact one will visit only some of them before going to another area. For a regular Brownian motion, the fractal dimension of the trace is $d_t = 1$ for $d = 1$, and $d_t = 2$ for $d \geq 2$: a Brownian motion in $d > 2$ is non-compact since the fractal dimension of the trace will be smaller than the space dimension, leading to a non-compact exploration.

We will see below how to link the compacity property with the fractal dimension of the embedding space d_f and the walk dimension d_w of a scale-invariant random walk.

Several targets and first-passage

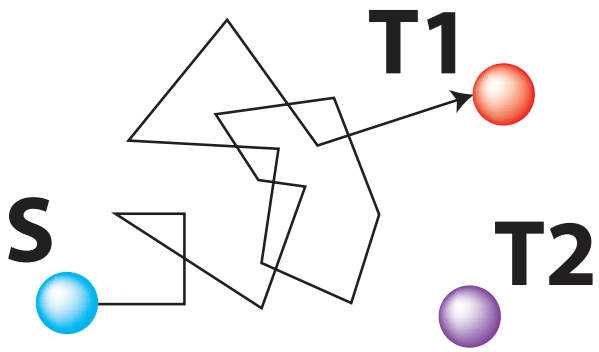

Figure 1.10: Scheme of a two target first-passage problem.

So far, we considered the case of a random walk with a single target. If several targets exist, new first-passage quantities appear. As shown in Figure 1.10, if more than one target is available, one can compute the "first-passage splitting probabilities", namely the probability that the random walker, starting from $\mathbf{r}_S$, hits the target $\mathbf{r}_i$ before any of the other targets $\mathbf{r}_{j\neq i}$. This splitting probability will be noted P_i, it correspond to the probability mass of the discrete random variable "index of the first target hitted". Normalization once again gives:

$$\sum_i P_i = 1 \tag{1.59}$$

Those probabilities allow to study quantitatively competitive reactions[169].

Another quantity is the conditional first-passage density, namely the first-passage density at target i, without touching any of the other targets $j \neq i$. This density, noted $\mathrm{FPT}_i(t)$ is the probability density function of the first-passage time at target i, knowing that the target i is the first target hitted. It is normalized

$$\int_0^\infty \mathrm{FPT}_i(t)dt = 1 \tag{1.60}$$

As previously, we can define the moments of conditional first-passage time:

$$\langle \mathbf{T}_i^n \rangle = \int_0^\infty t^n \mathrm{FPT}_i(t)dt \tag{1.61}$$

The last quantity we will focus on is the first exit time, namely the time needed to reach for the first time any of the targets. The related density $\mathrm{FET}(t)$ is normalized, and satisfy:

$$\mathrm{FET}(t) = \sum_i P_i \mathrm{FPT}_i(t) \Rightarrow \int_0^\infty \mathrm{FET}(t)dt = \sum_i P_i = 1, \tag{1.62}$$

where the P_i are the splitting probabilities. If the targets surround entirely a given volume, like in Figure 1.11, we retrieve the alternative notion of first exit time given below, namely the time needed to exit the volume for the first time.

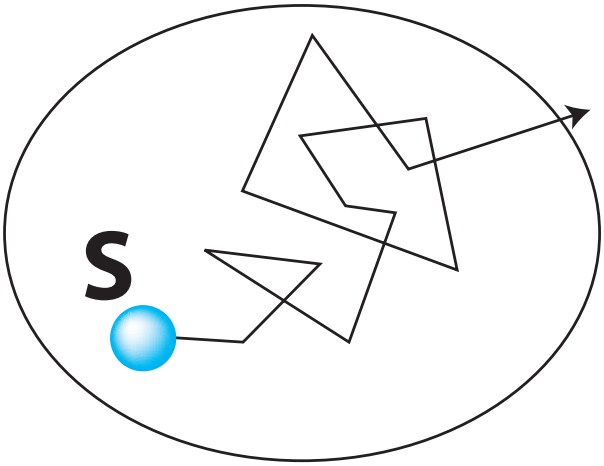

Figure 1.11: Scheme of a first exit problem.

Other first-passage quantities

We will now introduce some supplementary first-passage quantities, to offer a wide overview of what a first-passage observables could be.

We always talk about first-passage, but for random walk performing jumps in a continuous space, two quantities can be defined: the first-passage time and the first arrival time[123]. The first-passage time is the first time the random walker trajectory crosses the target, but this could occur during a jump, where the random walker does not hit the target. The first arrival time is in the contrary the first time the random walker hits the target. By definition, the first arrival time is always larger (or equal) than the first-passage time.

The first exit time is defined when the random walker starts in a given subset of space. As shown in Figure 1.11, the first time the random walker exits this subset is the first exit time. It is equivalent to a first-passage time if we define the subset boundary as the target.

The maximal excursion is a related concept: $r_{\max}(t)$ is the maximal distance from the origin the random walker have reached at time t. If we note $\mathbf{r}(t)$ the random walker position at time t, we can define this maximal excursion as:

$$r_{\max}(t) = \max_{0 \leq t' \leq t} \|\mathbf{r}(t') - \mathbf{r}(t=0)\| \tag{1.63}$$

This maximal excursion is a growing function of time, and is closely related to the first exit time. If the first exit time of a sphere centered on $\mathbf{r}(t=0)$ and of radius r_0 is t, then $r_{\max}(t) = r_0$. The first exit time focus on the time needed to exit a given subset while the maximal excursion focus on the maximal distance reached at a given time.

At last, we can introduce the occupation time, the number of time $\mathbf{N}_i$ a random walker visit a given site $\mathbf{r}_i$ before reaching the target site $\mathbf{r}_T$. This quantity is useful in the context of reactions occurring with a finite probability per unit of time[21, 35, 65]. We stress that the occupation time provides a finer information on the trajectory of the particle than the first passage time. We can retrieve this first-passage time $\mathbf{T}$ with the sum over all sites of the occupation time:

$$\mathbf{T} = \sum_i \mathbf{N}_i \tag{1.64}$$

We have introduced the first-passage observables, we will now see how to compute them analytically.

1.2.2 Renewal equation and Green functions

To compute the first-passage observables, we will introduce the Green functions, that will be extensively used in this thesis. Instead of just defining them, we will see how they appear "naturally" when computing the mean first-passage time.

To do so, we will consider a Markovian random walker. As previously, we note $P(\mathbf{r}_T, t|\mathbf{r}_S)$ the propagator, namely the probability that the random walker is in position $\mathbf{r}_T$ at time t starting from $\mathbf{r}_S$ at time $t = 0$, and $\mathrm{FPT}(t)$ the first-passage density from $\mathbf{r}_S$ to $\mathbf{r}_T$. The renewal equation links those two quantities:

$$P(\mathbf{r}_T, t|\mathbf{r}_S) = \delta_{t,0}\delta_{\mathbf{r}_T,\mathbf{r}_S} + \int_0^t \mathrm{FPT}(t')P(\mathbf{r}_T, t-t'|\mathbf{r}_T)dt' \tag{1.65}$$

This equation means that the probability to be at $\mathbf{r}_T$ at time t is equal to the probability to hit $\mathbf{r}_T$ for the first time at any time $t' \in [0, t]$, and to come back in the same position $\mathbf{r}_T$ at time $t - t'$. The first term of the right hand side is the correction if $t = 0$ and $\mathbf{r}_T = \mathbf{r}_S$. δ is here the Kronecker symbol: $\delta_{x,y} = 1$ if $x = y$, and 0 else. Since we have here a simple convolution, we can obtain a simple equation after a Laplace transform:

$$\widetilde{P}(\mathbf{r}_T, s|\mathbf{r}_S) = \delta_{\mathbf{r}_T,\mathbf{r}_S} + \widetilde{\mathrm{FPT}}(s).\widetilde{P}(\mathbf{r}_T, s|\mathbf{r}_T) \tag{1.66}$$

We can develop those Laplace transforms into s series, assuming that all moments exist:

$$\widetilde{\mathrm{FPT}}(s) = \int_0^\infty e^{-st}\mathrm{FPT}(t)dt = \sum_{n=0}^{\infty}(-1)^n \left(\int_0^\infty t^n \mathrm{FPT}(t)dt\right) s^n = \sum_{n=0}^{\infty}(-1)^n \langle \mathbf{T}^n \rangle s^n \tag{1.67}$$

where $\langle \mathbf{T}^n \rangle$ is the nth moment of the first-passage time.

We cannot use directly the same formalism for the propagator: in a confined environment, when $t \to \infty$, the propagator does not converge to 0 but to the stationary probability P_{stat}:

$$\begin{aligned}
\widetilde{P}(\mathbf{r}_T, s|\mathbf{r}_S) &= \int_0^\infty e^{-st} P(\mathbf{r}_T, t|\mathbf{r}_S)dt \\
&= \int_0^\infty e^{-st}\left(P(\mathbf{r}_T, t|\mathbf{r}_S) - P_{\mathrm{stat}}(\mathbf{r}_T)\right)dt + \frac{P_{\mathrm{stat}}(\mathbf{r}_T)}{s} \\
&= \sum_{n=0}^{\infty}(-1)^n \left(\int_0^\infty t^n \left(P(\mathbf{r}_T, t|\mathbf{r}_S) - P_{\mathrm{stat}}(\mathbf{r}_T)\right)dt\right) s^n + \frac{P_{\mathrm{stat}}(\mathbf{r}_T)}{s}
\end{aligned} \tag{1.68}$$

At last, we can obtain the mean first-passage time, by identification of the s term in equation (1.66):

$$\langle \mathbf{T} \rangle = \frac{\displaystyle\int_0^\infty \left(P(\mathbf{r}_T, t|\mathbf{r}_T) - P_{\mathrm{stat}}(\mathbf{r}_T)\right)dt - \int_0^\infty \left(P(\mathbf{r}_T, t|\mathbf{r}_S) - P_{\mathrm{stat}}(\mathbf{r}_T)\right)dt}{P_{\mathrm{stat}}(\mathbf{r}_T)} \tag{1.69}$$

We will define the pseudo-Green function as:

$$H(\mathbf{r}_T|\mathbf{r}_S) = \int_0^\infty \left(P(\mathbf{r}_T, t|\mathbf{r}_S) - P_{\mathrm{stat}}(\mathbf{r}_T)\right)dt \tag{1.70}$$

Using this notation, we can write the MFPT in a most compact way:

$$\langle \mathbf{T} \rangle = \frac{H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_S)}{P_{\text{stat}}(\mathbf{r}_T)} \tag{1.71}$$

This expression was found first by Aldous [6] (chapter 2, Lemma 12), then by Noh & Rieger [148] and later by [91]. We found during my Master thesis [66] that this expression was not only a formal notation: it has a huge potential since we can compute those pseudo-Green functions, either exactly or with very good approximations, for a very wide range of random walks.

1.2.3 Results already known

We will present two papers that we will intensively use in this manuscript. The first one [65] links splitting probabilities and related first-passage observables with pseudo-Green function, and gives exact expressions as well as some approximations for pseudo-Green functions. The second one [67] was part of my Master thesis, and gives an extension for occupation times and related observables.

Two targets problem

The first paper [65] introduces a way to link splitting probabilities and related quantities with pseudo-Green functions. So far, we know that MFPT can be expressed, for a Markovian random walker evolving on a regular lattice, as:

$$\langle \mathbf{T} \rangle = \frac{H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_S)}{P_{\text{stat}}(\mathbf{r}_T)} \tag{1.72}$$

When there is more than one target available, we can extend this result, and compute the splitting probabilities. To do so, we will consider a network where two targets $\mathbf{r}_{T_1}$ and $\mathbf{r}_{T_2}$ exist, and use the so-called "electrical analogy" developed in [79]. The relationship between the flux, the potential and the mean first-passage time have been investigated in [57].

We will consider that a constant incoming flux of particles J comes from $\mathbf{r}_S$, and that an outcoming flux of particles J_1 exits in T_1, and J_2 in T_2. We suppose that we have reached the equilibrium state, such that J, J_1 and J_2 are constant. In particular, since all particles are eventually absorbed, either by T_1 or T_2, we have $J = J_1 + J_2$. In this model, the splitting probability, namely the probability to reach target i before touching target $j \neq i$ is $P_i = J_i/J$. The total number of particles $\mathcal{N}$ present in the domain is $\mathcal{N} = J\langle \mathbf{T} \rangle$, where $\langle \mathbf{T} \rangle$ is the mean first-passage time by any of the target. Since we have reached equilibrium, we can write that the average number of particle in $\mathbf{r}_i$, $\rho(\mathbf{r}_i)$ satisfies:

$$\rho(\mathbf{r}_i) = \sum_j \omega_{ij}\rho(\mathbf{r}_j) + J\delta_{iS} - J_1\delta_{iT_1} - J_2\delta_{iT_2}. \tag{1.73}$$

This equation means that to be in $\mathbf{r}_i$, the particle must come from a neighbor j, except in S where we have to add the incoming flux J, and in T_i, where we have to remove the outcoming flux J_i. Here, δ is the Kronecker delta function: $\delta_{xy} = 1$ if $x = y$, 0 else.

To solve this problem, we use the pseudo-Green function H [14, 65], which satisfies:

$$H(\mathbf{r}_i|\mathbf{r}_j) = \sum_k \omega_{ik} H(\mathbf{r}_k|\mathbf{r}_j) + \delta_{ij} - \frac{1}{N}, \tag{1.74}$$

where N is the total number of sites of the lattice. This function is also symmetrical in its arguments, and the sum $\sum_i H(\mathbf{r}_i|\mathbf{r}_j)$ is a constant independent of j. We obtain:

$$\begin{aligned} JH(\mathbf{r}_i|\mathbf{r}_S) - J_1 H(\mathbf{r}_i|\mathbf{r}_{T_1}) - J_2 H(\mathbf{r}_i|\mathbf{r}_{T_2}) &= \sum_k \omega_{ik} \left(JH(\mathbf{r}_k|\mathbf{r}_S) - J_1 H(\mathbf{r}_k|\mathbf{r}_{T_1}) - J_2 H(\mathbf{r}_k|\mathbf{r}_{T_2}) \right) \\ &\quad + J\delta_{iS} - J_1 \delta_{iT_1} - J_2 \delta_{iT_2}. \end{aligned} \tag{1.75}$$

It can be seen that $JH(\mathbf{r}_i|\mathbf{r}_S) - J_1 H(\mathbf{r}_i|\mathbf{r}_{T_1}) - J_2 H(\mathbf{r}_i|\mathbf{r}_{T_2})$ and $\rho(\mathbf{r}_i)$ solve the same equation, so that up to a constant ρ_0, we have:

$$\rho(\mathbf{r}_i) = \rho_0 + JH(\mathbf{r}_i|\mathbf{r}_S) - J_1 H(\mathbf{r}_i|\mathbf{r}_{T_1}) - J_2 H(\mathbf{r}_i|\mathbf{r}_{T_2}). \tag{1.76}$$

ρ_0 can be determined using the fact that $\sum_i H(\mathbf{r}_i|\mathbf{r}_j)$ is a constant. If we sum the previous equation over all nodes, we obtain:

$$\sum_i \rho(\mathbf{r}_i) = N\rho_0 = \mathcal{N} = J\langle \mathbf{T} \rangle \Rightarrow \rho_0 = \frac{J\langle \mathbf{T} \rangle}{N} \tag{1.77}$$

We choose flux J_1 and J_2 so that $\rho(\mathbf{r}_{T_1}) = \rho(\mathbf{r}_{T_2}) = 0$. This means that T_1 and T_2 are perfect targets, any particle reaching the target is instantaneously removed. This choice gives at last three equations, remembering that $J_i = P_i J$:

$$\begin{cases} \rho_0 + JH(\mathbf{r}_{T_1}|\mathbf{r}_S) - JP_1 H(\mathbf{r}_{T_1}|\mathbf{r}_{T_1}) - JP_2 H(\mathbf{r}_{T_1}|\mathbf{r}_{T_2}) &= 0 \\ \rho_0 + JH(\mathbf{r}_{T_2}|\mathbf{r}_S) - JP_1 H(\mathbf{r}_{T_2}|\mathbf{r}_{T_1}) - JP_2 H(\mathbf{r}_{T_2}|\mathbf{r}_{T_2}) &= 0 \\ P_1 + P_2 &= 1 \end{cases} \tag{1.78}$$

We can deduce from this system P_1, P_2 and $\langle \mathbf{T} \rangle$, adopting a more concise notation for pseudo-Green functions ($H(\mathbf{r}_{T_i}|\mathbf{r}_{T_j}) = H_{ij}$, $H(\mathbf{r}_{T_i}|\mathbf{r}_S) = H_{iS}$ and $H(\mathbf{r}_{T_i}|\mathbf{r}_{T_i}) = H_{0i}$):

$$\begin{cases} P_1 &= \dfrac{H_{1S} + H_{02} - H_{2S} - H_{12}}{H_{01} + H_{02} - 2H_{12}} \\ P_2 &= \dfrac{H_{2S} + H_{01} - H_{1S} - H_{12}}{H_{01} + H_{02} - 2H_{12}} \\ \langle \mathbf{T} \rangle &= N \dfrac{(H_{01} - H_{1S})(H_{02} - H_{2S}) - (H_{12} - H_{2S})(H_{12} - H_{1S})}{H_{01} + H_{02} - 2H_{12}} \end{cases} \tag{1.79}$$

We can easily extend this formalism to an arbitrary number of targets: each target gives an equation, and normalization of splitting probabilities another one. With n targets, we have $n+1$ equations, just enough to get the n splitting probabilities P_i and the mean first-passage time $\langle \mathbf{T} \rangle$.

We can also obtain the conditional MFPT $\langle \mathbf{T}_k \rangle$, namely the mean first-passage time from $\mathbf{r}_S$ to $\mathbf{r}_{T_k}$ knowing that T_k is the first target reached. To do so, we first compute the average number of particle $\mathcal{N}_k$ in the domain that will eventually be absorbed in $\mathbf{r}_{T_k}$. We have $\mathcal{N}_k = J_k \langle \mathbf{T}_k \rangle$. This average number is given by:

$$\mathcal{N}_k = \sum_i \rho(\mathbf{r}_i) P_k(\mathbf{r}_i), \tag{1.80}$$

where $P_k(\mathbf{r}_i)$ is the splitting probability to reach target T_k first, starting from $\mathbf{r}_i$. This equation is exact, but can be quite complex to deal with. For the two targets case, we have:

$$\mathcal{N}_1 = \sum_i \frac{(H_{i1} - H_{i2} + H_{02} - H_{12})(\rho_0 + JH_{iS} - J_1 H_{i1} - J_2 H_{i2})}{H_{01} + H_{02} - 2H_{12}} \tag{1.81}$$

We can deduce the conditional MFPT using this formula

$$\langle \mathbf{T}_1 \rangle = \sum_i \left(\frac{H_{i1} - H_{i2} + H_{02} - H_{12}}{H_{01} + H_{02} - 2H_{12}} \frac{\langle \mathbf{T} \rangle + H_{iS} - P_1 H_{i1} - P_2 H_{i2}}{P_1} \right) \tag{1.82}$$

where P_1, P_2 and$\langle T \rangle$ are the one deduced previously.

Those expressions are valid for any Markovian random walker evolving on a regular lattice, and can be adapted to other lattices using a similar approach. Splitting probabilities and related observables can be written using pseudo-Green function, we thus need to compute or to estimate those functions.

We emphasize that the results obtained here are *exact*: if the pseudo-Green functions are known analytically, we get analytically the splitting probabilities or the conditional MFPTs. If we approximate the pseudo-Green functions, we can estimate those observables, as well as the error margin, using the error margin of the pseudo-Green functions.

Pseudo-Green functions

Reference [65] also provide exact expressions of pseudo-Green function H for regular hypercubic Euclidian networks, with periodic or reflective boundary conditions. With periodic boundaries, we can obtain a general result in d dimension using the definition of (1.74):

$$H(\mathbf{r}_i|\mathbf{r}_j) = \frac{\sum_k H(\mathbf{r}_i \pm \mathbf{e}_k|\mathbf{r}_j)}{2d} + \delta_{ij} - \frac{1}{N}, \tag{1.83}$$

Using a discrete Fourier transform, and $(\mathbf{e}_1, \ldots, \mathbf{e}_d)$ as an orthogonal base, we get:

$$\widehat{H}(\mathbf{q}|\mathbf{r}_j) = \frac{\sum_k \widehat{H}(\mathbf{q}|\mathbf{r}_j) \cos(\mathbf{q}.\mathbf{e}_k)}{2d} + e^{-\imath \mathbf{q}.\mathbf{r}_j} - \delta_{\mathbf{q},\mathbf{0}} \Rightarrow \widehat{H}(\mathbf{q}|\mathbf{r}_j) = \frac{e^{-\imath \mathbf{q}.\mathbf{r}_j} - \delta_{\mathbf{q},\mathbf{0}}}{1 - \frac{1}{2d}\sum_k \cos(\mathbf{q}.\mathbf{e}_k)} \tag{1.84}$$

where $\mathbf{q} = (2\pi k_1/X_1, \ldots, 2\pi k_d/X_d)$, with X_i the lattice size in the ith direction, and $k_i \in [0, X_i - 1]$. An inverse discrete Fourier transform finally gives:

$$H(\mathbf{r}_i|\mathbf{r}_j) = \frac{1}{N} \sum_{\mathbf{q} \neq \mathbf{0}} \frac{e^{\imath \mathbf{q}.(\mathbf{r}_i - \mathbf{r}_j)}}{1 - \frac{1}{2d}\sum_k \cos(\mathbf{q}.\mathbf{e}_k)} \tag{1.85}$$

$N = \prod_i X_i$ is here the volume. We can see that the pseudo-Green functions only depends on the vector $\mathbf{r}_i - \mathbf{r}_j$, and not on the position of $\mathbf{r}_i$, as expected for a lattice with periodic boundary conditions.

We can retrieve this result using directly the propagator. We consider a d-dimensional Euclidian lattice of size $X_1 \times \ldots \times X_d$, with periodic boundary conditions. We can note that $P_{\text{stat}} = 1/N = 1/\prod_i X_i$. If we look at the propagator, we have, for a nearest neighbor random walk on an Euclidian lattice:

$$P(\mathbf{r}, t|\mathbf{r}_S) = \frac{1}{2d} \sum_{i=1}^{d} \left(P(\mathbf{r} - \mathbf{e}_i, t-1|\mathbf{r}_S) + P(\mathbf{r} + \mathbf{e}_i, t-1|\mathbf{r}_S) \right) \tag{1.86}$$

If we perform a discrete Fourier transform, we get:

$$\begin{aligned}\widehat{P}(\mathbf{q},t|\mathbf{r}_S) &= \frac{1}{d}\sum_{i=1}^{d}\cos\left(2\pi\mathbf{q}.\mathbf{e}_i\right)\widehat{P}(\mathbf{q},t-1|\mathbf{r}_S)\\ &= \left(\frac{1}{d}\sum_{i=1}^{d}\cos\left(2\pi\mathbf{q}.\mathbf{e}_i\right)\right)^t\widehat{P}(\mathbf{q},t=0|\mathbf{r}_S),\end{aligned} \tag{1.87}$$

where $\mathbf{q}=\sum_i k_i\mathbf{e}_i/X_i$, with $k_i\in[0,X_i-1]$. The random walker is initially in $\mathbf{r}_S$, we thus have $\widehat{P}(\mathbf{q},t=0|\mathbf{r}_S)=\exp(-2\imath\pi\mathbf{q}.\mathbf{r}_S)$. We can now perform an inverse discrete Fourier transform:

$$P(\mathbf{r},t|\mathbf{r}_S)-P_{\text{stat}}(\mathbf{r})=\sum_{\mathbf{q}\neq\mathbf{0}}\left(\frac{1}{d}\sum_{i=1}^{d}\cos\left(2\pi\mathbf{q}.\mathbf{e}_i\right)\right)^t e^{2\imath\pi\mathbf{q}.\mathbf{r}-\mathbf{r}_S}, \tag{1.88}$$

We remove the $\mathbf{q}=\mathbf{0}$ term to take in account the $-P_{\text{stat}}$ term. To get the pseudo-Green function, we just have to integrate this result over time. Since we are here in discrete time, the integration is in fact a summation:

$$H(\mathbf{r}|\mathbf{r}_S)=\sum_{t=0}^{\infty}\left(P(\mathbf{r},t|\mathbf{r}_S)-P_{\text{stat}}(\mathbf{r})\right)=\sum_{\mathbf{q}\neq\mathbf{0}}\frac{e^{2\imath\pi\mathbf{q}.\mathbf{r}-\mathbf{r}_S}}{1-\frac{1}{d}\sum_{i=1}^{d}\cos\left(2\pi\mathbf{q}.\mathbf{e}_i\right)}, \tag{1.89}$$

We retrieve the same expression as the one obtained starting with the functional equation satisfied by pseudo-Green functions.

This expression can be transposed to reflective boundary lattices, using the image method. Details can be found in reference [65], but the basic idea is to consider for each wall that an "image" of the network lies behind the wall, and that instead of being reflected by a straight wall, the random walker can evolve in the "image" network. This approach works only for "straight" walls (parallelepipedic domains), and gives in fact back the periodic boundary problem, that have been solved previously.

With the assumption that a random walker hitting a boundary stay in the same positions, we obtain for instance for a $X\times Y$ rectangle,

$$\begin{aligned}H(\mathbf{r}|\mathbf{r}') &= \frac{4}{N}\sum_{m=1}^{X-1}\sum_{n=1}^{Y-1}\frac{\cos\left(\frac{m\pi x'}{X}\right)\cos\left(\frac{n\pi y'}{Y}\right)\cos\left(\frac{m\pi x}{X}\right)\cos\left(\frac{n\pi y}{Y}\right)}{1-\frac{1}{2}\left(\cos\left(\frac{m\pi}{X}\right)+\cos\left(\frac{n\pi}{Y}\right)\right)}\\ &+\frac{4}{N}\sum_{m=1}^{X-1}\frac{\cos\left(\frac{m\pi x'}{X}\right)\cos\left(\frac{m\pi x}{X}\right)}{1-\cos\left(\frac{m\pi}{X}\right)}+\frac{4}{N}\sum_{n=1}^{Y-1}\frac{\cos\left(\frac{n\pi y'}{Y}\right)\cos\left(\frac{n\pi y}{Y}\right)}{1-\cos\left(\frac{n\pi}{Y}\right)},\end{aligned} \tag{1.90}$$

with$\mathbf{r}=(x,y)$ and where the coordinates x and y are half-integers going from $1/2$ to $X-1/2$ or $Y-1/2$.

At last, in more general domains, one can approximate H by the infinite-space lattice Green function G_0 [14], G_0 being evaluated as $G_0(\mathbf{r}|\mathbf{r}')=3/(2\pi\|\mathbf{r}-\mathbf{r}'\|)$ for $\mathbf{r}\neq\mathbf{r}'$, and $G_0(\mathbf{r}|\mathbf{r})=1.516...$ in three dimensions, and $G_0(\mathbf{r}|\mathbf{r}')=-(2/\pi)\ln|\mathbf{r}-\mathbf{r}'|$ for $\mathbf{r}\neq\mathbf{r}'$, and $G_0(\mathbf{r}|\mathbf{r})=1.029...$ in

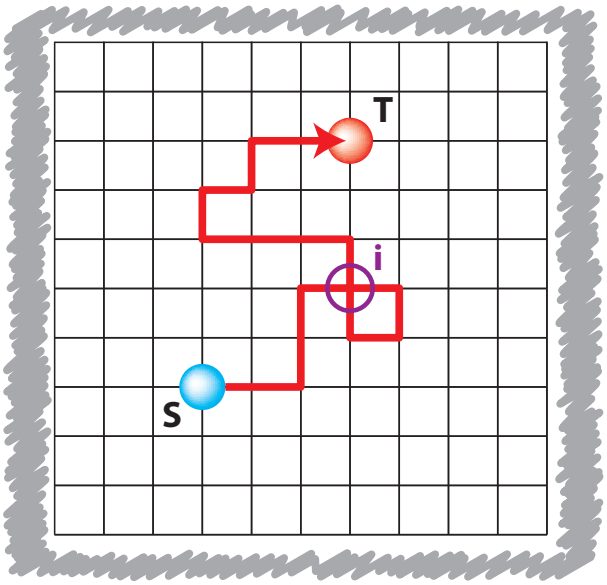

Figure 1.12: Schematic picture of the occupation time problem.

two dimensions. More accurate approximations can be found [65] for pseudo-Green function in general domains near a boundary, but the above approximations are good enough to capture the qualitative behavior of the pseudo-Green function on usual networks.

Those evaluations are crucial for first-passage time observables. We have shown how to link *analytically* first-passage properties with pseudo-Green functions: being able to get exactly pseudo-Green functions leads directly to an analytical expression of first-passage observables. If we only have a very good approximation of pseudo-Green functions, we will still being able to estimate the first-passage observables with a very good level of confidence, since the only approximation will be in the pseudo-Green functions estimate. We will use this to extract the MFPT or splitting probabilities dependence with lattice volume, with distances, or with lattice topology, knowing the pseudo-Green functions.

Occupation times

The second paper [67] has been written during my Master thesis, and use the previous results to compute new first-passage observables, like the mean occupation time. We will go deeper in details concerning this paper, since the solution found here will somehow be transposed later in the manuscript.

The occupation time is the number of times a given site i of a lattice has been visited by a random walker. The statistics of this general quantity has been studied for long, both by mathematicians [6, 111] and physicists [19, 21, 35, 63, 47, 136, 36, 97]. Occupation time has proven to be a key quantity in various fields, ranging from astrophysics [87], transport in porous media [41] or diffusion limited reactions [20]. The point is that as soon as the sites of a system have different physical or chemical properties, it becomes crucial to know precisely how many times each site is visited by the random walker.

Figure 1.12 shows a schematic picture of the problem: the random walk begins at the site S, and the occupation time $\mathbf{N_i}$ is the number of times it visits the site i before reaching the target T. In this picture, $\mathbf{N_i} = 2$.

Occupation time is for instance used to compute the MFPT in a random-trap model: in such model, the random walker wait a time τ_i in position $\mathbf{r}_i$. The relation with the occupation time is the following: the MFPT $\langle \mathbf{T} \rangle$ at the target $\mathbf{r}_T$ starting from site $\mathbf{r}_S$ can be written down as

$$\langle \mathbf{T} \rangle = \sum_{i=1}^{N} \langle \mathbf{N_i} \rangle \tau_i, \tag{1.91}$$

where N is the volume of the confining system, $\mathbf{N_i}$ is the number of times the site $\mathbf{r}_i$ has been visited before the target is reached and $\langle \ldots \rangle$ stands for the average with respect to the random walk. If $\forall i, \tau_i = 1$, we retrieve the relation between occupation time and MFPT in an usual random walk:

$$\langle \mathbf{T} \rangle = \sum_{i=1}^{N} \langle \mathbf{N_i} \rangle, \tag{1.92}$$

Concerning the distribution of the MFPT with respect to the disorder, that is with respect to the τ_i's, we are finally back to summing a deterministic number N of independent random variables $\langle \mathbf{N_i} \rangle \tau_i$ but non identically distributed (because of the factor $\langle \mathbf{N_i} \rangle$), which requires the determination of the *mean* occupation times $\langle \mathbf{N_i} \rangle$ introduced before.

The whole distribution can become necessary if we look at diffusion limited reactions in confined media. We can for instance consider a free diffusing reactant A that enters in a cavity and can react with a given fixed center $\mathbf{r}_i$. We assume that each time the walker reaches the reactive site $\mathbf{r}_i$, it has a probability p to react, which schematically mimics an imperfect reaction in confined geometry. Actually, numerous chemical reactions, ranging from trapping in supermolecules [10] to activation processes of synaptic receptors[107, 179] can be roughly rephrased by this generic scheme. Knowing the occupation time distribution in $\mathbf{r}_i$, we can compute the probability for A to react with the center $\mathbf{r}_i$ before exiting the cavity. More generally, for a random walker starting from a site $\mathbf{r}_S$, we can get the probability Q to react with $\mathbf{r}_i$ before reaching a target site $\mathbf{r}_T$, possibly different from $\mathbf{r}_S$. Partitioning over the number of times the reactive site $\mathbf{r}_i$ has been visited, we have:

$$Q = 1 - \sum_{k=0}^{\infty} P(\mathbf{N_i} = k)(1-p)^k. \tag{1.93}$$

Once again, the random variable $\mathbf{N_i}$ is involved, but that time the determination of the *entire distribution* $P(\mathbf{N_i} = k)$ is needed.

We present here the method of computation of the $\mathbf{N_i}$ statistics in confining geometry developed in [67]. In particular, we obtain explicitly the *exact distribution* in the case of parallepipedic confining domains.

We start with the computation of the mean $\langle \mathbf{N_i} \rangle$, assuming for the time being that the starting and target sites are different ($\mathbf{r}_S \neq \mathbf{r}_T$). We note ω_{ij} the transition probabilities from site j to site i. We have $\sum_i \omega_{ij} = 1$, and we take $\omega_{ij} = \omega_{ji}$. These general transition probabilities can take into account reflecting boundary conditions.

We will adapt the electrical analogy considered before: we start by considering an incoming flux J of particles in $\mathbf{r}_S$. Since the domain is finite, all the particles are eventually absorbed in $\mathbf{r}_T$, and in the stationary regime, there is an outgoing flux J of particles in $\mathbf{r}_T$. As previously, the mean particle density ρ_i satisfies the following equation:

$$\rho(\mathbf{r}_i) = \rho_0 + JH(\mathbf{r}_i|\mathbf{r}_S) - J_1 H(\mathbf{r}_i|\mathbf{r}_{T_1}) - J_2 H(\mathbf{r}_i|\mathbf{r}_{T_2}), \tag{1.94}$$

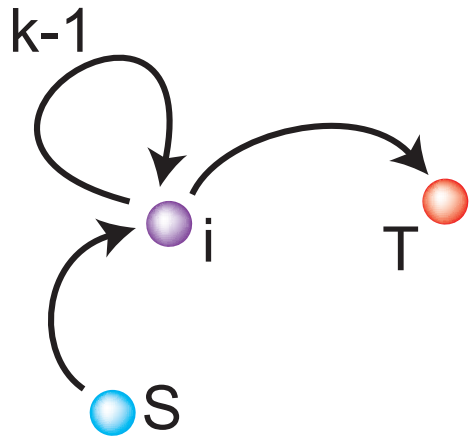

Figure 1.13: Scheme of the mechanism used to get the full occupation time distribution.

where ρ_0 is (we have only one target):

$$\rho_0 = \frac{J\langle \mathbf{T} \rangle}{N} = J\left(H_{TT} - H_{TS}\right) \tag{1.95}$$

We used here the fact that $\langle \mathbf{T} \rangle = N(H_{TT} - H_{TS})$, as shown in section 1.2.2. To find the mean occupation time, we can simply notice that the mean particle density $\rho(\mathbf{r}_i)$ is equal to $\langle \mathbf{N_i} \rangle J$.

$$\langle \mathbf{N_i} \rangle = \frac{\rho(\mathbf{r}_i)}{J} = H_{iS} - H_{iT} + H_{TT} - H_{ST}, \tag{1.96}$$

This equation satisfies the boundary condition $\rho(\mathbf{r}_T) = 0$.

Equality $\rho(\mathbf{r}_i) = \langle \mathbf{N_i} \rangle J$ is related to the ergodic property of the process: for the stationary problem considered, with a constant flux J exiting in $\mathbf{r}_T$ and incoming in $\mathbf{r}_S$, the stationary state is equivalent to a time average. Indeed, at any time t, we can consider that we superpose the trajectories started between $-\infty$ and t. The stationary probability to be in $\mathbf{r}_i$ is the time average of the probability to be in $\mathbf{r}_i$ at time $t' \in [0, \infty[$ for a single trajectory starting from $\mathbf{r}_S$ and being absorbed in $\mathbf{r}_T$ (up to a scaling factor J):

$$\rho(\mathbf{r}_i) = J \sum_{t=0}^{\infty} P(\mathbf{r}_i, t | \mathbf{r}_S) = J \langle \mathbf{N}_i \rangle_t \tag{1.97}$$

Since the process is ergodic, time average $\langle \ldots \rangle_t$ and ensemble average $\langle \ldots \rangle$ are the same.

Note that equation (1.96) also gives the mean occupation time of a subdomain, which is simply the sum of the mean occupation time of all the sites in the subdomain.

We can go further and obtain the entire distribution of the occupation time. The idea to tackle this *a priori* difficult problem is to use the splitting probabilities computed previously. To hit the site $\mathbf{r}_i$ k times, the random walker starting from $\mathbf{r}_S$ has first to hit $\mathbf{r}_i$ once, before reaching $\mathbf{r}_T$, and then to come back $k-1$ times without touching $\mathbf{r}_T$, and at last to hit $\mathbf{r}_T$. Figure 1.13 shows this mechanism: each arrow is associated with a splitting probability between $\mathbf{r}_i$ and $\mathbf{r}_T$, starting either from $\mathbf{r}_S$ or from $\mathbf{r}_i$.

Using this scheme, and denoting here $P_{ij}(i|S)$ the splitting probability to reach $\mathbf{r}_i$ before $\mathbf{r}_j$, starting from $\mathbf{r}_S$, we have $P(\mathbf{N_i} = 0) = P_{iT}(T|S)$, and for $k \geq 1$:

$$P(\mathbf{N_i} = k) = P_{iT}(i|S) \left[\sum_j \omega_{ji} P_{iT}(i|j) \right]^{k-1} \left[\sum_j \omega_{ji} P_{iT}(T|j) \right] \tag{1.98}$$

The three terms of this last equation correspond respectively to

- the probability to reach $\mathbf{r}_i$ before $\mathbf{r}_T$, starting from $\mathbf{r}_S$,
- the probability to return to $\mathbf{r}_i$ before reaching $\mathbf{r}_T$, starting from $\mathbf{r}_i$, to the power $k-1$,
- and the probability to reach $\mathbf{r}_T$ before returning to $\mathbf{r}_i$.

For the last two probabilities, we sum over all neighbors of $\mathbf{r}_i$: to return in $\mathbf{r}_i$, the random walker first has to go to a neighbor $\mathbf{r}_j$, with a probability ω_{ij}, and then to go back in $\mathbf{r}_i$ (or to go in $\mathbf{r}_T$) starting from $\mathbf{r}_j$.

Equation 1.98 can then be written

$$P(\mathbf{N_i} = k) = AB(1-B)^{k-1} \text{ for } k \geq 1, \tag{1.99}$$

with, using equations (1.79)

$$A \equiv P_{iT}(i|S) = \frac{H_{iS} + H_{TT} - H_{ST} - H_{iT}}{H_{ii} + H_{TT} - 2H_{iT}}, \tag{1.100}$$

and

$$\begin{aligned} B &\equiv \sum_j \omega_{ji} P_{iT}(T|j) = 1 - \sum \omega_{ji} P_{iT}(i|j) \\ &= \frac{\sum_j \omega_{ji} H_{Tj} - H_{iT} - \sum_j \omega_{ji} H_{ji} + H_{ii}}{H_{ii} + H_{TT} - 2H_{iT}} \\ &= \frac{1}{H_{ii} + H_{TT} - 2H_{iT}}, \end{aligned} \tag{1.101}$$

where we used equation (1.74) and the fact that $\sum_i w_{ij} = 1$. It can also be noted that $P(\mathbf{N}_i = 0) = 1 - A$. The distribution of the occupation times given by equations (1.99)-(1.101) is the main result of this subsection, and several comments are in order.

(i) Expressions of pseudo-Green function H given in equation (1.90) makes this result exact and completely explicit for parallepipedic domains.

(ii) Computing $\langle \mathbf{N_i} \rangle$ with this distribution gives back the expected result (1.96).

(iii) It can be noted here that B, which characterizes the decay of the probability distribution of $\mathbf{N_i}$, is independent of the source. In addition, qualitatively, the basic evaluations of H (namely $H = G_0$) give for B the following order of magnitude, if $\mathbf{r}_i$ and $\mathbf{r}_T$ are at a distance R:

$$B \simeq \begin{cases} \dfrac{1}{2G_0(0) - \dfrac{3}{\pi R}} & \text{in 3D,} \\ \dfrac{1}{2G_0(0) + \dfrac{4}{\pi} \ln(R)} & \text{in 2D,} \end{cases} \tag{1.102}$$

where $G_0(0) = G_0(\mathbf{r}|\mathbf{r})$ is a dimension-dependent constant, given above in the discussion on the evaluation of pseudo-Green function H. This shows that B decreases with the distance between i and T: a larger distance corresponds to a slower decay. But, while it tends towards 0 in two dimensions (which corresponds to a wide distribution of $\mathbf{N_i}$, and a large variance), it tends to a finite value in three dimensions. It can thus be said that the sites much further from the target than the source have, in three dimensions, a significant probability to be visited, but a low probability to be visited many times, whereas, in two dimensions, they have a low probability to be visited at all, but a comparatively high probability to be visited many times. This is connected with the transient or recurrent character of the unbounded random walk in two or three dimensions.

(iv) The results obtained here for different starting and target sites may easily be adapted to identical starting and target sites ($S = T$):

$$P(\mathbf{N_i} = 0) = 1 - B \; ; \; P(\mathbf{N_i} = k) = B^2(1-B)^{k-1} \text{ for } k \geq 1. \tag{1.103}$$

Note that this gives in particular a mean occupation time of 1 for all sites, a result which could be derived from an extension of Kac's formula [6, 65]. However, here, we obtain not only the mean occupation time but the entire distribution of this occupation time, which appears to vary from site to site: the further the site is from the target, the slower the probability distribution decays.

To conclude, we have computed the distribution of the occupation time of a given site $\mathbf{r}_i$, for a random walk in confined geometry, eventually trapped at a target. This distribution is exact, and completely explicit in the case or parallepipedic confining domains. While the mean occupation time, unsurprisingly, is higher when $\mathbf{r}_i$ is near the source and lower near the target (and uniform if the source and target are identical), the distribution of the occupation time is essentially exponential, with a slower decay when the point is far away from the target.

Before going to the next section, we insist on the fact that pseudo-Green function H are the key of first-passage observables. If we are able to compute or to estimate those functions, we will obtain directly the mean first passage time, the splitting probabilities as well as the entire occupation times distribution.

Quick summary

- We have introduced first-passage observables that will be used all along this manuscript ;
- with the renewal equation, we saw that pseudo-Green functions H were a way to compute first-passage observables ;
- we showed that those pseudo-Green functions could be computed analytically or estimated with infinite Green functions.

1.3 Beyond Brownian motion

So far, we focused on nearest-neighbor random walks on regular lattices, where $P_{\text{stat}} = 1/N$. Those random walks can be seen as a discretization of a regular continuous random walks on an

Euclidian space. We will introduced in this section ways to extend the results previously obtained on first-passage properties to a wider range of random walk mechanisms.

In the first sub-section, we will present the prominent (continuous) anomalous diffusion models. Those models will be the ground on which is built the chapter 2. They all have a mean square displacement scaling like $\langle r^2 \rangle \propto t^\alpha$ with $\alpha \neq 1$, a feature often observed in experimental trajectories. We will introduce the three sub-diffusive processes studied in chapter 2, namely continuous time random walks (CTRW), fractional Brownian motion (fBm) and diffusion on fractals, as well as a super-diffusive process used in chapter 3, the Lévy fligths.

The second sub-section will be dedicated to (discrete) random walks on self-similar networks, that will be used to simulate the continuous processes of anomalous diffusion. We will thus introduce in this sub-section the networks that will be used during this manuscript, as soon as numerical simulation on self-similar networks will be required to check an analytical result. It will mainly be deterministic fractals, like (u, v)–flowers or the Sierpinski gasket, and random fractals, like percolation networks or Erdös-Rényi networks.

At last, in the third sub-section, we will see how we can have a generic approximation of pseudo-Green functions for self-similar random walks, either discrete or continuous. This approximation allows to obtain the MFPT dependence in geometrical parameters, or in environment characteristics. Those results have been published in [66], and were part of my Master thesis.

1.3.1 Anomalous diffusion processes

The mean square displacement can sometimes not scale linearly with time. Such phenomenon is called *anomalous diffusion*:

$$\langle r^2 \rangle \propto t^\alpha \tag{1.104}$$

where α is the anomalous diffusion coefficient. We will talk about sub-diffusion if $\alpha < 1$ and super-diffusion for $\alpha > 1$. The chapter 2 of this manuscript will be focused sub-diffusion, often observed experimentally, but that can be explained by several models.

We will describe in this sub-section processes that take place in the usual Euclidian space, in continuous time, and that lead to anomalous diffusion. For each model, we will describe the mathematical aspect that will be used in chapter 2 (for sub-diffusive models) as well as an "intuitive" explanation of the origin of anomalous diffusion. Four models will be successively described: continuous time random walk (CTRW), that could be applied to discrete networks, diffusion on fractals, related to what could be obtained with the discrete networks previously introduced, fractional Brownian motion (fBm) and Lévy flights, that will be defined in continuous space.

CTRW

Continuous time random-walks (CTRW) is a first model for sub-diffusion, that relies on the presence of "traps", that immobilize the walker for a random time at each step. This random time is distributed according to a heavy tailed law in order to lead to sub-diffusion. This process is non Markovian, even if slightly: we just change here the "clock" of the random walk. If we assume that the waiting times at each step are 1, we find back a Markovian process. The clock modification due to the waiting times lead also to a long range correlation for a given random walk: time average and ensemble average does not lead to the same result, ergodicity is broken [16].

This model is used in a wide variety of fields [115, 172], ranging from charge carrier motion in amorphous semiconductors [187], over tracer diffusion in underground aquifers [186], up to

weakly chaotic systems [194]. CTRW can be introduced with their discrete version: let consider a random walker that perform a succession of jumps, each jump $\delta\mathbf{r}$ being a random variable whom probability density function is known. The jumps are assumed to be performed instantaneously, and between two jumps, the random walker rests a random time ψ, drawn from another probability distribution $\psi(t)$ [187]. Two random processes occurs at each step: the time spent in each position is given by a random waiting time ψ, and the length of the next jump by a random jump length $\delta\mathbf{r}$.

If the waiting time distribution admit a finite first moment, and if the jump distribution admit a finite second moment, the central limit theorem ensures that after a while, the CTRW model gives back a usual Brownian motion.

If the waiting time distribution is of the "heavy tailed" form

$$\psi(t) \underset{t\to\infty}{\sim} \frac{\alpha\tau^\alpha}{\Gamma(1-\alpha)t^{1+\alpha}}, \tag{1.105}$$

for $0 < \alpha < 1$, the mean waiting time $\langle\psi\rangle$ diverges, and the resulting process becomes sub-diffusive with mean square displacement satisfying

$$\langle r^2\rangle = K_\alpha t^\alpha. \tag{1.106}$$

The exponent α is then the one of the waiting time density (1.105). The generalized diffusion coefficient K_α is related to the second moment of the jump length distribution is $\langle\delta\mathbf{r}^2\rangle$

$$K_\alpha = \frac{\langle\delta\mathbf{r}^2\rangle}{2d\tau}. \tag{1.107}$$

The anomalous behavior in these models originates from a heavy tailed distribution of waiting times[201]: without this heavy tail, we find back a Brownian motion, with, after a while, $\langle r^2\rangle \propto t$. Waiting times with power-law distribution as in equation (1.105) has been observed experimentally, for instance for the motion of probes in a reconstituted actin network [228].

CTRW is a model adapted to describe a random walker that can be trapped by the environment, potentially for a very long time. For tracers diffusing in biological cells, traps can be out-of-equilibrium chemical binding configurations[181, 182], and the waiting times the dissociation times; traps can also be realized by the free cages around the tracer in a hard sphere like crowded environment, and the waiting times are the life times of the cages (see Figure 1.14).

To treat mathematically CTRW, we will use in chapter 2 their continuous limit, described by fractional diffusion equations[172, 229]. Assuming that $\psi \sim t^{-(1+\alpha)}$ for large t, with $0 < \alpha < 1$, we have:

$$\frac{\partial P(r,t)}{\partial t} = \frac{K_\alpha}{r^{d-1}}\frac{\partial}{\partial r}\left(r^{d-1}\frac{\partial}{\partial r}\left({}_0D_t^{1-\alpha}P(r,t)\right)\right), \tag{1.108}$$

where we used the Riemann-Liouville fractional operator:

$${}_0D_t^{1-\alpha}P(r,t) = \frac{1}{\Gamma(\alpha)}\frac{\partial}{\partial t}\int_0^t \frac{P(r,t')}{(t-t')^{1-\alpha}}dt', \tag{1.109}$$

We will see in chapter 2 that using subordination [227], we can transpose all results obtained for regular random walk, if we replace s by K_α/Ds^α in the Laplace space.

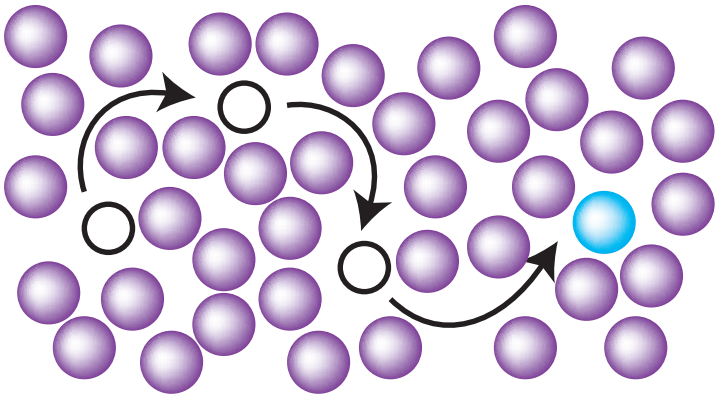

Figure 1.14: Random walk in a crowded environment with moving obstacles forming dynamical cages.

Diffusion on fractals

A second kind of model for sub-diffusion relies on spatial inhomogeneities. In such model, the entire space is not accessible, the random walker evolve on a space where bottlenecks and dead ends exists on all scales. A discrete version is exemplified by diffusion in deterministic or random fractals network, as illustrated by De Gennes's "ant in a labyrinth" [73]. This results in an effective sub-diffusion in the embedding space. While the fractal dimension d_f characterizes the geometry of the fractal, the diffusive dynamics involves the random walk exponent d_w.The latter is related to the anomalous diffusion exponent through $\alpha = 2/d_w$ [102] ($d_w \geq 2$ for sub-diffusion).

Fractals can be used to model complex networks, and have recently been suggested to mimic certain features of diffusion under conditions of molecular crowding [134, 140]. The anomalous behavior is in this case due to the presence of fixed obstacles [180], as shown in Figure 1.15, which creates numerous dead ends and leads to an anomalous diffusive behavior.

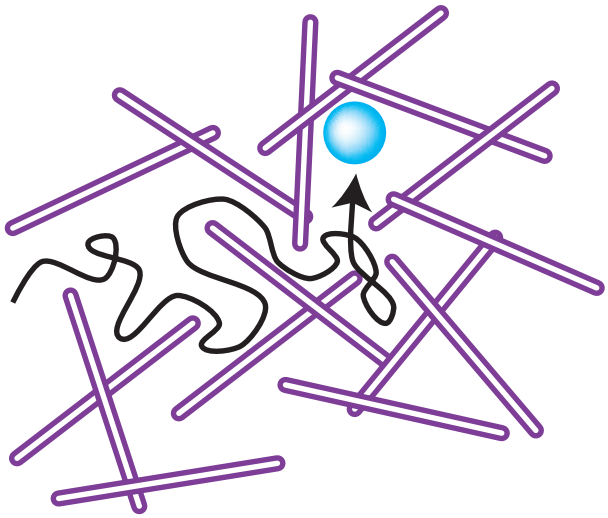

Figure 1.15: Random walk in a crowded environment with static obstacles.

For the continuous limit, corresponding for instance to a deterministic fractals network of

generation g, with $g \to \infty$, a partial differential equation has been proposed for the propagator [154]:

$$\frac{\partial P(r,t)}{\partial t} = \frac{K}{r^{d_f-1}} \frac{\partial}{\partial r} \left(r^{1-(d_w-d_f)} \frac{\partial P(r,t)}{\partial r} \right) \tag{1.110}$$

This equation has been found using scaling arguments, and will prove to be quite accurate, even if not analytical.

FBm

Fractional Brownian motion (fBm) is another model leading either to sub-diffusion or to super-diffusion. It has been introduced to take into account long-term spacial correlations in a random walk. In this model, the state of the system at time t is influenced by all states $t' \in]-\infty, t[$. FBm was introduced in 1D by Mandelbrot & Van Ness [137]: it is a normalized continuous Gaussian process $B^H(t)$, which starts at 0 ($B^H(t=0)=0$), has expectation zero for all $t \geq 0$, and whose auto-correlation function is:

$$\left\langle B^H(t) B^H(t') \right\rangle = \frac{1}{2} \left(|t|^{2H} + |t'|^{2H} - |t-t'|^{2H} \right) \tag{1.111}$$

$H \in [0,1]$ is called the Hurst index. Using this auto-correlation function, we obtain, for $t' = t$, that $\langle B^H(t)^2 \rangle = t^\alpha$, with $\alpha = 2H$. FBm therefore describes both sub-diffusion ($H < 1/2$) and super-diffusion ($H > 1/2$) up to the ballistic limit $\alpha = 2$.

This process is rather complex from the mathematical point of view. Since the process is infinitely correlated, we cannot obtain a (local) diffusion equation, as we had for CTRW or diffusion on fractals. Some tools have still been developed by mathematicians, for instance a representation using Ito' integrals [52]:

$$B_H(t) = \frac{1}{\Gamma(H+1/2)} \int_0^t (t-t')^{H-1/2} B(t') dt', \tag{1.112}$$

where $B(t)$ is an usual Brownian motion, and Γ the gamma function. We can check that $B_H(0) = 0$. We can also rewrite fBm in terms of fractional differential operators [189]:

$$B_H(t) =_0 D_t^{-(H+1/2)} B, \tag{1.113}$$

where ${}_0D_t^{-\alpha}$ is the Riemann-Liouville previously introduced, and B a classical Brownian motion.

This process is non Markovian: once again, time average and ensemble average do not lead to the same result. The propagator cannot be properly defined for fBm, since no partial differential equation exists for such process. We can still compute the one-point distribution of trajectories at time t starting from $t = -\infty$, and satisfying $B^H(t=0) = 0$, namely the pseudo-propagator starting from a "stationary" state (the increments have reached a stationary distribution, and the walk starts at 0):

$$P(x,t|0,0) = \sqrt{\frac{H}{\pi}} \frac{\Gamma(H+1/2)}{t^H} e^{-H\Gamma^2(H+1/2)\frac{x^2}{t^{2H}}} \tag{1.114}$$

This one-point distribution is Gaussian in space, and gives all moments of fBm, which will be used in chapter 2.

FBm is used to describe the motion of a monomer in a polymer chain [116] or single file diffusion [128]. FBm has recently been proposed to describe the diffusion in a crowded environment [207].

It is used as soon as one wants to introduce strong correlations (or anti-correlations) in the trajectory. For instance in the case of a particle evolving in a solvent, the particle motion displaces the solvent, and the resulting viscous force influences the future steps of the random walker. To know how the random walker at a time t will evolve, we need to know the solvent velocity field, which can be deduced from the past trajectory of the random walker.

FBm is defined in 1D, but following reference [214], we can extend fBm to several dimensions such that a d-dimensional fBm of exponent H is a process in which each of the coordinates follows a one-dimensional fBm of exponent H. The resulting d-dimensional fBm still satisfies relation $\langle r^2 \rangle \propto t^\alpha$, with $\alpha = 2H$.

Lévy flights

At last, we briefly introduce the Lévy flight model. In his discrete version, at each step, the random walker performs a random step drawn from a heavy-tailed distribution:

$$P(r) \underset{r\to\infty}{\sim} \frac{\alpha \tau^\alpha}{\Gamma(1-\alpha) r^{1+\alpha}} \tag{1.115}$$

This model is somehow similar to a CTRW, but it is now the jump distribution that is heavy tailed, and not the waiting time distribution. The generalized central limit theorem ensures that a sum of heavy tailed variable converges to a Lévy α-stable distribution. The characteristic function of this stable distribution is:

$$\phi(k) = \exp(-|ck|^\alpha), \tag{1.116}$$

where c is a scaling factor. If $\alpha = 2$, we find back the classical Brownian motion, with a variance $\sigma^2 = 2c^2$.

We can extend the discrete Lévy flight to a continuous process, where the propagator is a Lévy α-stable distribution, whose scaling factor is $t^{1/\alpha}c$ instead of c. The mean square displacement in such model in infinite, $\langle r^2 \rangle = \infty$. The main feature of Lévy flights is that they are self-similar, and that $r \propto t^{1/\alpha}$.

They have been used to describe albatrosses trajectories [218] for instance: the search process consists in straight ballistic phases (the jumps) that alternates with reorientation phases (between two jumps).

1.3.2 Self-similar networks

We now present scale-invariant networks that exhibit various fractal dimensions d_f and d_w. We recall that the fractal dimension d_f links the number of node $N(r)$ at distance smaller than r of a given node, $N(r) \propto r^{d_f}$. The walk dimension d_w links the typical distance r reached at time t with time through $t \propto r^{d_w}$. The electrical analogy [79] allows to compute the difference $d_w - d_f$: if we replace each link of a given network by a unitary resistance, then the equivalent resistance $R(r)$ between two points at a distance r fulfills $R(r) \propto r^{d_w - d_f}$.

The walk dimension d_w can be linked with the anomalous diffusion coefficient α: $r^2 \propto t^{2/d_w}$ leading to $\alpha = 2/d_w$.

As a general result, we can note that $d_w - d_f = 1$ for a scale-invariant tree, since the equivalent resistance between two points is directly the distance between those points: the absence of loop gives directly this result. Since $R(r) = r$, we get $d_w - d_f = 1$. Beside, for a tree, $\langle k \rangle = 2$.

We will see in the next sub-section how to link d_f, d_w and the compacity of the random walk.

Percolation networks

A general percolation network is created using a "regular" network, and occupying randomly each of the edges ("bond percolation") or each of the nodes ("node percolation") with a statistically independent probability p. We will here focus only on bond percolation.

We will call "clusters" a set of connected nodes on a percolation networks. When $p \to 0$, we have a large number of clusters on the percolation networks, and when $p = 1$, there is only one cluster. We call p_c the critical percolation threshold, such that starting from $p = 0$, a "giant" cluster appears for the first time. This giant cluster connect almost surely both extremities of the original network. Figure 1.16 shows how to construct a percolation cluster from a finite $2D$ Euclidian network: first we remove randomly the edges, then we identify the giant cluster, here connecting all extremities of the original network.

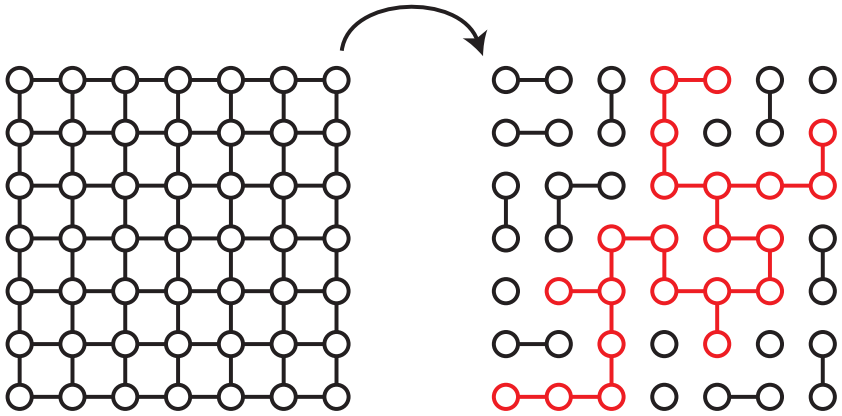

Figure 1.16: Creation of a percolation network starting from a 2D Euclidian network: the giant cluster is here in red.

In this manuscript, a percolation network will refer to the giant cluster of a bond percolation network on an Euclidian network in d dimension. Such network is random and self-similar: we will have to average over several network realization to get a meaningful result.

On the critical percolation cluster ($p = p_c$), the Alexander-Orbach conjecture[7] states that:

$$\frac{2d_f}{d_w} = \frac{4}{3} \tag{1.117}$$

For a $2D$ Euclidian network, using Euclidian distances, $p_c = 0.5$, and $d_f = 91/48$, which leads using the Alexander-Orbach conjecture to $d_w = 2.844\ldots$. For a $3D$ Euclidian network, once again with chemical distances, $p_c \simeq 0.2488\ldots$, $d_f = 2.523\ldots$[71], which gives $d_w = 3.783\ldots$. We will see in chapter 2 how to tranpose these results for chemical distances.

If $p < p_c$, the cluster have the same fractal dimension d_f, but is not "infinite", in the sense that, almost surely, it does not connect two extremities of the original network.

If $p > p_c$, the super-critical percolation cluster is fractal at small distances, with the same d_f than in the percolation cluster, and regular at higher distances, with the same d than the original network. The critical distance decreased as p increase.

Critical Erdös-Rényi networks

Erdös-Rényi networks can be defined as a percolation networks, but where the original network is a complete graph (a graph where each node is connected to all the other nodes). Equivalently,

it can be constructed as follows: we take N nodes, and for every pair of nodes (i, j), a link is created with probability p. The Erdös-Rényi network is then defined as the giant cluster, namely the larger set of connected nodes. The giant cluster is fractal only if $p < p_c$.

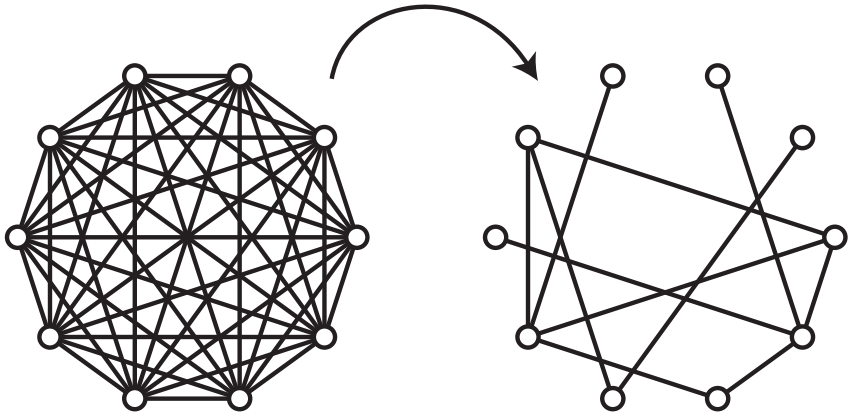

Figure 1.17: Percolation on a complete graph K_{10} leading to an Erdös-Rényi network.

The percolation threshold is $p_c = 1/N$, where N is the size of the original network. This leads to a network where the average connectivity $\langle k \rangle = 2$, the average being made on several realizations of the network. The fractal dimension for a critical percolation cluster is $d_f \simeq 1.9$ [203]. We computed numerically $d_w \simeq 2.9$. Erdös-Rényi network will in this manuscript refer to the critical percolation cluster on a complete graph ($p = p_c = 1/N$). Those networks are almost self-similar trees: if loops can occurs, we still have $\langle k \rangle \simeq 2$ and $d_w - d_f \simeq 1$.

Sierpinski gasket

As introduced previously (see Figure 1.6), the Sierpinski gasket is defined as a triangle where at each step, we remove an inner triangle in any full triangle. We have already shown that $d_f = \ln(3)/\ln(2)$, we can now compute $d_w - d_f$ using the electrical analogy. To do so, we compute the equivalent resistance between one edge and the two others. Figure 1.18 shows how to replace each link by a resistance R, and then how to simplify the resistance network to obtain an equivalent resistance.

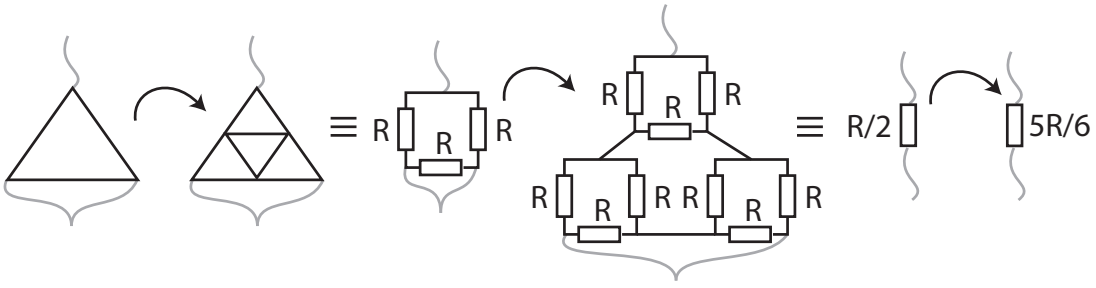

Figure 1.18: Equivalent resistance and $d_w - d_f$ computation for a Sierpinski gasket.

For the Sierpinski gasket, when the distance double, the resistance is multiplied by 5/3, leading to:

$$\frac{R(2r)}{R(r)} = \frac{5}{3} \Rightarrow 2^{d_w - d_f} = \frac{5}{3} \Rightarrow d_w - d_f = \frac{\ln(5/3)}{\ln(2)} \tag{1.118}$$

(u, v)–flowers

(u, v)–flowers are deterministic recursive networks described in reference [177]. The construction starts at generation $g = 0$ from two nodes linked. At each generation, every link is substituted by two paths of length u and v, containing thus respectively $u - 1$ and $v - 1$ nodes. Figure 1.19 shows the first generations of a $(3, 2)$–flower.

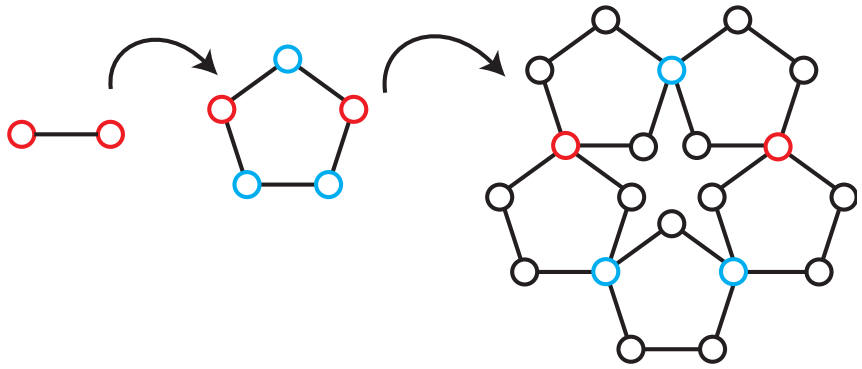

Figure 1.19: First three generations of a $(3, 2)$–flower, nodes created in the same generation having the same color.

At the gth generation, the link number is $L(g) = (u + v)^g$. The network size is:

$$N(g) = N(g-1)+(u+v-2)L(g-1) = 2+(u+v-2)\frac{(u+v)^g-1}{u+v-1} \underset{g\to\infty}{\sim} \frac{u+v-2}{u+v-1}(u+v)^g \quad (1.119)$$

To compute d_f and d_w with recursive deterministic networks, we can only focus on what happens to a given link. For a (u, v)–flowers, we replace a link by two paths. The distance between the two original point becomes u (we assume that $u \leq v$), and the node number becomes $u + v$. The fractal dimension is thus:

$$\frac{N(ur)}{N(r)} \sim (u + v) \Rightarrow N(r) \sim r^{d_f} \text{ with } d_f = \frac{\ln(u + v)}{\ln(u)} \quad (1.120)$$

At last, using the electrical analogy, starting from two points linked through a simple resistance, we get two resistance u and v in parallel, leading to an equivalent resistance of $uv/(u + v)$:

$$\frac{R(ur)}{R(r)} \sim \frac{uv}{u+v} \Rightarrow R(r) \sim r^{d_w-d_f} \text{ with } d_w = \frac{\ln(uv)}{\ln(u)} \quad (1.121)$$

One can note that those formula does not work for $u = 1$. In this case, the network is not fractal anymore, and other quantities have to be introduced to characterize the random walk [177].

We extended those networks to (u, v, w)–flowers, for which a third path is added. For those networks, if $1 < u \leq v \leq w$, we obtain similarly:

$$d_f = \ln(u + v + w)/\ln(u) \text{ and } d_w - d_f = -\ln(1/u + 1/v + 1/w)/\ln(u). \quad (1.122)$$

Random (u, v)–flowers

These networks are constructed like deterministic (u, v)–flowers as described in [211]: instead of transforming *every* link at each generation, only one link is substituted by two paths of length u and v. The resulting network is random and self-similar.

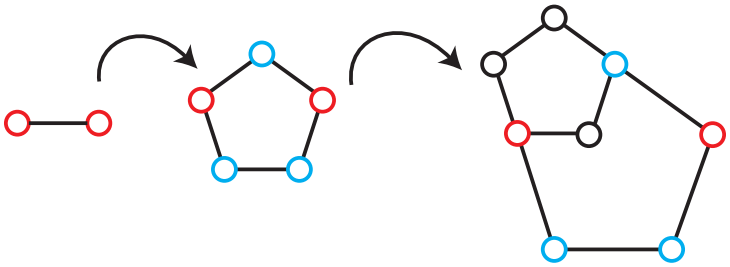

Figure 1.20: First three generations of a random $(3, 2)$–flowers, nodes created in the same generation having the same color.

Figure 1.20 shows how to construct such network. Dimensions d_f and d_w are determined numerically for those networks with a power-law fit of $N(r) \propto r^{d_f}$ and of $\langle r^2 \rangle \propto t^{2/d_w}$; for $(2, 2)$-random flowers for instance, we obtain a compact networks with $d_w - d_f \simeq 0.6$.

Networks of Kozma et al.

These networks, defined in [125], are classical Euclidian lattices in which long range links ("short-cuts") are added. A short-cut starts from each node with probability p, and leads to a node at a distance r where r is distributed according to a power law of index α. Figure 1.21 shows how to create a Kozma network starting from an Euclidian lattice.

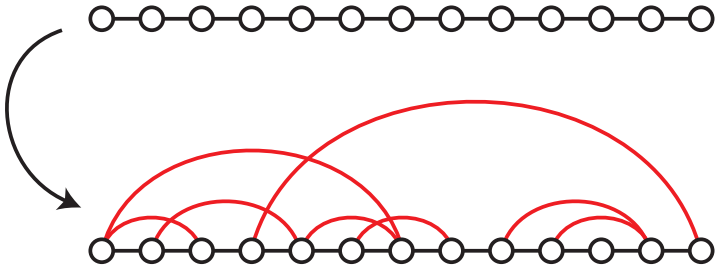

Figure 1.21: Creation of a Kozma network starting from a 1D Euclidian network.

We will focus on this manuscript on Kozma networks based on $1D$ Euclidian lattice. For such networks, the exploration is compact for $\alpha > 2$ and non-compact for $\alpha < 2$.

All those networks will be used to understand the influence of the network topology on first-passage properties: we will be able to see the difference between a tree and a network containing loops, between compact and non-compact networks, between fractal and non-fractal (but scale-invariant) networks, and so on. We will always perform nearest-neighbor random walks on such network: at each step, the random walker have an equal probability to choose any of the neighbor, namely a node linked directly to the node occupied by the random walker. The random walk is here discrete in time and space.

1.3.3 Pseudo-Green functions for self-similar RW

We pointed out several time in this chapter that the evaluation of pseudo-Green functions is crucial to obtain first-passage properties. We have already shown how to get exact expressions

of those pseudo-Green functions for regular lattice, we will here present a more general approximation, for scale-invariant random walk in confined media, either in discrete or in continuous space. Those results have been published in [66], and were part of my Master thesis.

This approximation relies on two assumptions:

(i) that we can, with a very good accuracy, take the *infinite* Green function instead of the confined pseudo-Green function to estimate first-passage properties,

(ii) that we can get those infinite Green functions using a scaling form of the propagator.

Since we will use very frequently in this manuscript the results given here, we will show that this approximation is confirmed by numerical simulations for several emblematic models.

Infinite space assumption

We consider a random walker moving in a bounded domain of size N. As previously, we note $P(\mathbf{r}, t|\mathbf{r}')$ the propagator, *i.e.* the probability density to be at site $\mathbf{r}$ at time t, starting from the site $\mathbf{r}'$ at time 0. We recall that:

$$\langle \mathbf{T} \rangle = \frac{H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_S)}{P_{\text{stat}}(\mathbf{r}_T)} \tag{1.123}$$

where H is the pseudo-Green function, defined as

$$H(\mathbf{r}|\mathbf{r}') = \int_0^\infty \left(P(\mathbf{r}, t|\mathbf{r}') - P_{\text{stat}}(\mathbf{r}) \right) dt \tag{1.124}$$

and P_{stat} is the stationary probability distribution. To go further, we have to estimate the pseudo-Green function H, which is indeed a cumbersome task, since it depends both on the walk's characteristics and on the shape of the domain. We propose to approximate H by its infinite-space limit, which is precisely the usual Green function G_0 :

$$H(\mathbf{r}|\mathbf{r}') \simeq G_0(\mathbf{r}|\mathbf{r}') = \int_0^\infty P_0(\mathbf{r}, t|\mathbf{r}')dt \tag{1.125}$$

where P_0 is the infinite space propagator. We stress that when inserted in equation (1.123), this form does not lead to a severe infinite space approximation of the MFPT, since all the dependence on the domain geometry is now contained in the factor $1/P_{\text{stat}}$. This approximation is the key step of the derivation presented here, and captures extremely well the confining effects on MFPTs in complex media.

Scaling law assumption

We first consider the case of a uniform stationary distribution $P_{\text{stat}} = 1/N$, which is realized as soon as the network is undirected and the number of connected neighbors of a node, the connectivity, is constant. This assumption does not apply to scale-free networks, which will be tackled later on. Following [71], we assume for P_0 the standard scaling:

$$P_0(\mathbf{r}, t|\mathbf{r}') \sim t^{-d_f/d_w} \Pi \left(\frac{|\mathbf{r} - \mathbf{r}'|}{t^{1/d_w}} \right) \tag{1.126}$$

where d_f is the fractal dimension ($N(r) \propto r^{d_f}$), and d_w the random walk dimension ($\langle r^2 \rangle \propto t^{2/d_w}$). This form ensures the normalization of P_0 by integration over the whole space.

We can here link compacity with d_f and d_w. We will assume that a compact random walk is equivalent to a recurrent random walk, namely a walk that has a probability 1 to come back eventually at the starting point. Equivalently, we can consider that the walk is recurrent as soon as the average number of return is infinite, and transient else [168]. The average number of return $\langle \mathbf{N}_r \rangle$ is approximately:

$$\langle \mathbf{N}_r \rangle = \int_1^\infty P_0(\mathbf{r}, t|\mathbf{r}) dt \propto \Pi(0) \int_1^\infty \frac{1}{t^{d_f/d_w}} dt \tag{1.127}$$

We start the integration at $t = 1$, since the scaling is valid only after a while. This does not affect the final result: the average number of return between $t = 0$ and $t = 1$ is necessarily finite. We deduce that if $d_f > d_w$, the average number of return is finite, the random walk is non-compact (or transient), and if $d_f < d_w$, the average number of return is infinite, the random walk is compact (or recurrent). The case $d_f = d_w$ will be called "marginal", even if formally compact.

Using the scaling 1.126, a derivation (detailed in [66]) yields as central result:

$$\langle \mathbf{T} \rangle \sim \begin{cases} N\left(A - Br^{d_w - d_f}\right) & \text{for } d_w < d_f \\ N\left(A + B\ln(r)\right) & \text{for } d_w = d_f \\ N\left(A + Br^{d_w - d_f}\right) & \text{for } d_w > d_f \end{cases} \tag{1.128}$$

for $r = |\mathbf{r}_T - \mathbf{r}_S|$ different from 0. Strikingly, the constants A and B do not depend on the confining domain. In addition, while A is related to the small scale properties of the walk, we emphasize that B can be written solely in terms of the infinite space scaling function Π.

For instance for non-compact exploration ($d_f > d_w$), we have:

$$\begin{cases} A = \displaystyle\int_0^\infty P_0(0, t) dt \\ B = \displaystyle\int_0^\infty \frac{\Pi(u^{-1/d_w})}{u^{d_f/d_w}} du \end{cases} \tag{1.129}$$

For compact exploration, we have to introduce a constant C_0, the limit for $x \to 0$ of $\Pi(x)$, and define Π^* as:

$$\Pi(x) = C_0 - \Pi^*(x) \tag{1.130}$$

We then get (see [66] for details):

$$\begin{cases} A = \displaystyle\lim_{R\to\infty} \left[\int_0^\infty \left(P_0(0, t) - P_0(R, t)\right) dt - BR^{d_w - d_f} \right] \\ B = \displaystyle\int_0^\infty \frac{\Pi^*(u^{-1/d_w})}{u^{d_f/d_w}} du \end{cases} \tag{1.131}$$

A similar result exists for the $d_w = d_f$ case.

These expressions show a very general scaling of the MFPT dependence on the geometrical parameters N and r. Some comments are in order:

(i) First, we point out that equation (1.128) gives the large N asymptotic of the MFPT being a function of N and r as independent variables.

(ii) Second, the equation (1.128) shows two regimes, which rely on infinite space properties of the walk: in the case of compact exploration ($d_w \geq d_f$) where each site is eventually visited, the MFPT behaves like $\langle \mathbf{T} \rangle \propto Nr^{d_w - d_f}$ ($\langle \mathbf{T} \rangle \propto N\ln(r)$ for $d_w = d_f$) at large distances, so that the dependence on the starting point always matters; in the opposite case of non-compact exploration, $\langle \mathbf{T} \rangle$ tends to a finite value for large r, and the dependence on the starting point is lost.

Numerical simulations

We now confirm these results by Monte Carlo simulations and exact enumeration methods applied to various models which exemplify the three possible cases ($d_w > d_f$, $d_w = d_f$ and $d_w < d_f$).

(i) The random barrier model in dimension $d = 2$ is defined by a lattice random walk with nearest neighbors symmetrical transition rates Γ distributed according to some distribution $\rho(\Gamma)$. Even for a power law distribution $\rho(\Gamma)$ the scaling function $\Pi(\xi)$ can be shown to be Gaussian ($d_f = d_w = 2$), which allows us to explicitly compute the constant B and obtain

$$\langle \mathbf{T} \rangle \sim N \left(A + \frac{1}{2\pi D_{\text{eff}}} \ln(r) \right). \tag{1.132}$$

Here D_{eff} is a diffusion constant depending on $\rho(\Gamma)$ which can be determined by a mean field approximation.

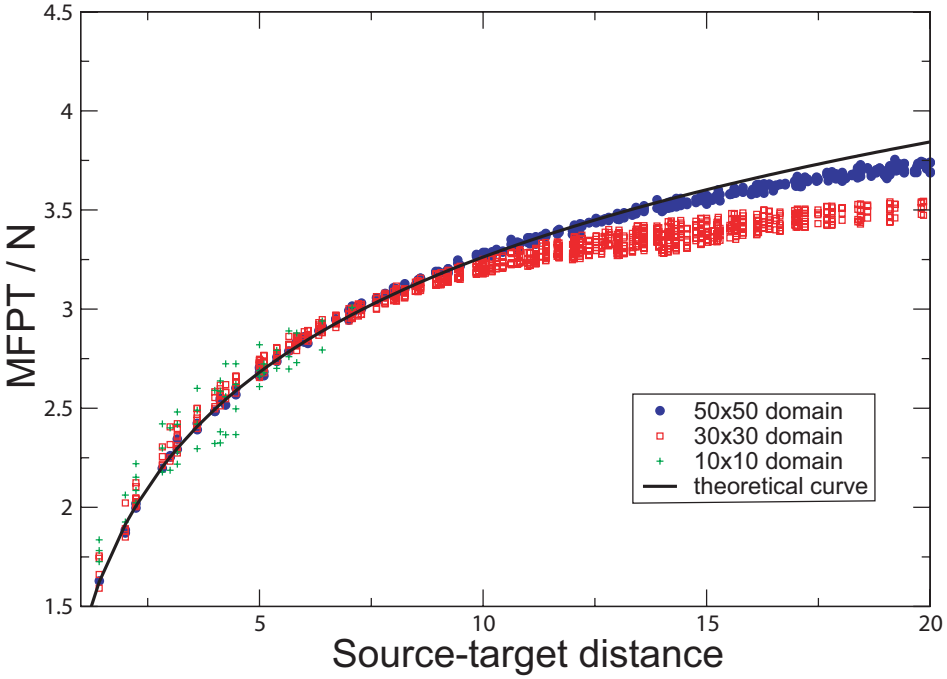

Figure 1.22: MFPT as a function of the source-target distance for random barrier model. Here, the transition rate distribution is $\rho(\Gamma) = (\alpha/\Gamma)(\Gamma/\Gamma_0)^\alpha$, with $\Gamma_0 = 1$ and $\alpha = 0.5$. The confining domain is an $L \times L$ square with the target in the middle. Three different domain sizes are shown, with numerical simulations of the MFPT rescaled by the volume N, averaged over the disorder. The theoretical curve (black line) is given by equation (1.132), where A is a fitting parameter.

(ii) The Sierpinski gasket of finite order, defined in the last sub-section, is a representative example of deterministic fractals. In this case $d_f = \ln(3)/\ln(2) < \ln(5)/\ln(2) = d_w$, so that

our theory predicts the scaling

$$\langle \mathbf{T} \rangle \sim N r^{(\ln(5)-\ln(3))/\ln(2)}. \tag{1.133}$$

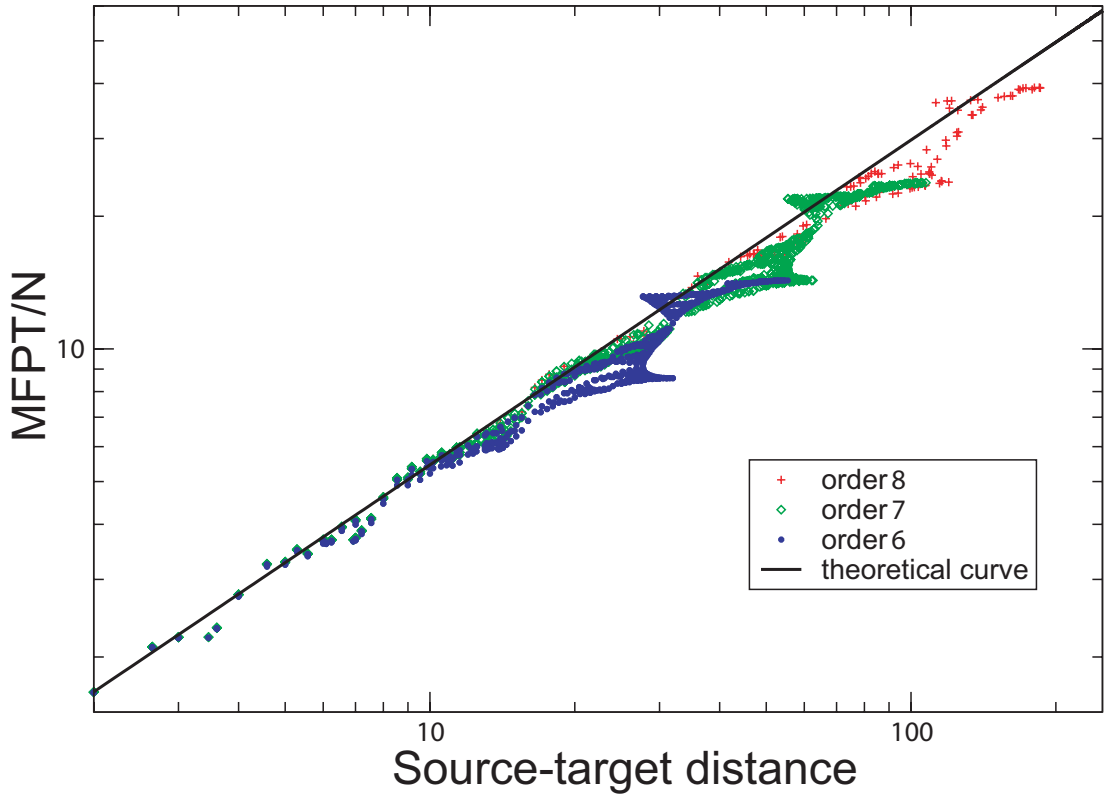

Figure 1.23: MFPT as a function of the source-target distance for finite Sierpinski gasket (log/log plot) for three different system sizes (orders 6, 7 and 8). For each set of points, the size of the Sierpinski gasket and the target point are fixed, and the starting point takes various positions on the network. The black line corresponds to the theoretical scaling $r^{d_w - d_f}$.

(iii) The Lévy flight model, defined previously, is based on a heavy-tailed distribution of jump lengths $p(l) \propto l^{-d-\beta}$ $(0 < \beta \leq 2)$. The walk dimension is now $d_w = \beta$, while the fractal dimension is the dimension of the Euclidian space d. In dimensions $d \geq 2$, or for $d = 1$ when $\beta < 1$, one has $d_f > d_w$ and our theory gives

$$\langle \mathbf{T} \rangle \sim N \left(A - B r^{\beta - d} \right). \tag{1.134}$$

Figures 1.22, 1.23 and 1.24 show that both the volume dependence and the source-target distance dependence are unambiguously captured by our theoretical expressions (1.128), as shown by the data collapse of the numerical simulations. We emphasize that the very different nature of these examples demonstrates that the range of applicability of our approach, which mainly relies on the length scale invariant property of the infinite space propagator (1.126), is wide.

The case of non uniform degree distribution will studied in detail in chapter 3.

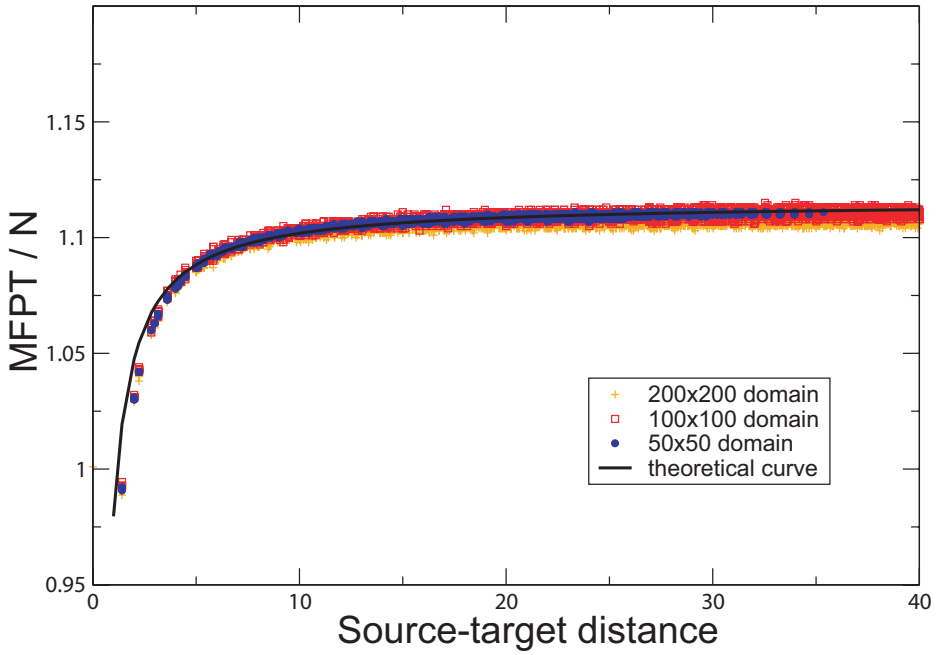

Figure 1.24: MFPT as a function of the source-target distance for Lévy flights on a square ($d = 2$), for three different system size (50×50, 100×100 and 200×200), with $\beta = 1$. The target is in the middle of the network. Simulation points are fitted with $\langle \mathbf{T} \rangle / N \sim \left(A - Br^{\beta - d}\right)$ (black curve).

Quick summary

- We have introduced anomalous diffusion processes, as well as discrete networks leading to anomalous diffusion;
- we have shown how to approximate pseudo-Green functions for scale-invariant processes.

Conclusion

During this first chapter, we have presented very basically what random walks and diffusion were, as well as the main mathematical tools that will be used during this manuscript. Those definitions will not always be recalled later, we will now assume that the reader is familiar with those concepts.

We have introduced many first-passage observables: depending on the problem, the relevant observable is not always the same. We will often focus on the mean first passage time (MFPT) because it is an easy way to estimate transport properties, but several others will be used along this manuscript. The key point of this long introduction has been to show how crucial the estimation of pseudo-Green functions was: once those functions are known, we can extract many first-passage observables using *exact* expressions.

At last, we presented the concept of anomalous diffusion, as well as some prominent continuous processes leading to such behavior, and some scale-invariant networks also sharing to the same feature. Those continuous processes and networks will be used several times during this manuscript without any further definition: the reader will have to come back in chapter 1 for details.

My Master thesis ended with the publication of [66], detailed in the last sub-section of this chapter. The possibility to estimate quickly and efficiently pseudo-Green functions is a powerful tool, that works with our assumptions for Markovian random walkers on regular scale-invariant networks, and the related continuous processes. We will during this manuscript try to extend this result for a wider range of processes: to anomalous diffusion (non Markovian random walks, experimental trajectories and related models) and to non regular networks (with different connectivities, persistence, moving targets and broken links). Each axe will form a chapter of this manuscript: in chapter 2 we will focus on anomalous diffusion, and in chapter 3 in first-passage properties on non-regular networks. To do so, we will keep the infinite space approximation (taking infinite Green functions instead of confined pseudo-Green function), and we will have to adapt the approximation concerning the scaling form of the propagator.

2 From experimental trajectories to (anomalous) diffusion mechanism, and the way back

Contents

In this chapter, we will focus on a feature often observed in experimental trajectories: anomalous diffusion. We will successively introduce a theoretical method to distinguish two sub-diffusive microscopical mechanisms using first-passage properties, then adapt it to effectively be able to analyze experimental trajectories, and at last introduce new models that reproduce some of the features observed during our analysis of experimental trajectories.

We will first introduce the purpose of our analysis, namely: why would we be interested in analyzing experimental trajectories? The first section will then present the initial approach we had when facing this challenge, mainly based on first-passage observables introduced in chapter 1. We will present the theoretical framework, and we will apply it to diffusion on fractals and CTRW. We developed an experimental set-up based on this approach, and shared it with experimentalists. After confronting our theory to experimental data, we had to change a bit the attack angle. Indeed the usual biological data sets were too small to allow us to extract with a good level of confidence the first-passage observables we were working with.

In the second section we will present other observables more adapted to small data sets, namely the displacement and maximal excursion moments, their ratios, and the growing shell test. This method will be tested on a "small" simulated data set, namely on 1,000 trajectories of 100 steps. This is already big for a biologist, but very small compared to the usual simulation data sets ($\sim$ 100,000 trajectories of 1,000 steps). We will analyze experimental data to test our method: simple Brownian diffusion of Quantum dots, and lipid granules data diffusing within *S. Pombe*. For the last data set, we observed several sub-diffusive regimes at different time scales, but none of the usual sub-diffusive models (CTRW, diffusion on fractals or fBm, introduced in chapter 1) could explain all the features observed.

In the last section, we will go back to theory, and take some prominent characteristics of the trajectories observed during the lipid granule analysis to focus on more adapted diffusive models. We will thus start from experiments to construct a meaningful model, from the experimental point of view, that can explained the observed characteristics. We will successively look at persistent random walks, with a very short memory (one step), then at random walks in an ordered porous media, and at last at random walks in a crowding environment.

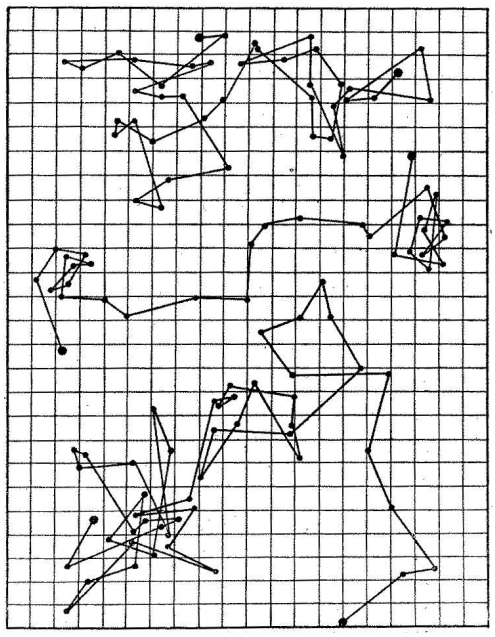

Figure 2.1: Three trajectories obtained by tracing a small grain of putty at intervals of 30 s. Reproduced from reference [157].

Introduction

With modern experimental tools single trajectories of particles down to the nanoscale can be obtained, even in complex systems such as biological cells [99, 191]. Experimental set-ups exist to obtain single trajectories of proteins, RNA or very small particles like quantum dots. What can we learn from those trajectories?

This idea to systematically measure individual trajectories of particles in order to obtain information about their ensemble behavior essentially goes back to Einstein's probabilistic interpretation of diffusion [82]. It was put to much use in the determination of the Avogadro–Loschmidt number in the beginning of the 20th century. In fact single trajectories of small granules in uniform gamboge emulsions obtained by fractional centrifuging were recorded and analyzed quantitatively by Jean Perrin in his seminal work on the deduction of Avogadros constant [157]. A few sample trajectories from Perrin's work are reproduced in Figure 2.1.

Those trajectories can be used to characterize the microscopical diffusion mechanism, or if we assume a certain diffusion mechanism, to obtain physical constants like diffusion coefficients. To get an accurate result, one has either to have a lot of trajectories, or very long ones, in order to get a meaningful ensemble average or time average.

It was Nordlund who first devised a set-up to obtain long time-resolved trajectories of single particles using a moving photographic film [149]. An example for such a trajectory of a small mercury sphere slowly falling in water is shown in Figure 2.2. The Brownian motion coupled with the drift motion is clearly visible: the stochastic motion superimposed on the drift produces the wave-like behavior.

For each trajectory Nordlund then plotted the mean square displacement and fitted the diffusion constant. For one single trajectory the result is reproduced in Figure 2.3. The unit of time on the horizontal axis is 1.481 second. Nordlund's result for the Avogadro number was remarkable: 5.91×10^{23}, within 2 % of the best current value. This is one of the first examples

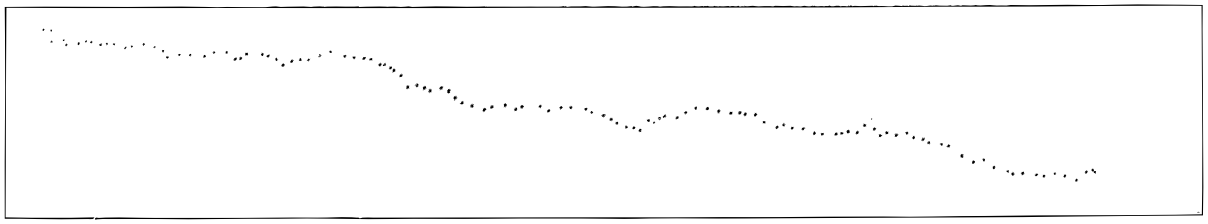

Figure 2.2: Stroboscopic trajectory of a small mercury sphere slowly falling in water, reproduced from [149].

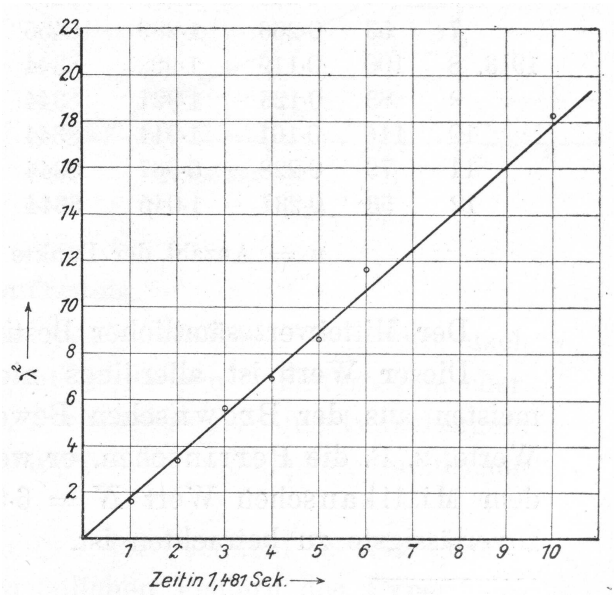

Figure 2.3: Mean square displacement obtained from a single trajectory by Nordlund [149].

of time averaging in single particle trajectories.

For a long time, it has been assumed that the diffusion of very small particles was Brownian. The only observable was then just the diffusion coefficient D, that can be obtained through the mean square displacement. One can perform:

(i) an ensemble average over a large set of single trajectories:

$$\langle r^2 \rangle(t) = \int_{-\infty}^{\infty} \|\mathbf{r}\|^2 P(\mathbf{r}, t) d\mathbf{r} = Dt, \tag{2.1}$$

where $\langle r^2 \rangle$ is the mean square displacement, D the diffusion coefficient, and $P(\mathbf{r}, t)$ the probability to be at position $\mathbf{r}$ at time t, starting from $\mathbf{r} = \mathbf{0}$ at time $t = 0$. In practice one often has only a finite set of trajectories, this integral becomes a simple sum:

$$\langle r^2 \rangle(t) = \frac{\sum_i \|\mathbf{r}_i(t)\|^2}{\sum_i 1} \simeq Dt, \tag{2.2}$$

where $\mathbf{r}_i(t)$ is the position in the ith trajectory of the random walker at time t, with $\forall i,\ \mathbf{r}_i(t = 0) = \mathbf{0}$.

Such analysis exploit only a small part of the available information in single trajectories: for a given time span, we only extract one r^2 value from each trajectory.

(ii) a time averaged mean square displacement, in order to increase the number of r^2 value we extract from one trajectory, at least for a time span small compared to the trajectory length:

$$\langle r^2 \rangle_t(\Delta, T) = \frac{\int_0^{T-\Delta} \|\mathbf{r}(t'+\Delta) - \mathbf{r}(t')\|^2 \, dt'}{T-\Delta}, \tag{2.3}$$

where $\mathbf{r}(t)$ is the random walker position at time t, T the trajectory length, and $\Delta < T$ the time span. We recall that $\langle \ldots \rangle_t$ is here a time average.

We can after this time average perform an ensemble average over all trajectories:

$$\left\langle \langle r^2 \rangle_t(\Delta, T) \right\rangle = \left\langle \frac{\int_0^{T-\Delta} \|\mathbf{r}(t'+\Delta) - \mathbf{r}(t')\|^2 \, dt'}{T-\Delta} \right\rangle = D\Delta \tag{2.4}$$

As previously, the integral as well as the ensemble average become summations in the analysis of a finite set of finite trajectories.

Using the time-average/ensemble-average technique allows one to obtain accurate estimates of the diffusion coefficient, at least for Brownian motion.

What about anomalous diffusion? Sub-diffusive single walker trajectories are indeed often observed: in the last few years, sub-diffusion has been observed in an increasing number of systems [172, 173], ranging from physics [122, 187] or geophysics [186] to biology [99, 212]. In particular, living cells provide striking examples for systems where sub-diffusion has been repeatedly observed experimentally, either in the cytoplasm [54, 99, 212, 232], the nucleus [160, 223] or the plasmic membrane [94, 127, 199]. However, the microscopical origin of sub-diffusion in cells remains debated, even if believed to be due to crowding effects in a wide sense as indicated by *in vitro experiments* [8, 9, 98, 228].

As explained in chapter 1, the sub-diffusive behavior deviates from the usual Brownian motion, and is usually characterized by a mean square displacement (MSD) which scales as $\langle r^2 \rangle \propto t^\alpha$ with $\alpha < 1$. α is here the anomalous diffusion exponent. How can we properly estimate this α exponent, and more importantly how can we determine which sub-diffusive models (all leading to the same α) explain the best the trajectory features?

Questions to be answered

- How can we extract a proper anomalous diffusion exponent α from experimental trajectories?
- For a given α, how can we determine the sub-diffusive microscopical mechanism?

2.1 First theoretical predictions: beyond the MSD

Anomalous diffusion can be obtained from a few models based on different underlying microscopical mechanisms. We presented in the first chapter several models, such as continuous time

random walk (CTRW) [115, 172], fractional Brownian motion (fBm) [137] or diffusion on fractals [73]. All those models lead to a mean square displacement $\langle r^2 \rangle$ scaling like t^α. If we only use the MSD of a given trajectory set, we will extract the sub-diffusion coefficient α, but we will not be able to determine what kind of microscopical model cause the sub-diffusion. We will present in this section the results published in reference [68], and see how first-passage observables introduced in the first chapter can give a finer information than the MSD. We will focus in this section solely on two possibilities: CTRW and diffusion on fractals. The analytical treatment presented here cannot be extended to fBm, as explained below. We will see in the next section how to deal with fBm.

CTRW and diffusion on fractals are relevant to describe anomalous diffusion in confined systems as cytoplasm or membrane of living cells. The cell is known to be a highly complex and inhomogeneous molecular assembly, composed of numerous constituents which may widely vary from one cell type to another. Here we wish to distinguish between two types of effects on transport in cellular medium. First, the overall density of free proteins and molecular aggregates is very high, in the cytoplasm as well as in the plasma membrane. In such crowded environment, a tracer particle is trapped in dynamical "cages" whose life times are broadly distributed at high densities, leading to an heavy tailed distribution of waiting times. This dynamical picture therefore fits the hypothesis of the CTRW model. Second, the cytoskeleton is made of semi-flexible polymeric filaments (such as F–actin or microtubules), which can be branched and cross-linked by proteins. This scaffold therefore acts as fixed obstacles constraining the motion of the tracer. Moreover, the cytoplasm can be compartmentalized by lipid membranes which further constrain the tracer. Such environment with obstacles can be described in a first approximation by a static percolation cluster.

While these two models lead to similar scaling laws for the MSD, their microscopical origins are intrinsically different and lead to notable differences in other transport properties. This has strong implications, in particular on transport limited reactions [131], which will prove to have very different kinetics in the two situations. As most of functions of a living cell are regulated by coordinated chemical reactions which involve low concentrations of reactants (such as transcription factors or vesicles carrying targeted proteins [5]), and which are limited by transport, understanding the origin of anomalous transport in cells and its impact on reaction kinetics is an important issue.

We will present – and analytically calculate – transport related observables, based on first-passage properties already introduced in the first chapter, which allow to discriminate between the CTRW and fractal models, and permit a quantitative analysis of the kinetics of transport limited reactions. We briefly recall the definition of those observables introduced in chapter 1:

- The first-passage time (FPT), which is the time needed for a particle starting from site S to reach a target T for the first time. We will focus on both the probability density function (PDF) of the FPT, and on its first moment, the mean FPT (MFPT).
- The splitting probability, which is the probability to reach a target T_i before reaching any other target $T_{j\neq i}$, in the case where several targets are available.
- The occupation time before reaction, which is the time spent by a particle at a given site T_1 before reaction with a target T_2. We will be interested in both the entire PDF of the occupation time, and the mean occupation time.

We will thus evaluate theoretically non trivial first-passage characteristics of transport in disordered media in any dimensions, while so far mainly effective one-dimensional geometries have

been investigated [168]. In particular we calculate here the MFPT, splitting probabilities and occupation time distribution of a random walk on percolation clusters, and discuss the potential implications of these results on reactions kinetics in living cells.

We will in the first sub-section set the theoretical framework and give explicit analytical expressions of the first-passage observables, which are summarized in equations (2.16), (2.17) and (2.18). We will present our formalism, based, as in the first chapter, on pseudo-Green functions and on their approximation. In the next two sub-section, we will apply this formalism to the model of diffusion on fractals, which fulfills our assumption, and to the CTRW model, which is non-Markovian. We will thus show in the third sub-section how we can deal with non-Markovian random walks, with the example of CTRW. In the fourth sub-section, we will suggest experiments which could help to discriminate the microscopical origin of sub-diffusion, using the results obtained for diffusion on fractals and CTRW. We will present a detailed experimental set-up, and we will see how realistic our approach is.

2.1.1 Theoretical framework

General formalism

Using techniques presented in the first chapter [64, 65, 66], we derive general analytical expressions of the first-passage observables using pseudo-Green functions. We consider a Markovian random walker moving in a bounded domain of N sites, with reflecting walls.

Let $P(\mathbf{r}, t|\mathbf{r}')$ be the propagator, *i.e.* the probability density to be at site $\mathbf{r}$ at time t, starting from the site $\mathbf{r}'$ at time $t = 0$, whose evolution is described by a master equation [216]

$$\frac{\partial P}{\partial t} = \mathcal{L}P \tag{2.5}$$

with a given transition operator $\mathcal{L}$. Fractional Brownian motion does not follow such equation, local in time: we have to take into account the whole trajectory between $-\infty$ and t to determine what happens at $t + dt$. This is why we will here focus on CTRW and diffusion on fractal, before seing in the following section how to deal with fBm.

For the sake of simplicity we assume that the walker performs symmetric jumps and that the stationary distribution is homogeneous

$$\lim_{t\to\infty} P(\mathbf{r}, t|\mathbf{r}') = P_{\text{stat}} = \frac{1}{N}. \tag{2.6}$$

As shown in the first chapter, the renewal equation gives an exact expression for the MFPT, provided it is finite:

$$\langle \mathbf{T} \rangle = N \left(H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_S) \right), \tag{2.7}$$

where H is the pseudo-Green function [14] of the domain :

$$H(\mathbf{r}|\mathbf{r}') = \int_0^\infty \left(P(\mathbf{r}, t|\mathbf{r}') - \frac{1}{N} \right) dt. \tag{2.8}$$

Two targets problem

We can also obtain, using pseudo-Green functions, splitting probabilities P_i as well as mean time $\langle \mathbf{T} \rangle$ to hit any of the targets, assuming that the random walker can be absorbed either by a target T_1 at $\mathbf{r}_1$, or a target T_2 at $\mathbf{r}_2$:

$$\frac{\langle \mathbf{T} \rangle}{N} = P_1 H(\mathbf{r}_1|\mathbf{r}_1) + P_2 H(\mathbf{r}_1|\mathbf{r}_2) - H(\mathbf{r}_1|\mathbf{r}_S). \tag{2.9}$$

Using the approach of chapter 1, we can use this equation together with the similar equation obtained by inverting 1 and 2, and the condition $P_1 + P_2 = 1$, to obtain a linear system of 3 equations for the 3 unknowns P_1, P_2, and $\langle \mathbf{T} \rangle$, which can therefore be straightforwardly determined. In particular the splitting probability P_1 reads:

$$P_1 = \frac{H_{1S} + H_{22} - H_{2S} - H_{12}}{H_{11} + H_{22} - 2H_{12}}, \tag{2.10}$$

where we used the notation $H_{ij} = H(\mathbf{r}_i|\mathbf{r}_j)$. This formula extends the result presented in chapter 1, obtained for simple random walks [64, 65], to the case of general Markov processes.

As shown in the first chapter, the splitting probabilities allow us to obtain the entire distribution of the occupation time [67] $\mathbf{N}_i$ at site i for general Markov processes:

$$P(\mathbf{N}_i = k) = \begin{cases} 1 - E_1 & \text{if } k = 0 \\ E_1 E_2 \left(1 - E_2\right)^{k-1} & \text{else} \end{cases} \tag{2.11}$$

where

$$E_1 = \frac{H_{iS} + H_{TT} - H_{ST} - H_{iT}}{H_{ii} + H_{TT} - 2H_{iT}}, \tag{2.12}$$

and

$$E_2 = \frac{1}{H_{ii} + H_{TT} - 2H_{iT}}. \tag{2.13}$$

In particular, the mean occupation time is then given by

$$\langle \mathbf{N}_i \rangle = H_{iS} - H_{iT} + H_{TT} - H_{ST}. \tag{2.14}$$

We stress that equation (2.11) gives the exact distribution of the occupation time for all regimes. It follows in particular that the large time asymptotic of the occupation time distribution is exponential. Actually one can argue in the general case that the FPT is also exponentially distributed at long times. This comes from the fact that the transition operator $\mathcal{L}$ has a strictly negative discrete spectrum for a finite volume N (see [216]). This exponential decay has been confirmed using another approach for a general scale-invariant random walk [151].

Large volume approximation

Equations (2.7), (2.10) and (2.14) give exact expressions of the first-passage observables as functions of the pseudo-Green function H. The key point is then, as shown in the first chapter [66], to estimate H. We can here assume that the problem is scale invariant and we use for P_0 the standard scaling [71] that we already used in chapter 1 to approximate the pseudo-Green functions:

$$P_0(\mathbf{r}, t|\mathbf{r}') \propto t^{-d_f/d_w} \Pi\left(\frac{\|\mathbf{r} - \mathbf{r}'\|}{t^{1/d_w}}\right), \tag{2.15}$$

where the fractal dimension d_f characterizes the accessible volume $V_r \propto r_f^d$ within a sphere of radius r, and the walk dimension d_w characterizes the distance $r \propto t^{1/d_w}$ covered by a random walker in a given time t. The form (2.15) ensures the normalization of P_0 by integration over the whole fractal set. Note that the mean square displacement is then given by $\langle r^2 \rangle \propto t^\alpha$ with $\alpha = 2/d_w$.

The derivation presented in the first chapter [66] yields for the MFPT:

$$\langle \mathbf{T} \rangle \sim \begin{cases} N(A - Br^{d_w - d_f}) & \text{for } d_w < d_f \\ N(A + B \ln r) & \text{for } d_w = d_f \\ BNr^{d_w - d_f} & \text{for } d_w > d_f \end{cases}, \tag{2.16}$$

where explicit expressions of A and B have already been given.

In fact, the above analysis of the pseudo-Green functions also permits to obtain explicit expressions of the splitting probabilities and mean occupation times:

$$P_1 \sim \begin{cases} \dfrac{A + B(r_{1S}^{d_w-d_f} - r_{2S}^{d_w-d_f} - r_{12}^{d_w-d_f})}{2(A - Br_{12}^{d_w-d_f})} & \text{for } d_w < d_f \\ \dfrac{A + B\ln\left(\dfrac{r_{2S}r_{12}}{r_{1S}}\right)}{2(A + B\ln(r_{12}))} & \text{for } d_w = d_f \\ \dfrac{1}{2}\left(\left(\dfrac{r_{2S}}{r_{12}}\right)^{d_w-d_f} - \left(\dfrac{r_{1S}}{r_{12}}\right)^{d_w-d_f} + 1\right) & \text{for } d_w > d_f \end{cases} \tag{2.17}$$

and

$$\langle \mathbf{N}_i \rangle \sim \begin{cases} A + B(r_{iS}^{d_w-d_f} - r_{iT}^{d_w-d_f} - r_{ST}^{d_w-d_f}) & \text{for } d_w < d_f \\ A + B\ln\left(\dfrac{r_{iT}r_{ST}}{r_{iS}}\right) & \text{for } d_w = d_f \\ B\left(r_{iT}^{d_w-d_f} + r_{ST}^{d_w-d_f} - r_{iS}^{d_w-d_f}\right) & \text{for } d_w > d_f \end{cases}, \tag{2.18}$$

where $r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$ is different from 0. Note that the entire distribution of $\mathbf{N}_i$ is obtained similarly by estimating E_1 and E_2 as defined by equations (2.12) and (2.13).

We recall that the constants A and B do not depend on the confining domain and can be written solely in terms of the infinite space scaling function Π of equation (2.15). We point out that in the case of compact exploration ($d_w > d_f$) the expression of splitting probabilities is fully explicit and does not depend on Π.

Equations (2.16), (2.17) and (2.18) therefore elucidate the dependence of the first-passage observables on the geometric parameters of the problem, and constitute the central theoretical result of this section.

2.1.2 Diffusion on fractals

The general formalism developed above applies to any kind of Markovian self-similar random walk. Diffusion on fractals is an example, critical percolation clusters being a representative example of random fractals [45, 71, 102].

Critical percolation cluster

Here we consider the case of bond percolation on a Euclidian lattice of d–dimensions. As presented in the first chapter, when p is above the percolation threshold p_c, an infinite cluster exists. In 3D, $p = p_c$, $d_f \simeq 2.58\ldots$ and $d_w \simeq 3.88\ldots$ [45]. Those value slightly differ from those presented in chapter 1, due to the method of computation used, but $d_w - d_f = 1.3\ldots$ is still the same. The motion is sub-diffusive with $\alpha = 2/d_w \simeq 0.51$. For a given critical percolation cluster, namely for a given configuration of the disorder, the theoretical development of previous paragraph holds, and the first-passage observables are given by the exact expressions (2.7), (2.10) and (2.14). However, the variations between different realizations of the disorder have to be taken into account, and averaging has to be performed in order to obtain meaningful quantities: the scaling form of the propagator (2.15) is only valid after a disorder averaging. We thus have to check that expressions (2.7), (2.10) and (2.14) still hold after this disorder averaging.

Disorder averaging

We note by $\bar{X}$ the average of X over the disorder, and assume that all configurations have the same volume N, which is a non restrictive condition in the large N limit since N is self-averaging. Equations (2.7), (2.10) and (2.14) then show that averaging the first-passage observables is equivalent to averaging the pseudo-Green function, and therefore to the propagator in virtue of (2.8). In the case of a random walk on a critical percolation cluster it has been shown that the propagator has a multifractal behavior [45].

This means that the propagator $P(r,t)$ has a very broad distribution, and is not self-averaging: its typical value is not its average value, which is dominated by rare events. In particular a scaling form of the averaged propagator is not available. However, this difficulty can be by-passed if one considers the chemical distance x, *i.e.* the length (in step number) of the shortest path between two points. Indeed in the chemical space, the propagator does have a simple fractal scaling [45, 102] and in the infinite volume limit the averaged propagator $\overline{P}_0(x,t)$ satisfies the scaling form (2.15) (see [45]). Note that this property is shared by most of random fractals [45], and makes the chemical distance a powerful tool to calculate disorder averages. The formalism derived for self-similar Markovian random walks can therefore be employed, and the scaling laws of the MFPT, splitting probability and mean occupation time *averaged over the disorder* are given in chemical space by equations (2.16), (2.17) and (2.18), where r is to be replaced by the chemical distance x. Note that in the chemical space, the fractal dimension is given by $d_f^c = d_f/d_{\min}$ and walk dimension is $d_w^c = d_w/d_{\min}$. Exponent c indicate that we use chemical distances. The dimension $d_{\min}$ is the fractal dimension of chemical paths and permits to recover the dependence on the euclidian distance r through the scaling [71] $x \propto r^{d_{\min}}$, with $d_{\min} \simeq 1.24\ldots$ in the case of the 3D Euclidian lattice [71].

Simulations

Simulations of first-passage observables are shown in Figure 2.4. All of the embedding domains have reflecting boundary conditions, and we averaged for each set of chemical distances $\{x_{ij}\}$ the desired observable over all configurations of source and targets yielding the same set $\{x_{ij}\}$.

Inset a) of Figure 2.4 is the MFPT for random walks on 3-dimensional critical percolation clusters. For each size of the confining domain (20^3, 25^3 and 30^3), the MFPT, normalized by the number of sites N, is averaged for a given chemical distance over both the different target and starting points, and over several critical percolation clusters. The black plain curve corresponds to the prediction of equation (2.16) with $d_w^c - d_f^c \simeq 1$.

Inset b) of Figure 2.4 shows the splitting probability for random walks on 3-dimensional critical percolation clusters. The splitting probability P_1 to reach the target T_1 before the target T_2 is averaged both over the different target points T_2 and over the percolation clusters. The chemical distance $\|\mathbf{r}_S - \mathbf{r}_{T_1}\| = 10$ is fixed whereas the chemical distance $\|\mathbf{r}_S - \mathbf{r}_{T_2}\| = \|\mathbf{r}_{T_1} - \mathbf{r}_{T_2}\|$ changes. The black plain curve corresponds to the explicit theoretical equation (2.17) with $d_w^c - d_f^c \simeq 1$.

Inset c) of Figure 2.4 shows the occupation time for random walks on critical percolation clusters. For each size of confining domain, the occupation time of site T_1 before target T_2 is reached for the first time is averaged over the different target points T_2 and over the percolation clusters. The chemical distance $\|\mathbf{r}_S - \mathbf{r}_{T_1}\| = 10$ is fixed whereas the chemical distance $\|\mathbf{r}_S - \mathbf{r}_{T_2}\| = \|\mathbf{r}_{T_1} - \mathbf{r}_{T_2}\|$ changes. The black plain curve corresponds to the prediction of equation (2.18) with $d_w^c - d_f^c \simeq 1$.

At last, inset d) of Figure 2.4 shows the MFPT for random walks on percolation clusters above criticality for a 25^3 confining domain. The MFPT, normalized by the number of sites N,

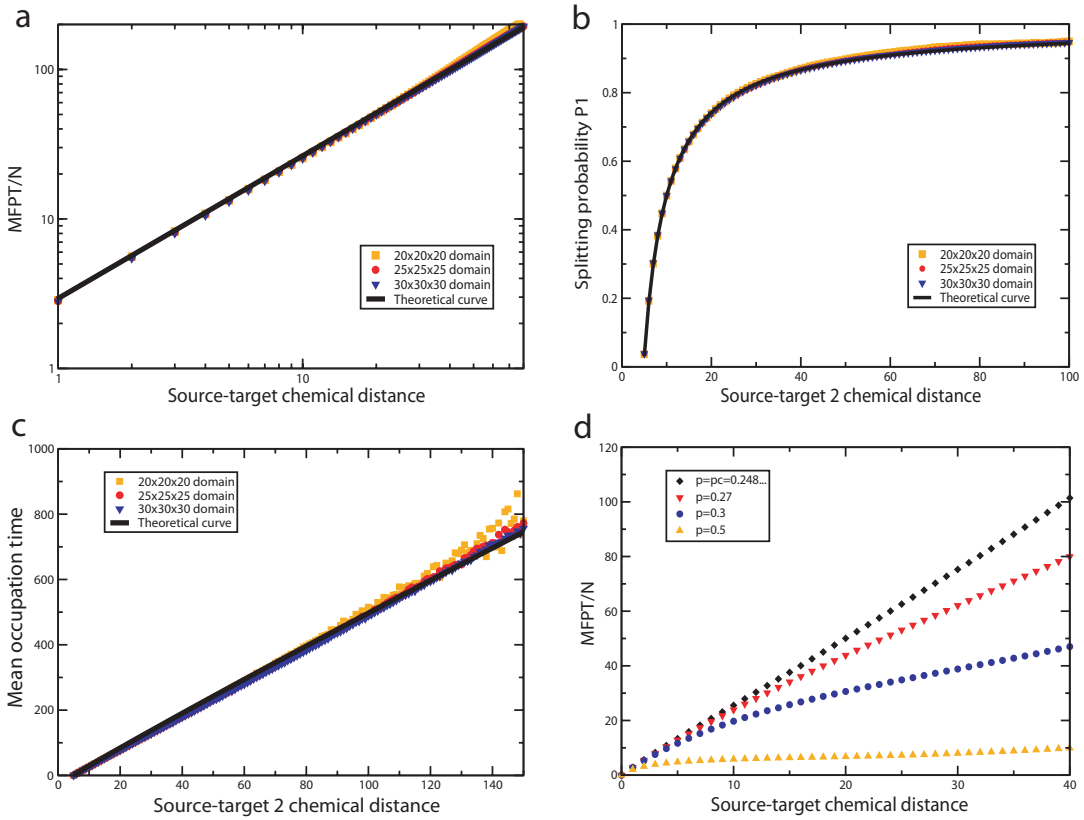

Figure 2.4: Numerical simulation of first-passage observables for random walks on three-dimensional percolation clusters. (a) MFPT for random walks on 3D critical percolation clusters. For each size of the confining domain, the MFPT, normalized by the number of size N, is averaged both over the different target and starting points separated by the corresponding chemical distance, and over the disorder. The black curve corresponds to the prediction of equation (2.16) with $d_w^c - d_f^c \simeq 1$. (b) Splitting probability for random walk on 3D critical percolation clusters. The chemical distance $ST_1 = 10$ is fixed whereas the chemical distance $ST_2 = T_1T_2$ varies. The black plain curve corresponds to the explicit theoretical expression (2.17) with $d_w^c - d_f^c \simeq 1$. (c) Occupation time for random walks on critical percolation clusters. For each size of confining domain, the occupation time of site T_1 before target T_2 is reached for the first time is averaged over the different target point T_2 and over the percolation clusters. The chemical distance $ST_1 = 10$ is fixed whereas the chemical distance $ST_2 = T_1T_2$ varies. The black plain curve corresponds to the prediction of equation (2.18) with $d_w^c - d_f^c \simeq 1$. (d) The MFPT for random walks on percolation clusters above criticality for a $25 \times 25 \times 25$ confining domain. The MFPT, normalized by the number of sites N, is averaged both over the different target and starting points separated by the corresponding chemical distance, and over the percolation clusters.

is averaged both over the different target and starting points separated by the corresponding chemical distance, and over the percolation clusters.

Figure 2.4 (a,b,c) shows that the simulations fit very well the expected scaling. Both the volume dependence and the source-target distance dependence are faithfully reproduced by our theoretical expressions, as shown by the data collapse of the numerical simulations.

If the bond concentration p is above the percolation threshold p_c, a correlation length $\xi \propto (p - p_c)^{-\nu}$ appears, where $\nu = 0.87..$ for $d = 3$. At length scales smaller than ξ, the percolation cluster is fractal, with the same fractal dimension d_f as the critical percolation cluster, and diffusion is anomalous. At length scales larger than ξ, the fractal dimension of the percolation cluster recovers the space dimension d and diffusion is normal [71].

Along the lines of the previous section, we thus expect the pseudo-Green function H to scale as $r^{d_w - d_f}$ for $r < \xi$, and as r^{d-2} for $r > \xi$. More explicitly, on the example of the MFPT we expect for the 3–dimensional cubic lattice

$$\langle \mathbf{T} \rangle \simeq \begin{cases} BNr^{1.36...} & \text{for } r < \xi \\ N\left(A' - \dfrac{B'}{r}\right) & \text{for } r > \xi \end{cases} . \tag{2.19}$$

Similarly, the other first-passage observables display a cross-over between these two regimes around ξ. The simulations do show very well the transition between the two regimes (see Figure 2.4 d)).

2.1.3 CTRW

A continuous time random walk (CTRW) is not necessarily Markovian unlike the fractal case, and therefore the above methodology cannot be applied directly. Moreover, the MFPT diverges if we use an heavy tailed distribution for the waiting time distribution. FPT distribution for CTRWs has however been obtained recently in [62]. We here briefly recall these results, and derive analytical expressions of the other observables.

Heavy tail distribution

As introduced in chapter 1, CTRW is a standard random walk with random waiting times, drawn from a distribution $\psi(t)$. We here consider heavy tailed distributions

$$\psi(t) \underset{t \gg \tau}{\sim} \frac{\alpha \tau^\alpha}{\Gamma(1-\alpha) t^{1+\alpha}} \tag{2.20}$$

The mean waiting time, $\langle \psi \rangle$, diverges for $\alpha < 1$ and the walk is sub-diffusive since the MSD scales like $\langle r^2 \rangle \propto t^\alpha$ (see [172, 187]). Here τ is a characteristic time in the process. We focus on the representative case of a Levy α-stable distribution [111], which satisfies equation (2.20) and whose Laplace transform is $\hat{\psi}(s) = \exp(-\tau^\alpha s^\alpha)$ $(0 < \alpha < 1)$.

first-passage observables for CTRW

We now derive the relation between the FPT to the site $\mathbf{r}_T$, starting from $\mathbf{r}_S$ for the standard discrete-time random walk ($\psi = 1$ for all steps) and the CTRW.

Denoting $\mathrm{FPT}(t)$ the probability density of the FPT for the CTRW, and $Q(n)$ the probability density of the FPT for the discrete-time random walk, n being the number of steps, one has

$$\mathrm{FPT}(t) = \sum_{n=1}^{\infty} Q(n)\psi_n(t), \tag{2.21}$$

where $\psi_n(t)$ is the probability that the sum of the n waiting times of the trajectory is t. We recall that each waiting time is distributed according the same distribution $\psi(t)$. $\psi_n(t)$ is here a convolution of n identical distributions $\psi(t)$, leading in Laplace space to:

$$\widehat{\psi_n}(s) = \widehat{\psi}^n(s) = \exp(-n\tau^\alpha s^\alpha) \tag{2.22}$$

Equation (2.21) can be rewritten after Laplace transformation as

$$\widehat{\mathrm{FPT}}(s) = \widehat{Q}\left(e^{-n\tau^\alpha s^\alpha}\right), \tag{2.23}$$

where

$$\widehat{Q}(z) = \sum_{n=1}^{\infty} Q(n) z^n \tag{2.24}$$

is the Z-transform of the discrete-time random walk first-passage density.

Several comments are in order.

(i) First, the small s limit shows that the MFPT is infinite, and the long-time behavior of $\mathrm{FPT}(t)$ is directly related to the MFPT of the discrete-time simple random walk:

$$\mathrm{FPT}(t) \underset{t\to\infty}{\sim} \frac{\alpha\tau^\alpha}{\Gamma(1-\alpha)t^{1+\alpha}}\langle n\rangle. \tag{2.25}$$

$\langle n\rangle$ is here the average number of steps needed to reach the target for the original discrete-time random walk, equivalent to a MFPT for this walk. It should be noted that as soon as $\widehat{Q}(z)$ is exactly known (such as for $d = 3$ in the large N limit, see [62]), the entire distribution of the FPT can be obtained.

(ii) Second, as splitting probabilities are time independent quantities, they are exactly identical for CTRW and standard discrete time random walks, and are therefore given by equation (2.17) with the space dimension d and the walk dimension $d_w = 2$.

(iii) Third, the same decomposition as equations (2.21), (2.23) holds for the distribution $\mathrm{FPT}_i(t_i)$ of the occupation time t_i of site i, where the distribution of the occupation time $F(\mathbf{N}_i)$ for the discrete-time random walk has to be introduced. This yields to

$$\mathrm{FPT}_i(t) \underset{t\to\infty}{\sim} \frac{\alpha\tau^\alpha}{\Gamma(1-\alpha)t^{1+\alpha}}\langle \mathbf{N}_i\rangle. \tag{2.26}$$

Interestingly, as $F(\mathbf{N}_i)$ is explicitly given by equation (2.11), the entire distribution of the occupation time can be derived.

We emphasize that a proper definition of the mean values of the first-passage observables (namely the MFPT and the mean occupation time) is provided by introducing a truncated

first-passage observable	CTRW model	Fractal model
FPT distribution	$\propto \frac{1}{t^{\alpha+1}}$	$\propto e^{-Ct}$
(Conditional) mean FPT	$\sim N\left(1-\frac{C}{r}\right)$	$\sim CNr^{\alpha}$
Splitting probability P_1	$\frac{1+C\left(r_{1S}^{-1}-r_{2S}^{-1}-r_{12}^{-1}\right)}{2\left(1-Cr_{12}^{-1}\right)}$	$\frac{1}{2}\left(\left(\frac{r_{2S}}{r_{12}}\right)^{\alpha}-\left(\frac{r_{1S}}{r_{12}}\right)^{\alpha}+1\right)$
(Conditional) $\langle \mathbf{N}_1 \rangle$	$1+C\left(r_{1S}^{-1}-r_{1T}^{-1}-r_{ST}^{-1}\right)$	$C\left(r_{1T}^{\alpha}+r_{ST}^{\alpha}-r_{1S}^{\alpha}\right)$

Table 2.1: Comparison of first-passage observables for CTRW and fractal models for $d = 3$.

distribution (with cut-off t_c) of waiting times in place of $\psi(t)$. As this allows to define a mean waiting time

$$\langle \psi \rangle = C \int_0^{t_c} t\psi(t)dt, \tag{2.27}$$

(where C normalizes the truncated PDF), the MFPT is then given by

$$\langle \mathbf{T} \rangle = \langle \psi \rangle \langle n \rangle, \tag{2.28}$$

and the mean occupation time reads

$$\langle t_i \rangle = \langle \psi \rangle \langle \mathbf{N}_i \rangle. \tag{2.29}$$

Note that our results show that the first-passage observables scale with the geometric parameters N and r exactly as a simple random walk. Their scaling dependence is therefore given by equations (2.16), (2.17), (2.18), where d_f is the space dimension d and $d_w = 2$.

2.1.4 How to discriminate CTRW and diffusion on fractals?

We will now use those analytical results to propose a way to distinguish between CTRW and diffusion on fractals, using the first-passage properties. We will first present our solution, then some extension of the previous results, and in the end an experimental set-up that allows to measure the proper first-passage observables.

Solution (in theory)

The first-passage observables derived earlier make it possible to distinguish between the two models of sub-diffusion, as summarized in Table 2.1. In this table, constant C has to be redefined on each panel.

We note that:

(i) The first-passage time has a finite mean and exponential tail for the fractal model, while it has an infinite mean and a power-law tail in a CTRW model. Analyzing the FPT

distribution tail therefore provides a first tool to distinguish the two models. As experiments can only find the first-passage up to a certain time, we need to use the above mentioned truncated means to define the MFPT for CTRW. In this case the scaling of the MFPT for CTRW with the source–target distance is the same as for a simple random walk, and can be distinguished from the scaling of the MFPT on random fractals. These two scalings are strikingly different for $d = 3$: the CTRW performs a non-compact exploration of space ($d_w = 2 < 3 = d$) leading to a finite limit of the MFPT at large source-target distance, while exploration is compact for a random walker on the percolation cluster ($d_w > d_f$) leading to a scaling in $r^{d_w - d_f}$ of the MFPT. We highlight that this feature could have very strong implications on reaction kinetics in cells. Indeed, in the cases where the fractal description of the cell environment is relevant, our results show that reaction times crucially depend on the source–target distance r. The biological importance of such dependence on the starting point has been recently emphasized in [121], on the example of gene colocalization. On the other hand, when the CTRW description of transport is valid, reaction times do not depend on the starting point at large distance r.

(ii) The splitting probabilities for the CTRW model and for the fractal models have different scalings with the distance between the source and the targets. As mentioned previously the difference is more pronounced for $d = 3$: the probability to reach the furthest target T_2 vanishes as $r^{-(d_w - d_f)}$ for the fractal model, r being the distance $\|\mathbf{r}_S - \mathbf{r}_{T_1}\|$ with the notations of Figure 2.5, while it tends to a constant for $d = 3$ according to the CTRW model. As discussed above, this could have important consequences for the kinetics of competitive reactions in cells.

(iii) As for the occupation time, both its distribution and the scaling of the conditional mean with the distances $\|\mathbf{r}_S - \mathbf{r}_{T_1}\|$ and $\|\mathbf{r}_S - \mathbf{r}_{T_2}\|$ can be used to distinguish between models. The advantage of the mean occupation time is that it can still discriminate between the models after averaging over initial conditions, and could therefore be used even with a concentration of tracers.

Some theoretical extensions

The presented theoretical framework can be extended to cover more realistic situations.

First, sub-diffusion could result in some systems from a combination of both the dynamical (CTRW) and static (diffusion on fractal) mechanisms. Interestingly, our approach can be adapted to study the example of CTRWs on a fractal which models such situations [37]. Indeed, the same decomposition as in equation (2.21) holds in this case and shows that the dependence of the first-passage observables (defined with truncated means if needed) on the source-target distance is exactly the same as in the case of a standard discrete-time random walk on the fractal, and therefore gives access to the dimensions d_w and d_f of the fractal. In turn, the FPT distribution tail is in this case reminiscent of the single step waiting time distribution defining the CTRW as shown by equation (2.25) (see also [37]). First-passage observables therefore permit in principle to isolate and characterize each of the CTRW and fractal mechanisms even when they are both involved simultaneously.

Second, in various systems sub-diffusion occurs over a given time scale or length scale, crossing over to the regular diffusive behavior (see [228] for instance). Both models can be adapted to capture this effect. In the fractal model the fractal structure persists up to the crossover length scale (which is the correlation length ξ in percolation clusters above criticality), and the waiting time distribution for the CTRW model has a Lévy-like decay until the crossover timescale, after

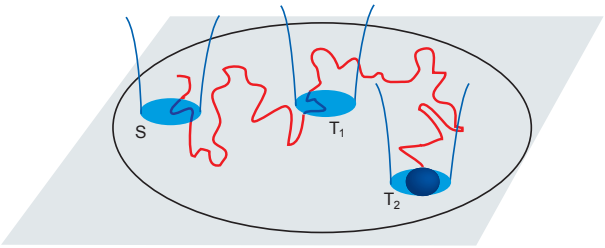

Figure 2.5: Schematic proposed set-up to measure first-passage observables.

which the decay is faster so that the mean waiting time becomes finite. The MFPT will exist in both of these modified models, but the CTRW model leads to a normal scaling of the MFPT with the volume and the source-target distance: namely , it corresponds to the results of the simple random walk, with the same time step as the mean waiting time. On the other hand, a truncated fractal structure would lead to the same scaling on larger scales, but to a scaling as in equation (2.18) at smaller scales. The small-distance behavior of the MFPT can thus discriminate the two models. The same conclusion holds for the splitting probabilities and occupation times: the small-length behavior will also differ.

Suggested (experimental) setup

We can at last present a potential experimental utilization of first-passage observables.

The schematic set-up that we propose to measure these observables relies on single particle tracking techniques (see Figure 2.5). We consider a single tracer, either a fluorescent particle or a nanocrystal, moving in a finite volume such as a living cell, a microfluidic chamber or vesicle. A laser excitation defines the starting zone S. As soon as the tracer enters S a signal is detected and a clock is started. Similarly, a second laser excitation defines the target zone T_1, and allows the measurement of the FPT of the tracer at T_1. In the same way, a third laser excitation can detect a second target T_2: counting the time spent by the tracer in T_2 before reaching T_1 gives exactly the occupation time. Splitting probabilities are straightforwardly deduced.

Experimental limitations

When presenting this idea to biologists, the setup of Figure 2.5 has been considered as rather difficult to realize.

The usual data are just direct observations through fluorescent microscopy of GFP-fused proteins. Turning on the fluorescent probe in a confined volume S is possible, but only for a limited number of probes. The easiest way to start the random walk is to consider that it begins as soon as the fluorescent probe leaves a given volume S. This volume can be defined using confocal microscopy.

After that, one has to follow the fluorescent probe, and see if it crosses a second volume T_1. Once again, a confocal microscopy can define precisely a small volume, and detect, like for the fluorescence correlation spectroscopy setup, if a random walker cross this volume of not.

With confocal microscopy, we can in the same time follow the whole trajectory, with several fluorescent probes evolving in the same time in the observation field, and define the volume S and T_1. But the accuracy will be rather bad, either in time if the space scanning is accurate, or in space if the scanning is fast. We could use two different lasers, one for the mesoscopic observation

of the whole trajectory, a second for the precise definition of S and T_1, but it becomes a complex setup.

At last, we have to stop somehow the random walk when the fluorescent probe reach the volume T_2. This can be achieved with a stronger laser exposition in this volume.

If the proposed setup could be somehow realized, three major limitations hinder to use it to discriminate between CTRW and diffusion on fractals:

(i) The random walkers can only be followed on a limited observation field. One has to take care that the confinement has to be smaller than this observation field.

(ii) The fluorescent probe can be destroyed before ever reaching T_2 through photobleaching. The amount of trajectories longer than a given time t will decay exponentially. We will thus have a bias in the statistic, short trajectories being over-represented compared to long trajectories.

(iii) The data set one can hope for such setup, even with a lot of experiments, is around $1,000$ events. This data set is to small to estimate correctly the first-passage density for instance, and discriminate between an exponential and a power-law tail.

In the end, the suggested experimental setup could be performed, but the resulting data set would not allow to discriminate properly between CTRW and diffusion on fractal. First-passage observables are good candidates to discriminate those two mechanisms, but the quantity studied so far do not exploit sufficiently the available information. Biologists work with trajectories, every single point of this trajectory should be used in the analysis, and not only the global time spent between two points, or the number of time a fluorescent probe passes through a given volume. New observables, based on first-passage properties, and using the whole trajectory, have to be developed.

Summary

- First-passage observables can (theoretically) discriminate CTRW and diffusion on fractals (even truncated or mixed) ;
- An experimental setup can be imagined to measure those observables ;
- Due to experimental constraints, other observables have to be found to exploit more completely trajectory data sets.

2.2 Further leads to face experimental constraints

As explained previously, experimental data sets usually consists on a limited number of single particle trajectories (typically between 10 and 1,000), each trajectory being between 20 and 1,000 frames long, with some missing frames. The data quantity, as well as the accuracy on the random walker position, is not adequate to use first-passage observables to a given target, at least for the time being. Indeed, the statistic we could have for the first-passage time is too limited to be confident on the extrapolated result. One should note that the simulations presented in the first chapter, or in the previous section, average between 10,000 and 100,000 trajectories for

each couple (start,target). Hoping to get such accuracy is, once again for the time being, highly unrealistic with biological experiments.

So far, we dit not show a way to discriminate a fBm with CTRW or diffusion on fractals. Other observables have been developed to tackle this problem. We will briefly introduce the normalized variance and the ergodicity breaking parameter, based on the same idea, and that we transposed to suggest our moment ratio (see below).

For fBm, a sub-diffusive process that we have introduced in the first chapter, we can define the normalized variance as:

$$V = \frac{\left\langle \left(\langle r^2 \rangle_t(\Delta, T) \right)^2 \right\rangle - \langle \langle r^2 \rangle_t(\Delta, T) \rangle^2}{\langle \langle r^2 \rangle_t(\Delta, T) \rangle^2}, \tag{2.30}$$

where $\langle \dots \rangle$ is still the ensemble average, and $\langle \dots \rangle_t$ the time average. For fBm, this normalized variance is equal to [74]

$$V \sim k(H) \begin{cases} \frac{\Delta}{T}, & \text{for } 0 < H < \frac{3}{4} \\ \frac{\Delta \ln(T)}{T}, & \text{for } H = \frac{3}{4} \\ \left(\frac{\Delta}{T}\right)^{4-4H}, & \text{for } \frac{3}{4} < H < 1 \end{cases} \tag{2.31}$$

The coefficient $k(H)$ shows a non-smooth transition at $H = 3/4$ with a divergence on reaching $H = 3/4$ both from below and above. The normalized variance tends to 0 for an ergodic process as $T \to \infty$: time average on an infinite trajectory and ensemble average over an infinite trajectory set are equal.

For CTRW where the waiting time distribution admit an α–heavy tail, we only have as theoretical result the limit when $T \to \infty$ of the normalized variance, the so-called ergodicity breaking parameter EB [104]:

$$\mathrm{EB} = \lim_{T\to\infty} V = \lim_{T\to\infty} \frac{\left\langle \left(\langle r^2 \rangle_t(\Delta, T) \right)^2 \right\rangle - \langle \langle r^2 \rangle_t(\Delta, T) \rangle^2}{\langle \langle r^2 \rangle_t(\Delta, T) \rangle^2} = \frac{2\Gamma(1+\alpha)^2}{\Gamma(1+2\alpha)} - 1 \tag{2.32}$$

This EB parameter range monotonically from 1 when $\alpha \to 0$ to 0 when $\alpha = 1$ (Brownian motion, ergodic).

Using the EB parameter and the normalized variance, we could discriminate between fBm, CTRW. The diffusion on fractals could be tackled in the same way (EB = 0 since the process is ergodic). Once again, those criteria come from theoreticians: they are not adapted to extract all the information contained in a given experimental data set.

Trying to analyze experimental trajectories, we developed an alternative way to distinguish microscopical mechanisms for anomalous diffusion, using as extensively as possible experimental data, namely single particle trajectories. All results of this section have been published in [210] (method and first analysis) and in [113] (extensive analysis).

For this section, we will consider that the ensemble averaged mean square displacement (MSD) follows:

$$\langle r^2(t) \rangle = \int_0^\infty \|\mathbf{r}\|^2 P(r,t) d\mathbf{r} = K_\alpha t^\alpha. \tag{2.33}$$

Here we assumed spherical symmetry and an isotropic environment, such that $P(r,t)$ is the probability density to find the particle a (radial) distance r away from the origin at time t after release of the particle at $r = 0$ at time $t = 0$. K_α is the generalized diffusion coefficient, equivalent of the classical diffusion coefficient D in the case of anomalous diffusion.

We will first present our method on the example of a realistic simulated data set, showing that the observables we will introduce allow to distinguish properly the three sub-diffusion mechanisms we will focus on, namely CTRW, fBm and diffusion on fractals. We will then apply this method to a real Brownian data set (Qdots freely diffusing in solution) in order to check the influence of experimental artifacts on the observables chosen. At last, we will apply our method to a sub-diffusive data set of lipid granules evolving in fission yeast (*S. Pombe*).

2.2.1 Simulations

To represent an example of realistic experimental data set, we generated several trajectories with different microscopical mechanisms. The aim was first to extract accurately the sub-diffusive exponent α, then to develop observables allowing to distinguish the mechanisms, using as extensively as possible the trajectories simulated. We decided to focus on three mechanism introduced in the previous chapter: CTRW, diffusion on fractals and fBm. These models are physically different, while they all lead to a mean square displacement of the form (2.33).

We generated time series of two-dimensional random walks by simulation of each of the three sub-diffusion models. The dimensionality mimics the fact that typically 2D trajectories are recorded in experiments. The anomalous diffusion exponent in each case was chosen as $\alpha = 0.7$, a value observed in a number of biological contexts [54, 55, 99, 190, 191, 212]. For each model we created $1,000$ trajectories, each containing 100 frames.

In the CTRW case we performed a 2D Monte Carlo random walk with a heavy tailed waiting time distribution of the form (1.105), with $\alpha \simeq 0.70$, and Gaussian jump length statistics of width $\langle \delta \mathbf{r}^2 \rangle^{1/2} = 1$.

To simulate diffusion on a fractal we generated Monte Carlo random walks on a 2D critical bond percolation cluster of size 250×250 ($d_f = 91/48$ and $d_w \approx 2.844$, so that $\alpha = 2/d_w \simeq 0.70$).

Finally, in the fBm case we used the Hosking method [109] to generate a 2D fBm with a Hurst exponent $H = 0.35$ ($\alpha = 2H = 0.70$).

We point out that fBm and CTRW are simulated in an unbounded environment, while the percolation network is, in our simulation, in a confined environment. Since we chose small time intervals in the analysis, this will not significantly affect our results.

The purpose of the simulations is to generate sub-diffusion time series of specific models. This enables us to prove that the tools developed in this section are indeed able to distinguish these mechanisms. The data sets are chosen relatively small in order to mimic typical experimental data sets.

2.2.2 Moments and ratio observables

We will introduce in this sub-section observables that differs between CTRW, fBm and diffusion on fractals. The idea is to obtain a result as simple as the one of ergodicity breaking: a value independent of the diffusion coefficient, of the time unit and of the space unit. The ratios we will introduce allow this simple analysis: to compare an "absolute" value to the one obtained for a given trajectory set. We will also introduce the maximal excursion moments, that allows a more accurate estimate of the anomalous diffusion exponent.

To get the α anomalous diffusion exponent, the usual way is to compute the ensemble-averaged mean square displacement. More generally, we can compute, for an experimental data set, the regular moments of the displacement:

$$\langle r^k(t)\rangle \simeq \frac{1}{\mathcal{N}(t)} \sum_{i=1}^{\mathcal{N}(t)} r_i^k(t), \tag{2.34}$$

where $\mathcal{N}(t)$ is the number of trajectories that are at least t frames long.

Definition of moments ratio

We will introduce in this paragraph the ratio that will be used for anomalous diffusion processes. We recall that we are looking for a quantity that does not depend on the specific observation parameter (diffusion constant, time unit or space unit), but only on the microscopical diffusion mechanism. We will first assume that we consider a classical Brownian motion, and we will show later that we can transpose the results obtained to anomalous diffusion.

If we consider an usual Brownian motion, we expect that $\langle r^k(t)\rangle \simeq 0$ if k is odd, and that $\langle r^k(t)\rangle \propto (Dt)^{k/2}$ if k is even. We can thus obtain a time-independent observable using ratio of regular moments. For instance:

$$\frac{\langle r^4(t)\rangle}{\langle r^2(t)\rangle^2} \propto \frac{D^2t^2}{(Dt)^2} = 1 \tag{2.35}$$

We can note that this ratio is a time-independent numerical value, that does not depend on the diffusion coefficient D, and that is not affected by a space rescaling. For an experimental data set, there is absolutely no fitting parameter here. It is "universal" in the sense that it only depends on the space dimension, for a Brownian motion, as shown below. Using Cauchy-Schwartz inequality, we know that this ratio is always greater than 1.

Such ratio is in fact a measure of the mean square displacement dispersion. If we use an infinite data set, we expect to have a perfectly linear mean square displacement ($\langle r^2(t)\rangle = Dt$). For a finite data set, each point $\mathbf{r}(t)$ becomes a random variable. The ensemble average is here a simple sum of random variables, and is itself a random variable, centered around Dt, but with a certain variance. This variance is directly given by:

$$\mathrm{Var}\left[r^2(t)\right] = \langle r^4(t)\rangle - \langle r^2(t)\rangle^2 \tag{2.36}$$

We can define a "relative" variance as follow:

$$\overline{\mathrm{var}}\left[r^2(t)\right] = \frac{\mathrm{Var}\left[r^2(t)\right]}{\langle r^2(t)\rangle^2} = \frac{\langle r^4(t)\rangle}{\langle r^2(t)\rangle^2} - 1 \tag{2.37}$$

The ratio $\langle r^4\rangle/\langle r^2\rangle^2$ that we have defined is thus linked to the relative variance: the greater this ratio, the more dispersed the data set.

Last subtlety: the exact ratio value we can compute, using the Brownian propagator, only holds for an infinite data set. For a finite data set, which often the case for experimental data sets, the relative variance – or equivalently the moment ratio – is itself a random variable, whom the average value is the ratio computed for an infinite data set. If we compute the ratio for a finite data set, we will thus observe that the value depends on time: the moments at two different times are two distinct random variables, with the same average value. But as time grows, we expect that the computed ratio will converge toward the predicted value. Indeed, the position

	1 D	2 D	3 D
$\langle r^2 \rangle$	Dt	Dt	Dt
$\langle r^4 \rangle$	$3\,(Dt)^2$	$2\,(Dt)^2$	$5/3\,(Dt)^2$
$\langle r^4 \rangle / \langle r^2 \rangle^2$	3	2	5/3

Table 2.2: Regular Brownian moments

$\mathbf{r}(t)$ is already a sum of t random variables (assuming the time span between two frames is 1), the random jumps: as time grows, with a given trajectory data set, we will average on more and more random jumps. The moments computed will thus converge toward the infinite ensemble average as time grows. Similarly, if we compute the ratio $\langle r^4 \rangle / \langle r^2 \rangle^2$, we expect that as time grows, this ratio converges toward the predicted value for an (infinite) ensemble average.

The moment ratio have an unique value for a Brownian motion in a given dimension. We will see below that the moment ratio takes distinct value, still time-independent, for each sub-diffusion mechanism, and will be good candidates to discriminate between those mechanisms.

Regular moments computation for Brownian motion

We can compute the moments for a regular Brownian motion [111]. To compute the second and the fourth moments we use a classical diffusion equation. In d dimensions and for spherical geometry this reads

$$\frac{\partial P(r,t)}{\partial t} = \frac{D}{2dr^{d-1}} \frac{\partial}{\partial r} \left(r^{d-1} \frac{\partial P(r,t)}{\partial r} \right). \tag{2.38}$$

D is the diffusion coefficient, $P(r,t)$ the probability to find the random walker at a distance r of the initial position at time t.

To compute the regular moments, for an arbitrary dimension d, we impose the normalization and initial conditions

$$\int_0^\infty P(r,t) r^{d-1} dr = 1 \text{ and } P(r,0) = \frac{\delta_+(r)}{r^{d-1}}, \tag{2.39}$$

where $\delta_+(r)$ is the one-sided Dirac δ function. The boundary conditions are normal, *i.e.* chosen such that $P(r \to \infty, t) = 0$.

After obtaining an expression for the propagator $P(r,t)$ we compute the nth moment

$$\langle r^n \rangle = \int_0^\infty r^n P(r,t) r^{d-1} dr. \tag{2.40}$$

Alternatively, this can be obtained by integration of r^n times the diffusion equation (2.38).

Results are summarized in Table 2.2, and the computation is detailed in Appendix 1.

We see that the ratio $\langle r^4 \rangle / \langle r^2 \rangle^2$ is a time-independent, D-independent and space unit-independent quantity, at least for a Brownian motion. It is a very good candidate to characterize precisely (without any fitting parameter) an experimental diffusion mechanism. We will see below how to adapt the calculus to transpose it to CTRW, fBm and diffusion on fractals.

Maximal excursion moments

In the course of searching different observables that could discriminate CTRW, fBm and diffusion on fractals, we looked at a way to find back first-passage properties. We saw in the previous section that first-passage properties were very good observables to distinguish CTRW and diffusion on

fractals. The problem was to conciliate first-passage properties with small data sets. We now focus on an observable related to first-passage properties, the maximal excursion, more adapted to extract as much information as possible from small data sets.

The maximal excursion $r_{\max}$ is the greatest Euclidian distance that the random walker reaches between $t = 0$ and t. This maximal excursion is illustrated in Figure 2.6: for a given time t, knowing the whole trajectory (and not only position at time t), we can compute $r_{\max}(t)$ and $r(t)$.

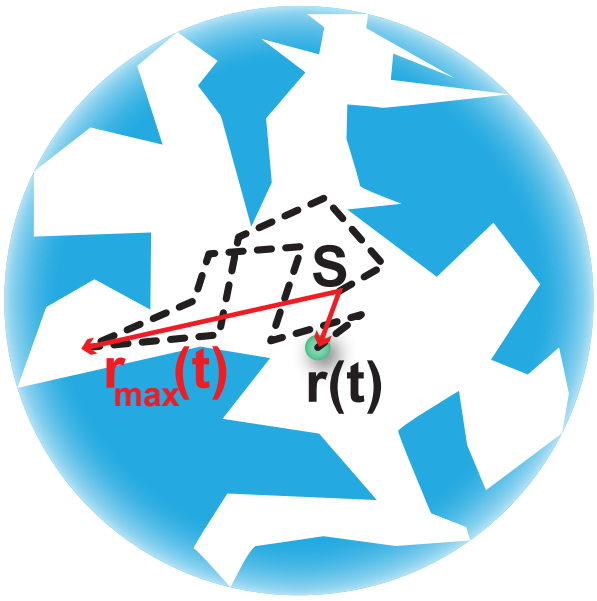

Figure 2.6: Scheme of the maximal excursion computation for a given trajectory.

As previously, we can define moments of this maximal excursion:

$$\langle r^n_{\max}(t)\rangle = \int_0^\infty r_0^n \Pr\left(r_{\max} = r_0, t\right) dr_0, \tag{2.41}$$

where $\Pr\left(r_{\max} = r_0, t\right)$ is the probability that the maximal distance from the origin that is reached up to time t, is equal to r_0.

This time, $\forall n \in \mathbb{N}$, $\langle r^n_{\max}(t)\rangle \propto (Dt)^{n/2}$. This ensemble average can be approximated, for a finite data set, by:

$$\langle r^k_{\max}(t)\rangle \simeq \frac{1}{\mathcal{N}(t)} \sum_{i=1}^{\mathcal{N}(t)} \left(\max_{0\leq t'\leq t} r_i(t')\right)^k \tag{2.42}$$

We can retrieve a time-independent ratio if we compute

$$\frac{\langle r^4_{\max}(t)\rangle}{\langle r^2_{\max}(t)\rangle^2} \propto \frac{D^2t^2}{(Dt)^2} = 1, \tag{2.43}$$

This ratio is a measure of the square maximal excursion relative variance. This ratio and the regular moment ratio do not measure the same thing: the regular moment ratio is an estimate of the position variance at a given time, this ratio is an estimate of the maximal excursion variance since the beginning of the random walk. We could hope to distinguish several mechanisms just with those two ratio: the relative variance of the position and of the maximal excursion are

	1 D	2 D	3 D
$\langle r^2_{\max}\rangle$	$1.83\ldots Dt$	$1.53\ldots Dt$	$1.40\ldots Dt$
$\langle r^4_{\max}\rangle$	$5.93\ldots (Dt)^2$	$3.49\ldots (Dt)^2$	$2.68\ldots (Dt)^2$
$\langle r^4_{\max}\rangle/\langle r^2_{\max}\rangle^2$	$1.77\ldots$	$1.48\ldots$	$1.36\ldots$

Table 2.3: Regular Brownian maximal excursion moments

expected to be different for correlated processes (like fBm and CTRW), and Brownian (uncorrelated) motion in fractal media. Indeed, for an anti-correlated process, the maximal excursion is expected to be lower than for an uncorrelated one: the probability to have several jumps in the same direction, leading to an extreme position, is lowered in the case of an anti-correlated process.

We expect that the ratios will be specific to each diffusion mechanism. We will show how to obtain those ratios, first for Brownian motion, before extending the result to CTRW and diffusion on fractals. fBm will be treated separately.

Maximal excursion moments computation for Brownian motion

To get maximal excursion moments for a Brownian motion, we first calculate the probability that at time t the distance from the origin traveled by the random walker is less than r_0: $r_{\max} \leq r_0$. To this end we consider an absorbing sphere at the radius $r = r_0$ and obtain the propagator in the domain $0 \leq r \leq r_0$ using the following boundary and initial conditions:

$$P(r_0, t) = 0 \text{ and } P(r, 0) = \frac{\delta_+(r)}{r^{d-1}}. \tag{2.44}$$

Any random walker that has reached before t a maximal excursion $r_{\max} \geq r_0$ has here been absorbed by the sphere $r = r_0$ before t. The probability to be within the sphere at time t is thus the probability to have a maximal excursion $r_{\max}$ at time t smaller than r_0.

$$\Pr\left(r_{\max} \leq r_0, t\right) = \int_0^{r_0} P(r, t) r^{d-1} dr. \tag{2.45}$$

The nth moment is obtained by integration of $r^n_{\max}$ times the derivative of the cumulative distribution, *i.e.* the density,

$$\begin{aligned} \langle r^n_{\max}\rangle &= \int_0^{\infty} r_0^n \frac{\partial\left(\Pr\left(r_{\max} \leq r_0, t\right)\right)}{\partial r_0} dr_0 \\ &= n \int_0^{\infty} r_0^{n-1} \left(1 - \Pr\left(r_{\max} \leq r_0, t\right)\right) dr_0. \end{aligned} \tag{2.46}$$

Results are summarized in Table 2.3, and the computation detailed in Appendix 1.

One can note that the ratio $\langle X^4\rangle/\langle X^2\rangle^2$ is smaller for maximal excursion moments than for regular moments. This means that the relative variance is smaller for maximal excursion that for displacement: one can extract, with a greater accuracy, the diffusion coefficient or the anomalous diffusion coefficient (here $\alpha = 1$), from the same data set, using maximal excursion.

This point is really important for small data sets: a power-law fit using mean square displacement could gives an anomalous diffusion coefficient $\alpha \neq 1$, while a power-law fit on the mean square maximal excursion gives a normal diffusion $\alpha \simeq 1$.

Extension to diffusion on fractal

So far we focused on the Brownian motion. Those computation can be extended to anomalous diffusion processes, at least to some extent.

We will first focus on diffusion on fractals. On such media, we can consider that we still have a classicam random walk, where the effective space dimension d can be any (positive) real number. As in the first chapter, we note d_f the (fractal) space dimension, with $V \propto r^{d_f}$, where V is the volume of a sphere of radius r, and d_w the walk dimension, that characterizes the mean square displacement, $\langle r^2 \rangle \propto Dt^{2/d_w}$. Those two dimensions are fractal-specific. We then use the generic diffusion equation from O'Shaughnessy–Procaccia [154]:

$$\frac{\partial P(r,t)}{\partial t} = \frac{D}{r^{d_f-1}} \frac{\partial}{\partial r} \left(r^{1-d_w+d_f} \frac{\partial P(r,t)}{\partial r} \right). \tag{2.47}$$

As for the Brownian motion, we can solve this equation, and find:

$$\langle r^{2k} \rangle = \frac{\Gamma\left(\frac{d_f+2k}{d_w}\right)}{\Gamma\left(\frac{d_f}{d_w}\right)} \left(D d_w^2 t\right)^{2k/d_w} \tag{2.48}$$

If we use $\alpha = 2/d_w$, this equation writes:

$$\langle r^{2k} \rangle = \frac{\Gamma\left(\frac{(d_f+2k)\alpha}{2}\right)}{\Gamma\left(\frac{d_f\alpha}{2}\right)} \left(\frac{4Dt}{\alpha^2}\right)^{k\alpha} \tag{2.49}$$

The diffusion equation can be used similarly for the absorbing sphere case, to retrieve the following result [34]:

$$\langle r_{\max}^k \rangle = A_{k,d_f,d_w} \left(\frac{d_w^2}{4} Dt\right)^{k/d_w} = A_{k,d_f,d_w} \left(\frac{Dt}{\alpha^2}\right)^{k\alpha/2}, \tag{2.50}$$

with

$$A_{k,d_f,d_w} = \frac{2^{1-d_f/d_w} 2k}{d_w \Gamma\left(\frac{k}{d_w}+1\right) \Gamma\left(\frac{d_f}{d_w}\right)} \int_0^\infty \frac{u^{(2k+d_f)/d_w-2}}{I_{d_f/d_w-1}(u)} du \tag{2.51}$$

Here, I_n is the modified Bessel function of the first kind of order n.

Extension to CTRW

We will now extend the Brownian result for CTRW. As presented in chapter 1, CTRW is in its discrete version a succession of jumps, where the random walker rest a random waiting time τ between two steps. We will assume that the waiting time distribution $\psi(\tau)$ scales as $\psi(\tau) \simeq \tau_0/\tau^{1+\alpha}$ with $0 < \alpha < 1$. To simplify the calculus, we will use the continuous limit of CTRW, namely the fractional diffusion, in order to retrieve a diffusion equation, as for the Brownian motion. This continuous process is governed by the following fractional diffusion equation [172]

$$\frac{\partial P(r,t)}{\partial t} = \frac{K_\alpha}{r^{d-1}} \frac{\partial}{\partial r} \left(r^{d-1} \frac{\partial}{\partial r} \left({}_0D_t^{1-\alpha} P(r,t) \right) \right). \tag{2.52}$$

Here we used the Riemann-Liouville fractional operator [172] introduced in chapter 1:

$${}_0D_t^{1-\alpha} P(r,t) = \frac{1}{\Gamma(\alpha)} \frac{\partial}{\partial t} \int_0^t dt' \frac{P(r,t')}{(t-t')^{1-\alpha}}, \tag{2.53}$$

The generalized diffusion constant is $K_\alpha = \langle \delta \mathbf{r}^2 \rangle / [2\tau_0^\alpha]$, where we used the (finite) jump length variance $\langle \delta \mathbf{r}^2 \rangle$.

The Laplace transform of the diffusion equation is therefore

$$sP(r,s) - P(r,0^+) = s^{1-\alpha} \frac{K_\alpha}{r^{d-1}} \frac{\partial}{\partial r} \left(r^{d-1} \frac{\partial}{\partial r} P(r,s) \right). \tag{2.54}$$

With this result, we can use the so-called subordination [227]. This means that we can replace in Laplace space s by $K_1 s^\alpha / K_\alpha$ where s is the Laplace variable conjugated to time t. Using this method, we obtain the ratio for both regular moments and maximal excursion statistics from the Brownian result, however, with different pre-factors

$$\langle r^k \rangle_{\mathrm{CTRW}} = \frac{\Gamma(k/2+1)}{\Gamma(\alpha k/2+1)} \frac{\langle r^k \rangle_{\mathrm{BM}}}{Dt} K_\alpha t^\alpha, \tag{2.55}$$

$$\langle r^k_{\max} \rangle_{\mathrm{CTRW}} = \frac{\Gamma(k/2+1)}{\Gamma(\alpha k/2+1)} \frac{\langle r^k_{\max} \rangle_{\mathrm{BM}}}{Dt} K_\alpha t^\alpha, \tag{2.56}$$

We can comment quickly this result. Since $\alpha < 1$, regular and maximal ratio are greater for a sub diffusive CTRW than for an usual Brownian motion:

$$\frac{\langle r^4 \rangle_{\mathrm{CTRW}}}{\langle r^2 \rangle^2_{\mathrm{CTRW}}} = \frac{2\Gamma(\alpha+1)^2}{\Gamma(2\alpha+1)} \frac{\langle r^4 \rangle_{\mathrm{BM}}}{\langle r^2 \rangle^2_{\mathrm{BM}}} \tag{2.57}$$

and similarly

$$\frac{\langle r^4_{\max} \rangle_{\mathrm{CTRW}}}{\langle r^2_{\max} \rangle^2_{\mathrm{CTRW}}} = \frac{2\Gamma(\alpha+1)^2}{\Gamma(2\alpha+1)} \frac{\langle r^4_{\max} \rangle_{\mathrm{BM}}}{\langle r^2_{\max} \rangle^2_{\mathrm{BM}}} \tag{2.58}$$

This means that the relative variance of the mean square displacement and of the mean square maximal excursion is greater for a CTRW than for a Brownian motion. For a CTRW, the random walker can stay for a very long time at the same place: the resulting trajectory is a succession of jumps separated by waiting times. A given data set exhibits a wider variance, since some of the walkers stay stuck for a very long time while others continue to evolve.

Extension to fBm

FBm is a non Markovian process: we cannot define a diffusion equation local in time. But we have an analog of propagator in free space [224], namely the one-point distribution of trajectories at time t, starting from $r = 0$ at $t = 0$:

$$P(r,t) = \frac{2}{\Gamma(d/2) \left(\frac{2K_\alpha}{d} t^\alpha \right)^{d/2}} \exp\left(-\frac{r^2}{\frac{2K_\alpha}{d} t^\alpha} \right), \tag{2.59}$$

which is equal to the propagator of Brownian motion after the substitution $t \to t^\alpha$. This expression is in fact the superposition of d independent fBm processes. It is not a real propagator, since we here assume that the fBm started in $t = -\infty$ and is in $r = 0$ at time $t = 0$: we thus waited that the increment reach the stationary probability before beginning the random walk. We have with equation (2.59) a pseudo-propagator of the walker distribution after a time t, starting for an "equilibrated" state. But we cannot use it as a classical propagator: knowing the position at time t' does not allow to get the position at time $t > t'$ by injecting $t - t'$ in equation (2.59).

	1 D	2 D	3 D
$\langle r^2 \rangle$	$K_\alpha t^\alpha$	$K_\alpha t^\alpha$	$K_\alpha t^\alpha$
$\langle r^4 \rangle$	$3\left(K_\alpha t^\alpha\right)^2$	$2\left(K_\alpha t^\alpha\right)^2$	$5/3\left(K_\alpha t^\alpha\right)^2$
$\langle r^4 \rangle / \langle r^2 \rangle^2$	3	2	5/3

Table 2.4: FBm regular moments.

H	fitted α ($= 2H$ in theory)	fitted α'
0.05	0.0991	0.2491
0.10	0.1993	0.3301
0.15	0.3015	0.4114
0.20	0.4003	0.4922
0.25	0.5015	0.5742
0.30	0.5971	0.6561
0.35	0.6979	0.7439
0.40	0.7982	0.8342
0.45	0.8962	0.9231
0.50	0.9975	1.0168

Table 2.5: FBm regular (α) and maximal excursion (α') second moment exponent in 2D (numerical simulations).

From equation (2.59) we can still compute at least the regular moments, with the analogous time substitution. The results are collected in Table 2.4. To compute the maximal excursion moments, one would need to know the solution of fBm in the presence of an absorbing boundary, a so far elusive quantity.

We estimated the maximal excursion second and fourth moment and the related ratio with numerical simulations. We generated 100,000 trajectories of 512 steps, in two dimensions, using the Hosking method [109], for H exponents from 0.05 ($\alpha = 0.1$) to 0.5 ($\alpha = 1$, Brownian motion), by step of 0.05.

We plotted in Figure 2.7 the second maximal excursion moment as a function of time. We observe that it behaves like a power-law for all H values, if we except the initial points, were the estimate of $r_{\max}$ is not very accurate. The surprising result is that the exponents are always greater than α. We note α' the second maximal excursion moment exponent, and we reported in Table 2.5 the fitted values of α and α'. α values are always very close to the theoretical ones ($2H$), and $H = 0.5$ gives back the Brownian motion ($\alpha = \alpha' = 1$). We see that α' behaves linearly with α, a linear regression gives, with a high correlation coefficient:

$$\alpha' \simeq 0.156 \pm 0.005 + (0.849 \pm 0.008)\alpha \tag{2.60}$$

Figure 2.8 shows the MME ratio as a function of time. We here obtain a more classical behavior, with a saturation for long times. This result was not obvious, it means that the fourth MME moment scales as $t^{2\alpha'}$. The saturation value grows with H.

We note that all curves start from 2 for $t = 1$, which is the MSD ratio in 2D ($r = r_{\max}$ for $t = 1$), and that the upper curve converge toward 1.49, the expected value for a Brownian motion in 2D ($H = 0.50$ is a Brownian motion). We do not have access to the limit $t \to \infty$, so we estimated the ratio using the minimal value, reached at time $t = 512$. Table 2.6 summarizes those results. Since the precision of the estimation is not very good, it is hard to recognize a

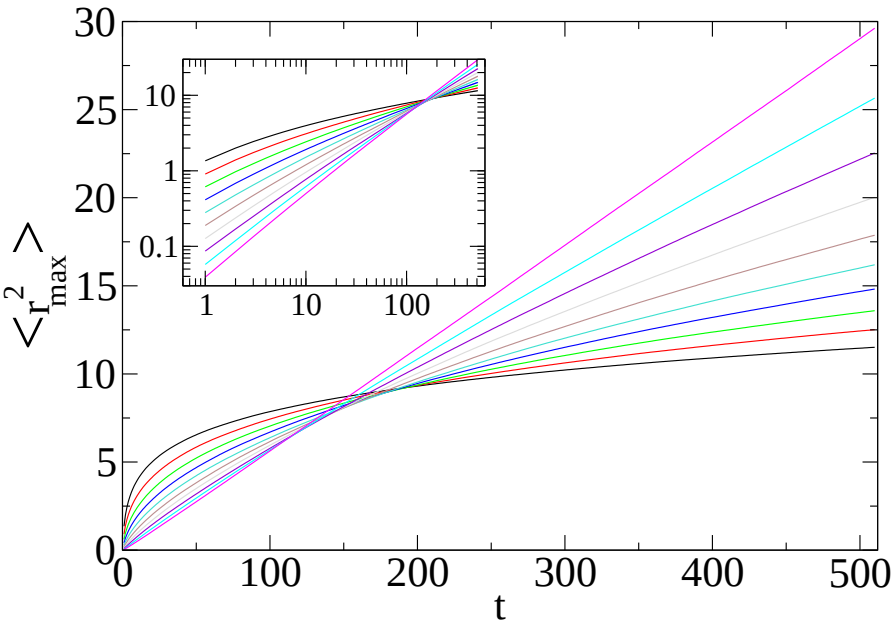

Figure 2.7: Second maximal excursion moment as a function of time for different fBm in 2 dimensions. The Hurst exponent varies from 0.05 (black curve) to 0.5 (magenta curve) by steps of 0.05. Simulations are made using the Hosking method [109]. The inset is a log-log plot.

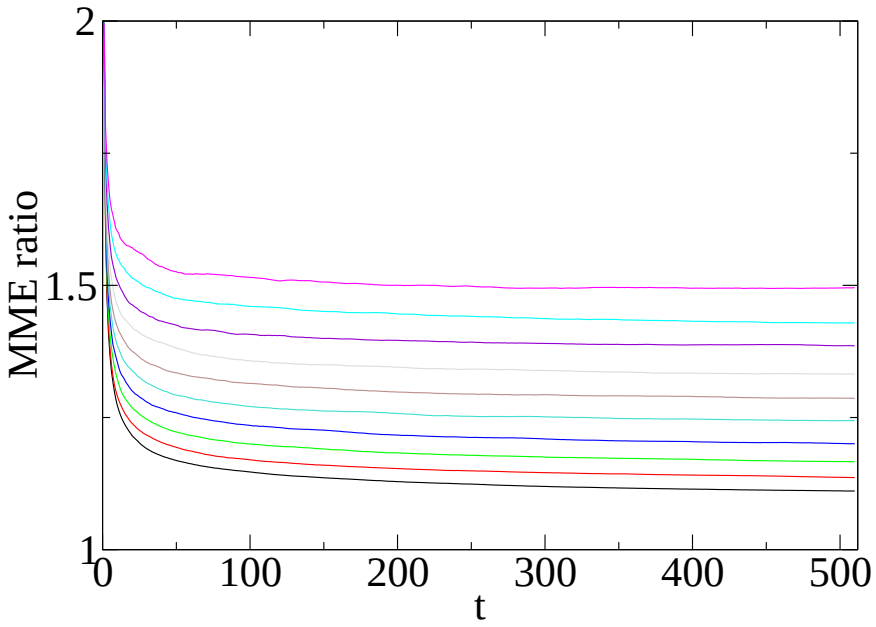

Figure 2.8: Maximal excursion ratio ($\langle r^4_{\max}\rangle / \langle r^2_{\max}\rangle^2$) for different fBm in 2 dimensions, as a function of time. The Hurst exponent varies from 0.05 (black curve) to 0.5 (magenta curve) by steps of 0.05. Simulations are made using the Hosking method [109].

H	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
Maximal excursion ratio	1.11	1.14	1.17	1.20	1.24	1.29	1.33	1.39	1.43	1.49

Table 2.6: Maximal excursion ratio ($\langle r^4_{\max}\rangle / \langle r^2_{\max}\rangle^2$) for different fBm in 2 dimensions at time $t = 512$.

particular behavior of this ratio with H. We propose a rough approximation using a shifted power-law, just for the sake to have a quick numerical value for any H:

$$\frac{\langle r^4_{\max}\rangle}{\langle r^2_{\max}\rangle^2} \simeq (1.05 \pm 0.01) H^{1.42 \pm 0.01} + (1.10 \pm 0.01) \tag{2.61}$$

We see that fractional Brownian motion have a relative variance, only for maximal excursion, lower than for a classical Brownian motion. Sub-diffusive fBm are anti-correlated: they tend to stay in the same area longer than a Brownian motion, and thus to have a less disperse envelope for maximal excursion. The exponent difference between displacement and maximal excursion is a clear mark of the fBm non Markovian nature.

Confrontation with numerical simulation

To see if the two ratio were enough to discriminate between CTRW, fBm and diffusion on fractal, we computed for the three simulated data sets the second moments as well as the ratios, for both displacement and maximal excursion.

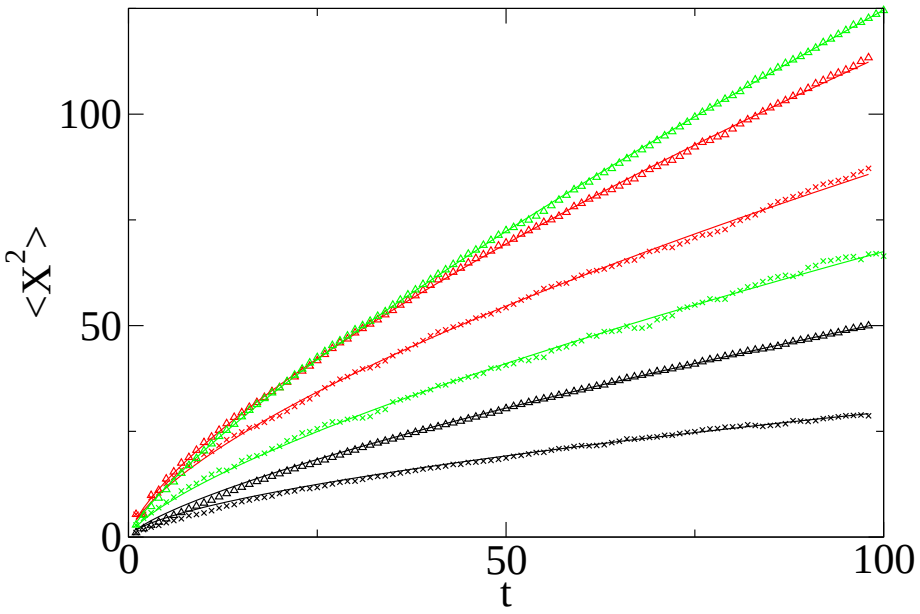

Figure 2.9: $\langle r^2(t)\rangle$ and $\langle r^2_{\max}\rangle$ as function of time t (arbitrary units) for the three simulated time series (1,000 trajectories of 100 steps each), each with anomalous diffusion exponent $\alpha = 0.7$. The black set corresponds to 2D percolation data, the red set to CTRW on Euclidian lattice, and the green set to fBm. For each set, the crosses $\times$ stand for $\langle r^2(t)\rangle$, and the triangles (Δ) for $\langle r^2_{\max}\rangle$.

	$\langle r^2 \rangle$	$\langle r^2_{\max} \rangle$
Fractals	$\alpha \simeq 0.64$	$\alpha \simeq 0.73$
CTRW	$\alpha \simeq 0.67$	$\alpha \simeq 0.71$
FBM	$\alpha \simeq 0.72$	$\alpha' \simeq 0.79$

Table 2.7: Anomalous diffusion exponent obtained by power-law fit for simulated data sets.

Figure 2.9 shows the result: for both displacement and maximal excursion second moment, we obtain power-law behavior. In this figure, diffusion on fractals is in black (percolation network), CTRW in red and fBm in green. Crosses ($\times$) represent the ensemble averaged mean square displacement, triangles (Δ) the ensemble averaged mean square maximal excursion.

The three data sets are simulated with the same $\alpha = 0.7$, the fitted anomalous diffusion exponent are given in Table 2.7. We can note that the exponent obtained with maximal excursion method is better, namely closer to the input value, for both fractal and CTRW data sets. We also expect different exponents α and α' for displacement and maximal excursion second moments in the fBm case, but this is not obvious with our simulated data set.

To refine the analysis, we plotted in Figure 2.10 the ratio $\langle X^4 \rangle / \langle X^2 \rangle^2$ for both quantities, and for the three data sets. In this figure, diffusion on fractals is in black (percolation network), CTRW in red and fBm in green. $+$ represent the displacement ratio, Δ the maximal excursion one. The continuous lines are the expected values for the displacement ratio, the short lines the expected values for the maximal excursion ratio.

For diffusion on fractals (black), we have a discrepancy between simulated and theoretical values. This might be related to the fact that simulations take place in a confined environment (percolation cluster on a 250×250 network): the random walker quickly reaches the boundaries and the convergence occurs toward the equilibrium distribution, and not toward the free space propagator.

The displacement ratio seems to oscillate much more than maximal excursion ratio. This could be explained by the intrinsic larger relative variance of displacement statistic.

The maximal excursion ratio seems to always decrease toward the infinite value. This can be partly explained by the sampling used in practical computation. When we calculate the maximal excursion, we use a diffusion equation that works for the continuous process. The maximal excursion we have at time t is the maximum of r on the segment $[0, t]$. When we simulate the time series, we do not have a continuous trajectory between 0 and t, but a discrete sampling over $[0, t]$. The "real" maximum at time t could have occur, for the continuous process, between two frames of the sampling, and we could have missed it. When we compare the theoretical results with the maximal excursion over the sampling, we expect discrepancies due to the sampling: we will always underestimate the maximal excursion, but the relative error will decrease with the sample size. As the frame number increases, we expect that the computed ratio slowly converges toward the "infinite" ratio.

We know that for the first frame, since we have only one point in the sampling, maximal excursion and displacement are the same: the two ratios are equal for this first frame. The displacement ratio being greater, the computed maximal excursion ratio decreases, as the frame number increases, toward the theoretical ratio.

If we take diffusion on fractals apart (black data sets), one can clearly distinguish CTRW and fBm using those ratio: CTRW ratios are above the Brownian one, respectively 2 for the displacement ratio and 1.79 for the maximal excursion ratio in 2D, fBm is below for maximal excursion ratio, and around 2 for the displacement ratio.

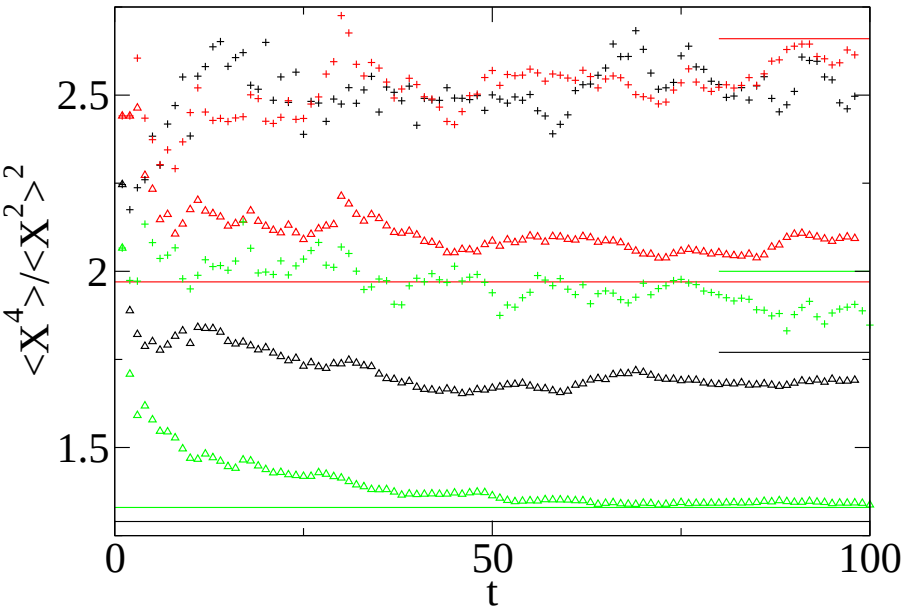

Figure 2.10: $\langle r^4(t)\rangle/\langle r^2(t)\rangle^2$ and $\langle r^4_{\max}(t)\rangle/\langle r^2_{\max}(t)\rangle^2$ as function of time t for the three simulated time series (1,000 trajectories of 100 steps each). The maximal excursion ratio for the diffusion on a 2D percolation cluster (black triangles) does not converge to the expected value 1.29 (continuous black line). The same behavior is observed for the displacement ratio (black +), for which the expected value is 1.77 (short black line). The maximal excursion ratio for the CTRW process (red triangles) converges slowly to 1.97 (continuous red line). The displacement ratio (red +) is more irregular and converges to 2.66 (short red line). At last, for fBm, the maximal excursion ratio (green triangles) converges to the estimated value of equation (2.61), 1.33 (continuous green line), and the displacement ratio (green +) oscillates around the Brownian value 2 (short green line).

QUICK SUMMARY

- Maximal excursion is more efficient than displacement to determine the anomalous diffusion exponent α and coefficient K_α ;
- Regular and maximal excursion ratios are different for diffusion on fractal, CTRW and fBm ;
- The exponent is not the same for displacement (α) and maximal excursion (α') for, and only for, fBm ;
- Ratios seem to be more efficient to distinguish mechanisms than second moments.

2.2.3 Growing shell analysis

Since the previous data did not allow to clearly distinguish diffusion on fractals with the other proposed mechanisms, we developed a new test dedicated to fractal detection. This test will be based on the scaling form of the propagator that depends on both d_w (or similarly the anomalous diffusion exponent α) *and* d_f. Knowing d_w from the second moments computed before, we can deduce d_f. We will first present how to use the scaling form of the propagator, before presenting how to perform the growing shell analysis, and confronting this test with numerical simulations.

Theoretical determination of d_f

We know that the probability density for a diffusing particle on a fractal satisfies the scaling relation [71] (we recall that $d_w = 2/\alpha$)

$$P(r,t) = t^{-d_f/d_w} P\left(\frac{r}{t^{1/d_w}}, 1\right) = t^{-\alpha d_f/2} P\left(\frac{r}{t^{\alpha/2}}, 1\right) \tag{2.62}$$

The same relation hold for a CTRW or a fBm (if we consider only the one-point distribution of trajectories) if we replace d_f by the Euclidian dimension d. Let us focus on the probability to be in a growing sphere of radius $r_0 t^{\alpha/2}$

$$\begin{aligned} \Pr\left(r \leq r_0 t^{\alpha/2}, t\right) &= \int_0^{r_0 t^{\alpha/2}} r^{d-1} P(r,t) dr \\ &= A(r_0) t^{\alpha(d-d_f)/2} \end{aligned} \tag{2.63}$$

Since the exponent α is known from the second moment fit we can extract d_f from the above relation.

Practical computation

With a finite data set, we can compute the probability to find the particle at time t in a (growing) shell of radius $r_0 t^{\alpha/2}$. Here r_0 is a free parameter. It should be chosen sufficiently large, such that for a given trajectory the probability to be within the shell is appreciably large. At the same time it should not be too large, otherwise the probability to be within the shell is always almost one. Choosing a small multiple of $\langle r(t=1)\rangle$ appears to be a good compromise. The probability

to be inside the shell then becomes

$$\Pr\left(r \leq r_0 t^{\alpha/2}\right) \simeq \frac{1}{\mathcal{N}(t)} \sum_{i=1}^{\mathcal{N}(t)} \Theta\left(r_i(t) - r_0 t^{\alpha/2}\right), \tag{2.64}$$

where $\Theta(r)$ is the Heaviside function, that equals 1 if $r \geq 0$, and 0 if $r < 0$. We expect that this probability scales as $t^{\alpha(d-d_f)/2}$. To fit the fractal dimension d_f we need the anomalous diffusion exponent α as input. We will use the value extracted from the second maximal excursion moment fits. The direct plot of the probability is quite easy to interpret : if the probability is constant, then $d = d_f$, if this probability slowly grows, then $d > d_f$, and the support is fractal ($d_f \neq d$). The dimension d is here the dimension of the trajectories (2 in our examples due to the projection onto the focal plane).

Confrontation with numerical simulations

Figure 2.11 shows the plot of the probability to be in a shell growing in $r_0 t^{\alpha/2}$ for the three simulated data sets (r_0 changes between the different sets). α is here the fitted value of the second maximal excursion moment for each set. Once again, the black set represent diffusion on fractals (critical percolation), the red one CTRW and the green one fBm .

We see clearly that for CTRW and fBm, the probability is approximately constant, and that for the diffusion on a percolation cluster, it grows with time, indicating that $d_f < d$, as it should be.

We performed power-law fit of each set to quantify this. We obtained $d - d_f \simeq 0.11$, which corresponds to $d_f \simeq 1.89$ (exact value $91/48 \simeq 1.896$) for diffusion on fractals (black set). CTRW gives $d - d_f \simeq 0.01$ instead of 0, and fBm $d - d_f \simeq -0.004$ instead of 0. This means that the growing shell analysis gives a very good indication on the fractality of the space where the random walk takes place.

Quick summary (in 2D)

	$(\langle r^2\rangle(t), \langle r^2_{\max}(t)\rangle)$	$\left(\frac{\langle r^4\rangle(t)}{\langle r^2\rangle^2(t)}, \frac{\langle r^4_{\max}\rangle(t)}{\langle r^2_{\max}\rangle^2(t)}\right)$	$\Pr\left(r \leq r_0 t^{\alpha/2}, t\right)$
Brownian motion	$(\propto t, \propto t)$	$(2, 1.49)$	A_0
Fractals	$(\propto t^\alpha, \propto t^\alpha)$	$(< 2, < 1.49)$	$\propto t^{\alpha(2-d_f)/2}$
CTRW	$(\propto t^\alpha, \propto t^\alpha)$	$(> 2, > 1.49)$	A_0
FBm	$(\propto t^\alpha, \propto t^{\alpha'})$	$(2, < 1.49)$	A_0

2.2.4 Application to experimental data sets

So far, we only considered simulated data sets, with an infinite precision: we did not deal with the classical experimental problems, like photobleaching (missing frames, trajectories of different lengths), displacement of the sample (drift, random moves of the microscope elements), observation in aqueous solutions (optical parasites, solvent motion). The random walker position is obtained after image analysis, and even this process is not perfect: the Gaussian fit can lead to new artifacts in the data.

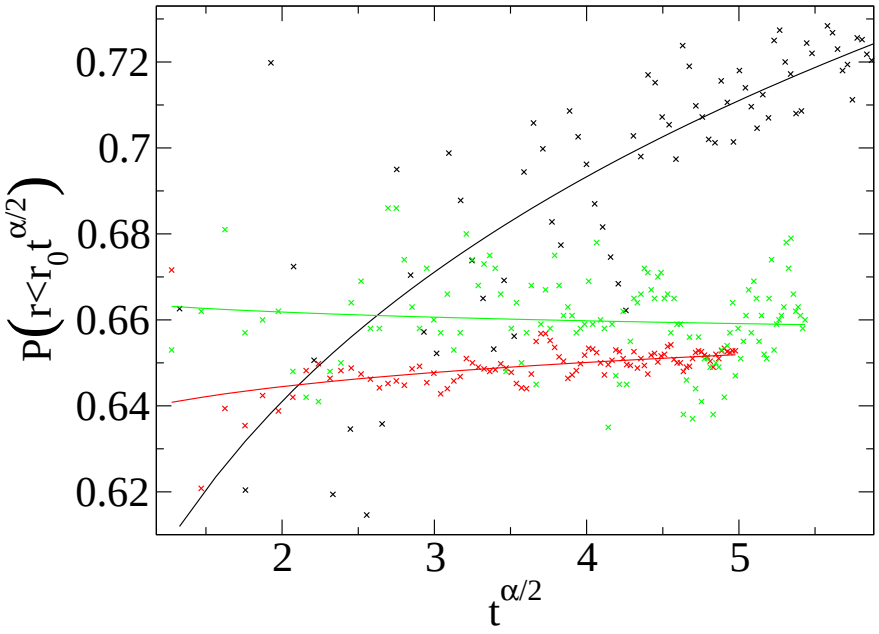

Figure 2.11: Probability to be in a growing shell of radius $r_0 t^{\alpha/2}$ as function of $t^{\alpha/2}$ for the three simulated sets (arbitrary units). This analysis is based on the previously fitted values of α. The 2D critical percolation cluster (black $\times$) produces $d - f_f \simeq 0.11$, *i. e.* $d_f \simeq 1.89$ (exact value $91/48 \simeq 1.896$). The CTRW set (red $\times$) gives $d - d_f \simeq 0.01$ instead of 0, and the fBm set (green $\times$) leads to $d - d_f \simeq -0.004$ instead of 0.

Since the method is dedicated to experimental analysis, we applied it on experimental data sets. The first one is simply an observation of Quantum dots (Qdots) freely diffusing in solution. The aim is to see if all the artifacts described previously have an influence on our analysis, and if so, in which proportion. The second data set is a biological observation of lipid granules in living cells. We will try to apply our formalism, to see if we can say something on the diffusion of lipid granules with our observables.

a) Analysis of Qdots data

Experimental data acquisition

The first set of data was obtained from fluorescence video tracking of single quantum dots freely diffusing in glycerol. The trajectories were recorded by a CCD camera, and then filtered by a particle tracker software. The software extracted the trajectories $(t, x(t), y(t))$ using a Gaussian fit. The details of the experimental setup are in Appendix 1.

The analysis is based on 67 trajectories, the longest of which consists of 210 frames. For these data, we expect to observe normal Brownian motion.

Analysis

The first test is simply to plot the second moment of both displacement and maximal excursion as a function of time. As shown in Figure 2.12, if we make an ensemble average of the 67 trajectories, we obtain for both $\langle r^2 \rangle$ and $\langle r^2_{\max} \rangle$ somehow a linear behavior. If we perform a power-law fit to confirm this, we obtain $\langle r^2 \rangle \propto t^{0.81}$ (red line) and $\langle r^2_{\max} \rangle \propto t^{1.02}$ (magenta line).

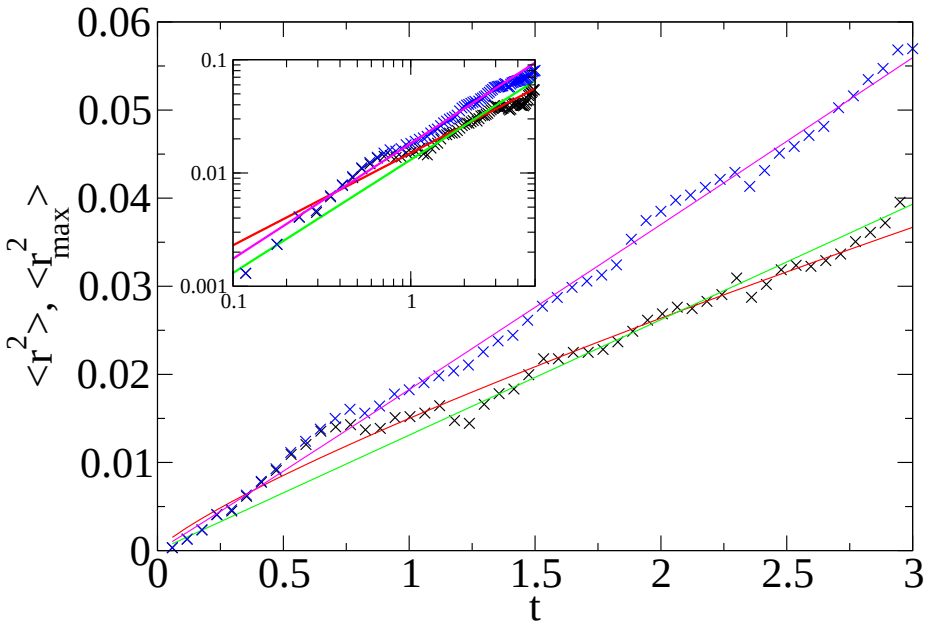

Figure 2.12: Analysis of an experimental set of 67 trajectories, the longest consisting of 210 points, for quantum dots freely diffusing in a solvent. Second moment of displacement (black $\times$) fitted by a power law leads to an exponent $\alpha = 0.81$ (red line). We also show a linear fit where α is set to 1 (green line, expected behavior for a Brownian motion). The maximal excursion second moment (blue $\times$) fitted by a power law lead to $\alpha = 1.02$ (magenta line). Time is in s; distances are in μm. (Inset) log-log plot of the same data.

This means once again that the maximal excursion has a better statistic: one can extract a more accurate α using the second maximal excursion moment than using the usual mean square displacement. With a larger data set, this distinction decreases, since the accuracy of the second moments exceed the data accuracy. Indeed, the relative variance decreases as the data set increases, while experimental errors are constant: the limitation does not come anymore from the observable used but directly from the data.

The second observation is that the fBm identification based on the difference between α and α' is not really usable for a limited data set. The difference we see is more linked to the statistic difference than to the microscopical diffusion mechanism.

At last, we performed a linear fit of the mean square displacement in Figure 2.12 (green line). This fit does look acceptable: the power law fit leading to $\alpha = 0.81$ is convincing at short time,

but if we had longer trajectories, we would certainly retrieve the classical Brownian behavior. This is an important point to emphasize: if we focus only on the mean square displacement with short trajectories, we could obtain an "anomalous" diffusion, while it is only due to a medium statistic quality, and a short-time artifact of the power-law fit. To confirm an anomalous diffusion, one should always check that other observables also lead to the same conclusion.

If we plot the ratio of displacement and of maximal excursion, we see in Figure 2.13 that the ratio converge toward the expected values for a Brownian motion (continuous lines). With trajectories of different lengths, we observe a slower convergence, since only few trajectories are averaged for long times. We discarded here the $t > 120$ frames.

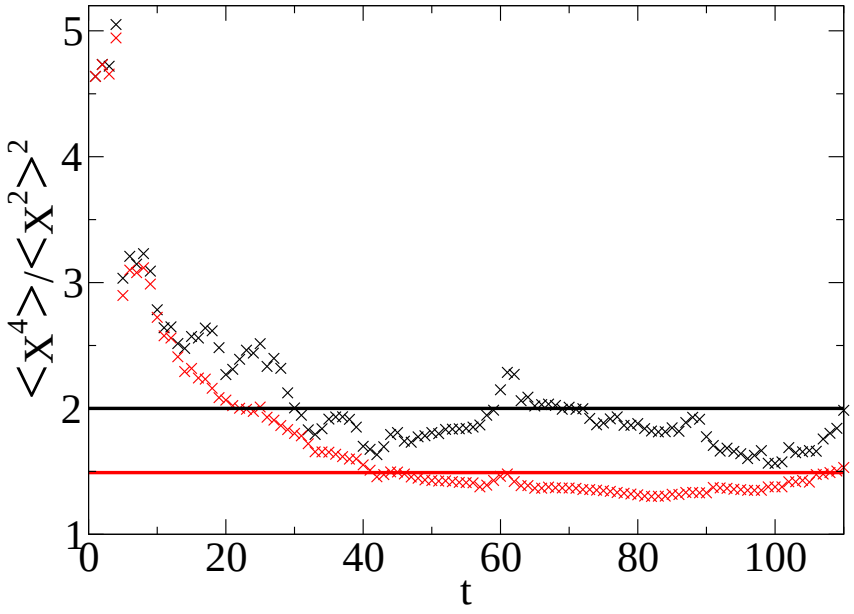

Figure 2.13: Analysis of an experimental set of 67 trajectories, the longest consisting of 210 points, for quantum dots freely diffusing in a solvent. Ratio of displacement (black $\times$) and of maximal excursion (red $\times$) as a function of time (in frames), compared to the expected values for a Brownian motion, namely 2 for the displacement ratio (black line), and 1.49 for the maximal excursion ratio (red line).

At last, we can apply the growing shell analysis, to see the method accuracy. Figure 2.14 shows that we obtain a relatively flat probability. With a small data set, we have some noise in the result, but a "blind" power-law fit leads to $d - d_f = -0.012$, close to 0 as expected. This test seems to be quite reliable even for a small and noisy data set.

b) Lipid granules

The real purpose of the whole method is to extract information on the microscopical diffusion mechanism using a finite set of data. We applied it to a real biological problem: lipid granules in yeast cells.

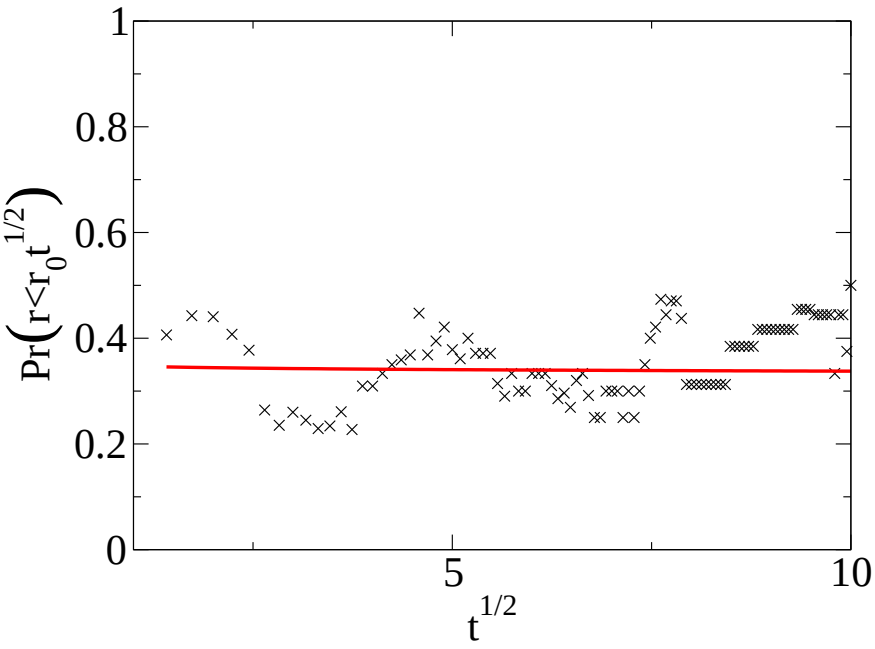

Figure 2.14: Analysis of an experimental set of 67 trajectories, the longest consisting of 210 points, for quantum dots freely diffusing in a solvent. Probability to be in a growing shell of radius $r_0 t^{\alpha/2}$ as function of $t^{\alpha/2}$ ($\alpha = 1$, t in frames). The power law fit (red line) leads to $d - d_f = -0.012$, close to 0 as expected.

Experimental data acquisition

In this second set of much longer trajectories (5,500 to 19,400 frames) the particle positions were acquired by video tracking of lipid granules in yeast cells by Christine Selhuber-Unkel and Lene B. Oddershede at Niels Bohr Institute (Copenhagen, Denmark). They used *Schizosaccharomyces pombe* (*S. pombe*, D817) fission yeast cells, expressing a GFP-fused marker of the nuclear and plasma membrane systems [77]. Spherical lipid granules are endogenously present in the cytoplasm of these cells [174]. The granules are refractive enough to be visualized with bright-field microscopy.

The *S. pombe* fission yeast cells appear cylindrical with outer dimensions of approximately 4 μm by 12 μm. The cell is surrounded by a rigid cell wall. The granules which were tracked in the present experiment are densely packed lipids of almost spherical shape and of approximately 300 nm diameter. These granules are located uniformly throughout the cytoplasm, except in the nucleus. The granules perform thermal fluctuations, these fluctuations being somewhat hindered by the presence of cytoskeletal elements such as microtubules, actin, and membranous structures. Due to the coupling to their viscoelastic environment one would expect a sub-diffusive behavior of the granules. It is, however, possible that a minority of the granules are actively moved, either by molecular motors (though, this has not been proved in literature to our best knowledge) or by cytoplasmic streaming.

Analysis

This set contains video tracking of eight different lipid granules moving in yeast cells. The trajectories are between 5,515 and 19,393 frames long. As we had few long trajectories, before an ensemble average we first directly analyzed the eight trajectories using the time-averaged mean square displacement. We compute the following quantity:

$$\langle r^2 \rangle_t(\Delta, T) = \frac{1}{T-\Delta} \sum_{i=1}^{T-\Delta} \|\mathbf{r}(i+\Delta) - \mathbf{r}(i)\|^2 \tag{2.65}$$

If the microscopical mechanism is ergodic, namely if time average and ensemble average lead to the same result, this time-average yields an accurate result, since we average over all the trajectory, even with only eight distinct trajectories. As shown in Figure 2.15, we obtain a distinct sub-diffusive behavior with an exponent close to 0.4. It is interesting to see that the data exhibit a scatter in amplitude and considerable local variation of slope. Such features were also observed previously; see, for instance, [54, 99].

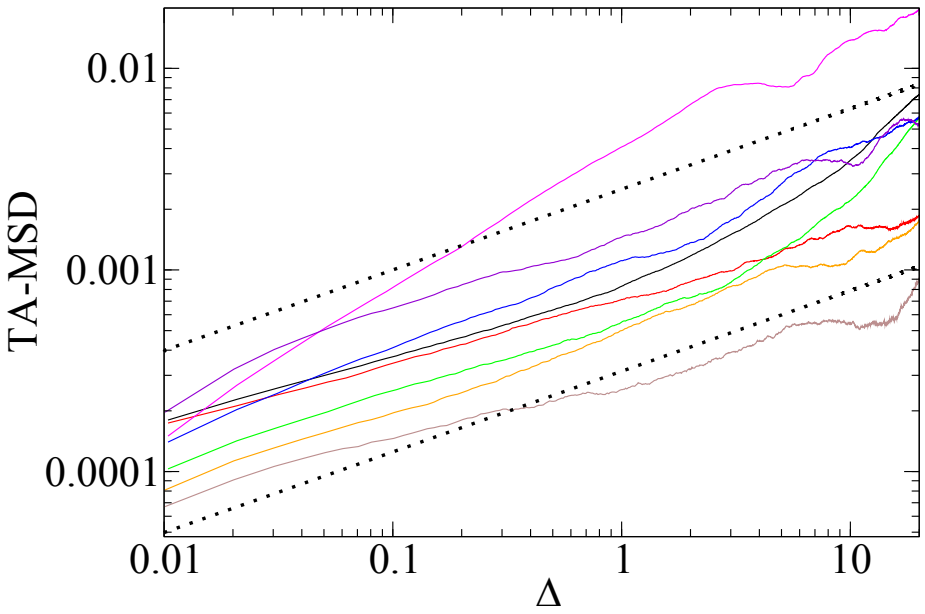

Figure 2.15: Lipid granules diffusing in a yeast cell. Eight trajectories, between 5,515 and 19,393 frames' long. Log-log plot of the time-averaged mean square displacement as a function of the lag time (continuous lines), and $A_0 t^{0.4}$ (dotted lines). Lag time is in s, and distances in μm.

The scattering of the diffusion coefficient may possibly be related to aging effects [11]. Those effects have so fare been related to CTRW, but for such mechanism, one would expect a linear behavior for the time-average mean square displacement. Another explanation could simply be that the granules size varies between two trajectories, each trajectory being the observation of a

given lipid granules. Similarly, one could argue that each trajectory correspond to a given cell organization, with a given cytoskeleton, that depends on the cell phase in the cell cycle. In the wider set that will be studied in the next sub-section, we will look more closely on the cell phase influence on the diffusion properties of lipid granules.

We also note that one of the curves shows a much steeper slope than do the others (magenta line in Figure 2.15). Here, it is not the diffusion coefficient that changes, but the anomalous diffusion exponent. We here clearly suspect a modification in the cell organization: we do expect a Brownian behavior for lipid granules in a aqueous solvent, and we observe a clear sub-diffusive behavior within a cell. The cell impact could be modulated, for instance depending on the cell cycle phase (G_0, G_1, S, G_2 or M): the cytoskeleton for instance is not the same during cell division and during interphase. If the cytoskeleton interacts with lipid granules, either with simple volume exclusion or with chemical interaction, the anomalous diffusion exponent could depend on the cytoskeleton organization, and thus on the cell cycle phase.

We extended the time-average analysis to the second maximal excursion moment and again obtained a clear sub-diffusive behavior, but with an exponent close to 0.5, as shown in Figure 2.16. This time-averaged second maximal excursion moment is computed as follows:

$$\langle r^2_{\max}\rangle_t(\Delta, T) = \frac{1}{T-\Delta}\sum_{i=1}^{T-\Delta} \max_{i\leq t\leq i+\Delta} \|\mathbf{r}(t) - \mathbf{r}(i)\|^2 \tag{2.66}$$

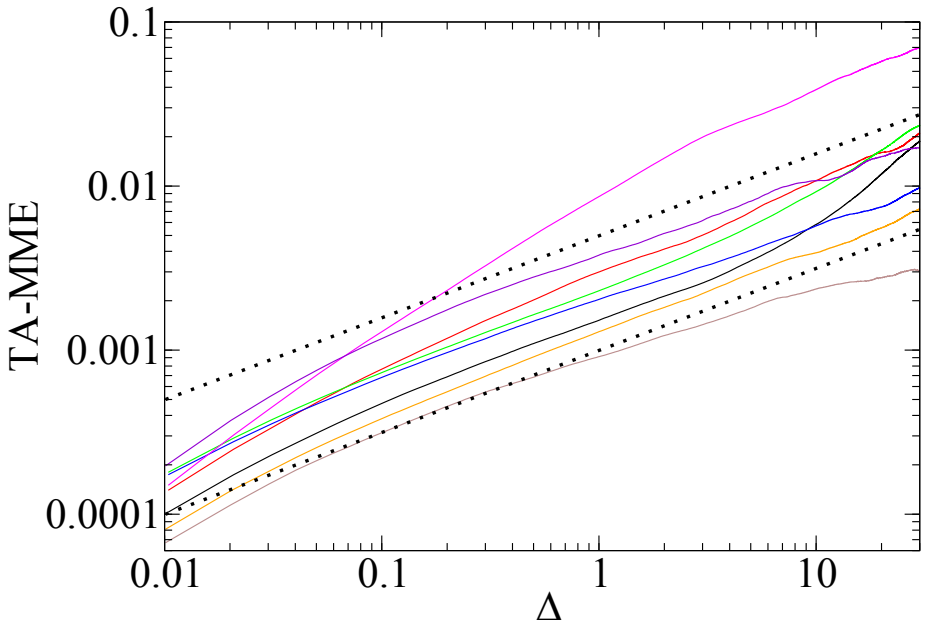

Figure 2.16: Lipid granules diffusing in a yeast cell. Eight trajectories, between 5,515 and 19,393 frames' long. Log-log plot of the time-averaged mean maximal excursion for the eight lipid granules trajectory (continuous lines), and $A_0 t^{0.5}$ (dotted lines). Lag time is in s, and distances in μm.

Once again, we have a scatter in amplitude. The initial slope variation ($0 < t < 0.1$ s) is due to the inaccuracy in the MME estimation when there are only few frames to average, typically less than 10 frames. A greater exponent for MME than for regular moment could be due to an inaccuracy in the fit. However, it may indeed point toward an underlying fBm process.

Those two analysis show a clear sub-diffusive behavior. The time-average results are still difficult to interpret, because the diffusing process is not necessarily ergodic, which is the case of fBm [74] and CTRW [16]. To tackle this problem, we decided to cut the few long trajectories to a large number of short (and independent) trajectories. As the different trajectories were not all recorded at the same frequency (96.5 and 99.1 frames/s), we kept only the greater set (96.5 fps), containing five trajectories, and we split those into 526 short trajectories of 100 steps each. These trajectories are non-overlapping and one may view them as the result of 526 separate observations.

Surprisingly, we retrieve the exponent 0.41 ± 0.01 using the mean square displacement, and the value 0.53 ± 0.02 from the second maximal excursion moment, as shown in Figure 2.17. In this figure, the black dots represent the mean square displacement, the red dots the second maximal excursion moment, for the set of 536 trajectories of 100 steps long.

We repeated this analysis with a set of trajectories of 150 frames (350 trajectories), corresponding to the black and red crosses in Figure 2.17. We observe that the choice of trajectories of 100 steps long has no influence on the value of the anomalous coefficient.

Because one of the trajectory (the magenta line in Figures 2.15 and 2.16) shows a much steeper slope, we excluded it for the rest of the analysis. We did once again the analysis for 445 trajectories of 100 steps long (blue and magenta dots in Figure 2.17), and for 296 trajectories of 150 steps long (blue and magenta crosses in Figure 2.17).

An interesting observation is the following: assuming that the underlying stochastic process is indeed an fBm, equation (2.60) for $\alpha = 0.41$ predicts a value $\alpha' = 0.50$ for the maximal excursion statistics, in quite good agreement with the fitted value. This finding is quite suggestive in favor of fBm as the stochastic process governing the particle motion.

Because the trajectories correspond to different granules, in different cells, we also studied them separately: each trajectory was split into stretches of 100 steps. For each granule, we plotted in Figure 2.18 the displacement and the maximal excursion ratios. They are somewhat noisy, but for each granule the maximal excursion ratio is clearly below the Brownian one (1.49): it ranges between 1.20 and 1.40. The regular moment ratio is slightly above the Brownian value (2), between 1.7 and 2.5. We also plotted the ratio for the whole set of 100 steps (thick lines), which gives approximately the same result as those obtained for individual trajectories. From these ratios, we obtain another clue pointing at an underlying fBm mechanism: the maximal excursion ratio is, on average, below the value for Brownian motion, and the displacement ratio is close to the Brownian value. These maximal excursion ratios are not very precise, but seem to range somewhat above the expected value for fBm with a $\alpha = 0.41$: equation (2.61) gives 1.21 ± 0.02.

The test with the growing sphere shown in Figure 2.19 is, once again, somewhat noisy; however, it clearly shows that the probability to be in a sphere, growing like $t^{\alpha/2}$, attains a constant value. This excludes the possibility that the process corresponds to diffusion on a fractal.

We did not really expect the diffusive mechanism to be diffusion on a fractal, since lipid granules are rather big particles in a cell. A fractal media could hardly exists over more than one order of magnitude.

The above analysis demonstrates that the tools proposed in this section allow us to classify the stochastic process underlying the motion of the measured single particle trajectories of the granules. We observe that the lipid granule motion in a yeast cell shares several distinct features

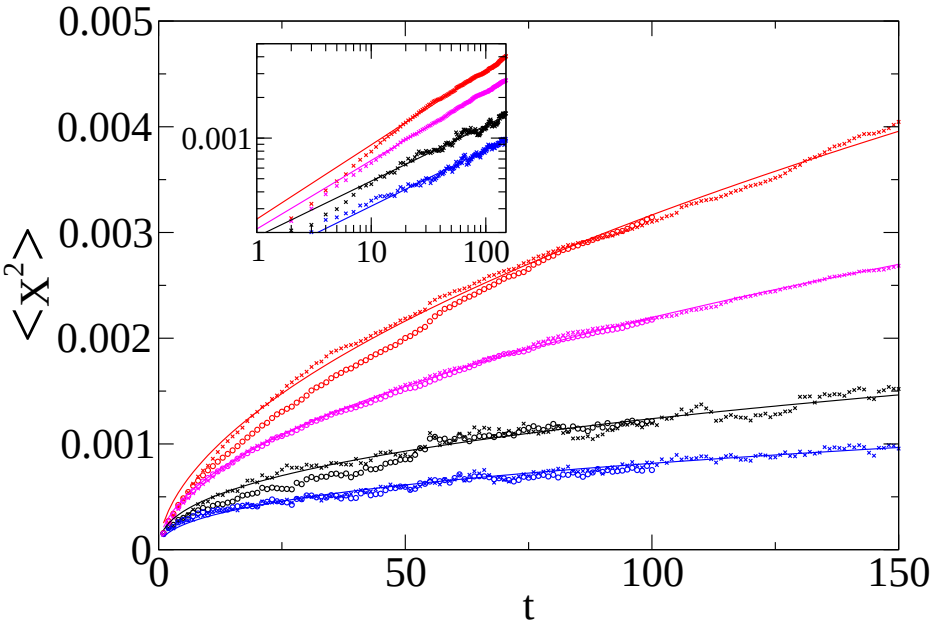

Figure 2.17: Lipid granules diffusion in a yeast cell. 526 subtrajectories of 100 steps extracted from the experimental set of 5 trajectories, which are between $5,515$ and $19,393$ frames' long. Ensemble average of the mean square displacement (black circles) fitted by a power law ($\alpha = 0.41$, black line) and ensemble average of the second maximal excursion moment (red circles), fitted with a power law ($\alpha = 0.55$, red line). We checked that creating 350 trajectories of 150 steps instead of 100 does not change the exponents obtained (crosses instead of circles). Because one of the trajectories has a steeper slope than the others, we repeated the same analysis without this trajectory. The new subset contained 445 trajectories of 100 steps (blue and magenta circles), or 296 of 150 steps (blue and magenta croisses). The power law fit now lead to $\alpha = 0.42$ for the MSD (blue line) and to $\alpha = 0.51$ for the maximal excursion (magenta line). Time is in s and distances in μm. (Inset) log-log plot of the same data.

with the fBm process. Namely, fBm explains the finding of different scaling exponents between the displacement and the maximal excursion second moment, including their actual values connected by equation (2.60). It is also consistent with a Brownian regular moment ratio, and an maximal excursion ratio lower than the Brownian ratio, as shown in Figure 2.18. The recorded data were also shown to be incompatible with diffusion on a fractal.

The question arises: Could CTRW function as a potential mechanism? The scatter between different single trajectories observed in the time-averaged second moments is reminiscent of the weak ergodicity breaking for CTRW sub-diffusion with diverging characteristic waiting time, as studied in the literature [104, 135].

However an alternative explanation may simply be that various cell environments and various granule sizes lead to a scattering of the diffusion coefficient. It should be noted that even between successive recordings the cellular environment may change slightly, influencing the motion of the

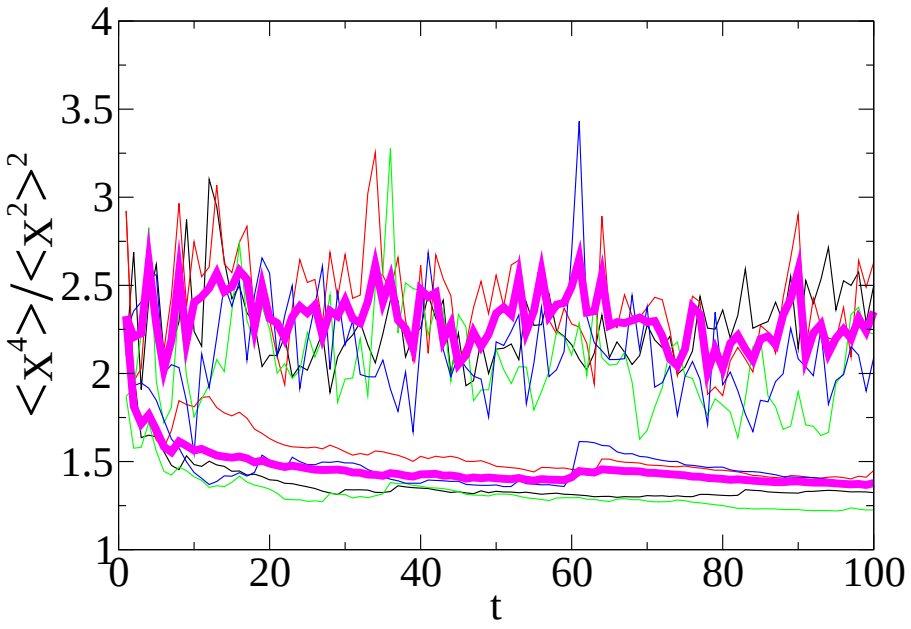

Figure 2.18: Lipid granules diffusing in a yeast cell. Eight trajectories, between 5,515 and 19,393 frames' long. For each trajectory, the displacement ratio (upper curves) and the maximal excursion ratio (lower curves) are plotted, as well as an average over the eight trajectories (thick magenta line). Time is in frame.

observed particle. The CTRW hypothesis, however, is not consistent with the moment ratio test: the expected ratio for $\alpha = 0.4$ would be 3.38 for the regular one, and 2.50 for the maximal excursion one – far above the observed values.

Summary

In this section, we proposed a new way to analyze single trajectories:

- Mean square displacement can be completed with maximal excursion moments, which has a better statistic to extract the anomalous diffusion coefficient and exponent.
- Ratios $\langle X^4 \rangle / \langle X^2 \rangle^2$ are a signature of the microscopical mechanism.
- Growing shell test is a good and simple way to leave out diffusion on fractals.
- For long trajectories, time-average is often used. Slicing a long trajectory in several independent short ones is a simple and efficient trick to retrieve a "classic" data set.

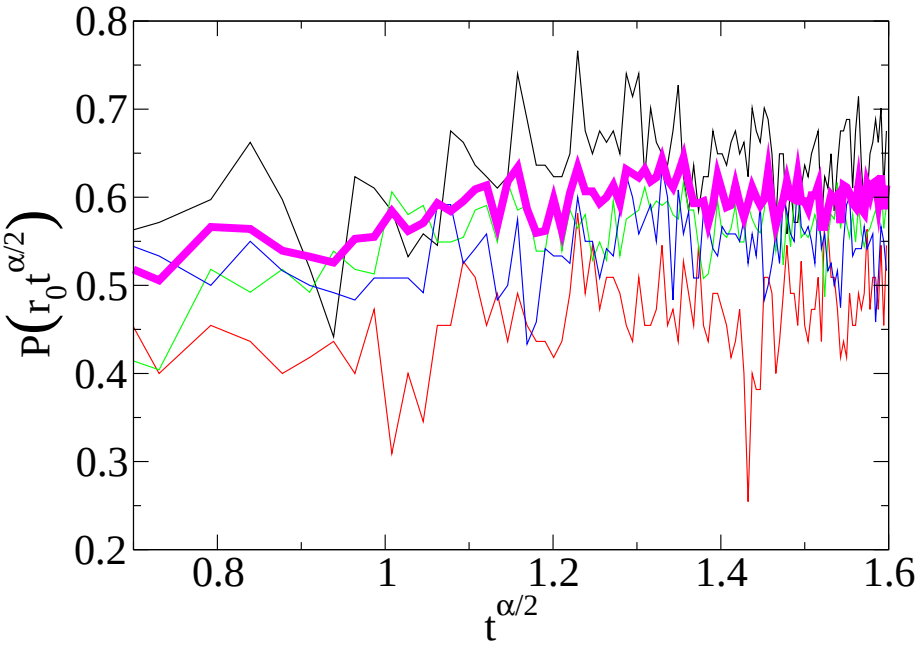

Figure 2.19: Lipid granules diffusing in a yeast cell. Eight trajectories, between 5,515 and 19,393 frames' long. For each trajectory, the probability to be in a growing sphere of radius $r_0 t^{\alpha/2}$ as a function of $t^{\alpha/2}$ is plotted, as well as an average over the eight trajectories (thick magenta line).

2.2.5 Extensive analysis of lipid granules

After analyzing the first lipid granule data set, we suspected the diffusive mechanism to be a fBm. To assess this assumption, we gathered more data, on a wider time range.

The set was extended with more trajectories obtained through optical microscopy, and with a wider time span, since we added data extracted by an optical tweezers setup. Those data have been gathered by Christine Selhuber-Unkel and Lene Oddershede at Niels Bohr Institute in Copenhagen (Denmark). The protocols are detailed in Appendix 1.

We note that during the optical tweezers tracking the granule particles become increasingly confined in the optical tweezers' trapping potential. The time-averaged mean square displacement of an ergodic process would therefore saturate to a stationary value. CTRW sub-diffusion, as shown in our data, is non-stationary and shows a power-law growth of the time-averaged mean square displacement.

Short-time analysis

We first focus on the short-time behavior, when the granule motion is recorded in an optical tweezers setup. In the experiment the trap is initially centered onto the granule such that no force is exerted. When the granule starts to move away from the trap center it experiences a restoring Hookean force [170]. The measured time-averaged mean square displacement $\langle r^2 \rangle_t(\Delta, T)$ from two different cell stages are shown in Figure 2.20. The voltage signal from the photodiode is

directly proportional to the particle position relative to the trap center. But imprecise knowledge of the optical properties in the cell hampers an accurate conversion to absolute length. In Figure 2.20 a distinct turnover is observed from an initial linear growth $\langle r^2\rangle_t \propto \Delta$ to a power-law behavior $\langle r^2\rangle_t \propto \Delta^\beta$ with $\beta \simeq 0.15 \ldots 0.20$.

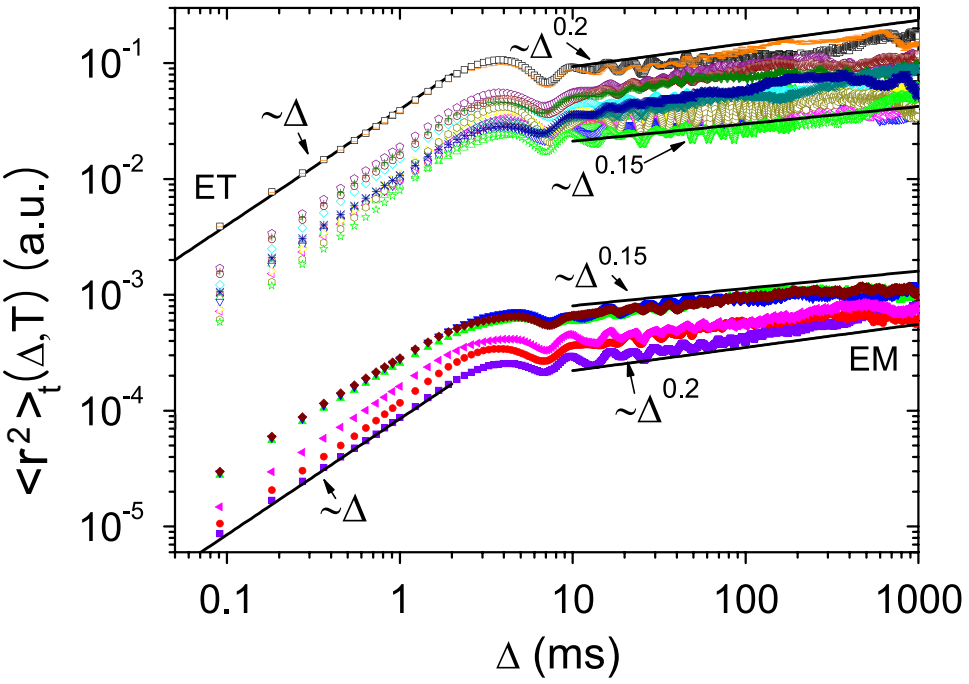

Figure 2.20: Time-averaged mean square displacement from individual trajectories of lipid granules in *S. pombe* in early mitotic (EM) cells (lower curves) and in early telophase (ET) cells (upper curves), measured by optical tweezers. A distinct turnover from $\langle r^2\rangle_t \propto \Delta$ to $\propto \Delta^\beta$ ($\beta \simeq 0.1 \ldots 0.2$) occurs. The two thick lines in the middle show the averages of the ET (red circles) and EM (black squares) data set. The overlaid thick black line is the result of CTRW simulations in a harmonic potential. ET and EM curves are shifted vertically.

The size of the lipid granules is about 300 nm. In addition to the amplitude scatter between different trajectories expected from CTRW theory[50, 51, 104], the fluctuations of the data in Figure 2.20 are due to natural granule size variations and different optical conditions for each trajectory. We note that oscillations around the turnover may arise for sub-diffusion in an underdamped medium [48, 49].

An analogous data set has already been analyzed previously through *ensemble* average, and anomalous diffusion was observed with $\alpha \simeq 0.80 \ldots 0.85$ in a range of 0.1 to 3 msec [191]. In contrast, in Figure 2.20 the initial behavior of $\langle r^2\rangle_t$ corresponding to this time range does not exhibit any apparent anomaly but scales like $\langle r^2\rangle_t \propto \Delta$. At longer lag time Δ due to the trap force one would expect $\langle r^2\rangle_t$ to saturate to a stationary thermal value; instead, the regime $\langle r^2\rangle_t \propto \Delta^\beta$ appears. This behavior is consistently observed in different cell stages (Figure 2.20).

Such a peculiar behavior is consistent with CTRW sub-diffusion: For free motion one finds $\langle r^2\rangle_t \propto \Delta/T^{1-\alpha}$, whose Δ scaling is independent of α, while the corresponding ensemble average follows $\langle r^2\rangle \propto \Delta^\alpha$ [104, 135]. Under confinement a turnover to the power-law $\langle r^2\rangle_t \propto (\Delta/T)^{1-\alpha}$ occurs [50, 51, 146]. This second power-law regime is terminated when Δ approaches the total

measurement time T, causing a dip in $\langle r^2 \rangle_t$ back to the plateau of the ensemble average. The observed characteristic turnover behavior is intimately connected to CTRW ageing and ergodicity breaking [11, 16, 40, 131, 167, 235]. Correlated motion [75] could not explain the observed behavior $\langle r^2 \rangle_t \propto \Delta$ turning over to $\langle r^2 \rangle_t \propto \Delta^\beta$.

The following features point toward CTRW sub-diffusion as the stochastic mechanism for the granule motion at short times:

(i) The time average $\langle r^2 \rangle_t$ initially scales linearly with Δ, albeit the ensemble average shows sub-diffusion ($\langle r^2 \rangle \propto \Delta^\alpha$) in comparable time ranges.

(ii) At longer times a turnover to the power-law $\langle r^2 \rangle_t \propto \Delta^\beta$ occurs instead of the convergence to a plateau, which would necessarily occur for an ergodic process; the anomalous diffusion exponent $\alpha \simeq 0.80 \ldots 0.85$ observed in the ensemble average is consistent with the exponents $\beta \simeq 0.15 \ldots 0.20$ observed in Figure 2.20 based on the relation $\beta = 1 - \alpha$ [50, 51, 146], as well as the slopes of the long time data (see below).

We can complement this analysis by evaluating $\langle r^2 \rangle_t$ as function of the total measurement time T: no ageing is observed, as shown in Figure 2.21, contrasting the scaling $\langle r^2 \rangle_t(T) \propto T^{\alpha-1}$ predicted for CTRW sub-diffusion with diverging mean waiting time [50, 51, 104, 135].

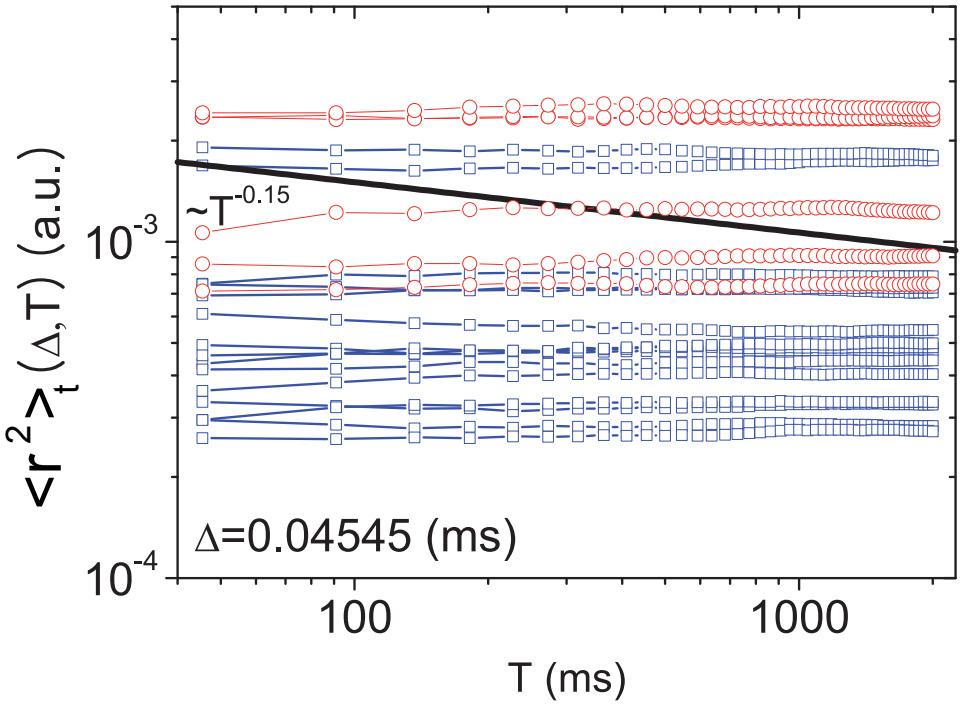

Figure 2.21: Time-averaged mean square displacement for two set of lipid granules in *S. pombe*, either in early mitotic cells (red curves) or in early telophase (blue curves), as a function of measurement time T. The overlaid thick black line is the result of CTRW simulations in a harmonic potential for $\alpha = 0.85$.

If we look directly at the trajectories, we observe a behavior that is neither an fBm nor a CTRW. Figure 2.22 shows a typical trajectory along a given coordinate ($x(t)$), and a simulation of a trajectory with the same anomalous coefficient $\alpha = 0.8$ for a fBm or a CTRW. fBm

leads to an antipersistent motion, really noisy around a stationary position. CTRW consists on very pronounced stalling events. Experimental trajectory is somehow a mix between those two mechanisms: some "jumps" with small but noisy excursions during the waiting times.

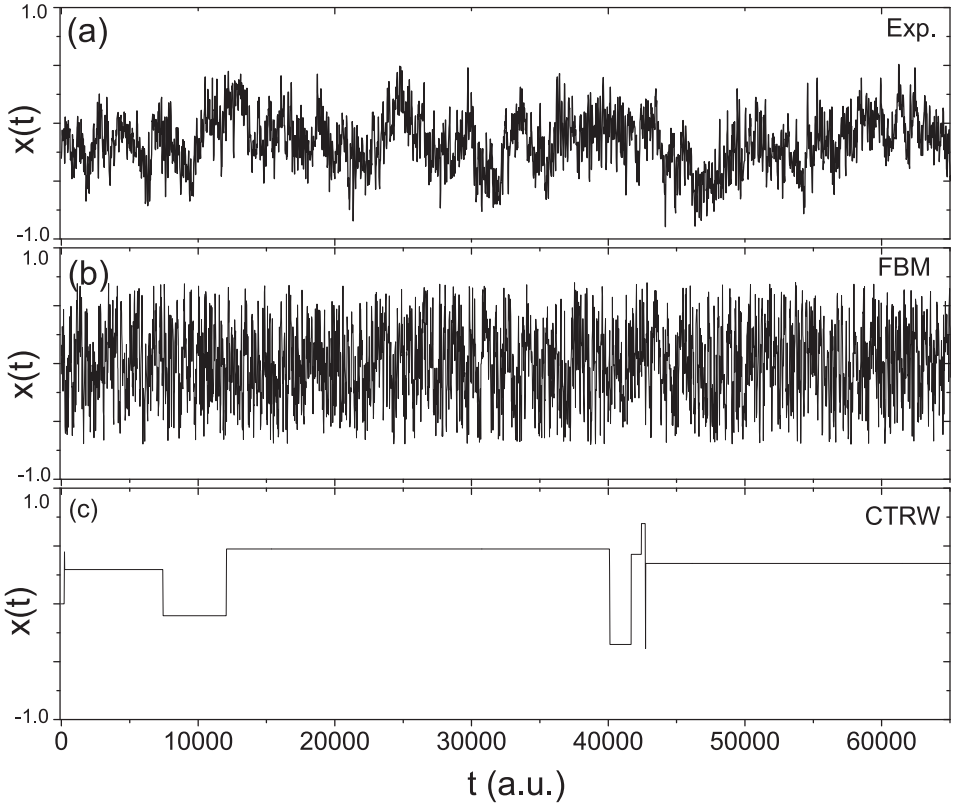

Figure 2.22: Time series $x(t)$ of the experimentally recorded lipid granule motion in *S. pombe* (a), and from simulations of fBm (b) and CTRW (c) with $\alpha = 0.2$.

At short time, we thus have a motion that shares some features with a sub-diffusive CTRW, but not all of them. The exact mechanism is probably an intrication of the "pure" CTRW and of the "pure" fBm model.

Longer time analysis

For longer time, we focus on motion of the lipid granules recorded by video particle tracking. Figure 2.23 shows the time-averaged mean square displacement as a function of lag time Δ. Initially the slope is around $\alpha \simeq 0.8$ or slightly below, consistent with the short time data.

Several of the curves turn to a gentler slope at around $t \sim 100$ ms, some curves eventually switch to normal diffusion ($\alpha = 1$) at $t \sim 100$ s. The regime change occurs for $\langle r^2 \rangle_t \sim 10^{-3} \mu\text{m}^2$ leading to a typical distance of 30 nm: it is not related to cell membrane confinement effect, since the cell is approximately 4 μm×12 μm. But it could be linked to the typical size of the available space for a diffusing object of 300 nm diameter (lipid granule) trapped in a cytoskeleton "cage".

Significant deviations are observed, which, for a living systems, is not surprising. Both the granule size and the materials properties of the cellular environment may change on these time scales, for instance depolymerization or repolymerization of the cytoskeleton.

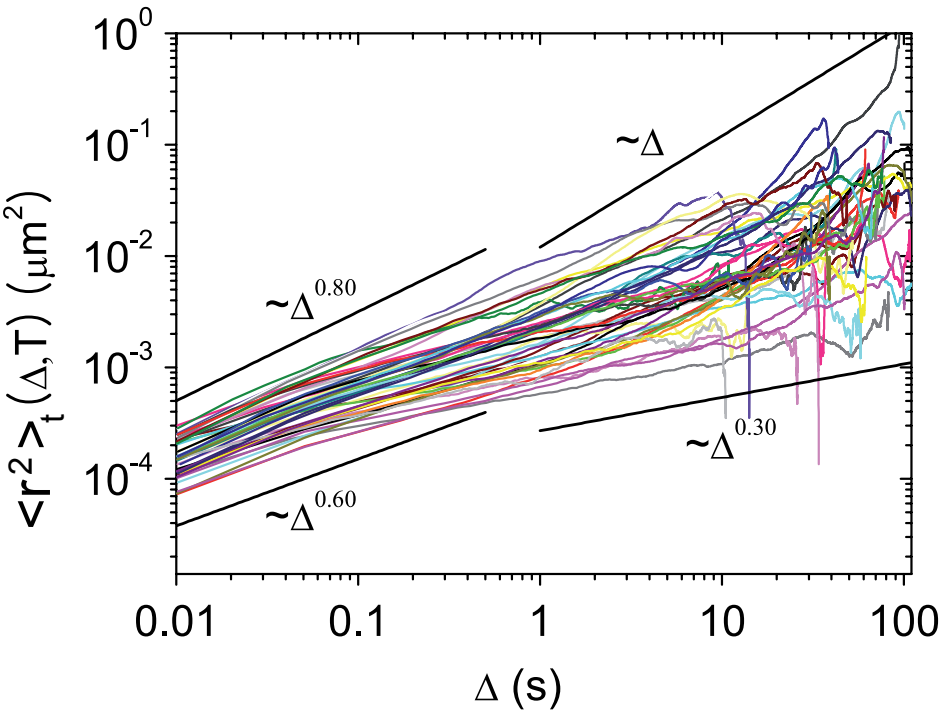

Figure 2.23: Time-averaged mean square displacement for lipid granules in *S. pombe* (cells in interphase) from video tracking data. An initial slope around $\alpha \simeq 0.8$ is found, turning over to a gentler slope. Several trajectories later exhibit a pseudo-Brownian behavior $\alpha \simeq 1.0$.

If we plot the displacement and the maximal excursion ratios, we observed that the value are not consistent with a CTRW. The displacement ratio (upper curves) are really noisy, and slightly above the expected value for a Brownian motion in two dimensions (2). The maximal excursion ratio (lower curves) slowly decrease, but are clearly under the expected value for a Brownian motion (1.49), which is a characteristic of fBm.

At longer times the motion is best described by sub-diffusive fBm, although a conclusive statement in this time range is more difficult due to the fact that cellular processes appear to be superimposed on the motion.

Physically, the CTRW-like motion may be associated with the interaction between granules and the semiflexible filaments of the cytoskeleton similar to the observations in [228]. While the behavior becomes more erratic in the long time data a turnover to a gentler slope is observed before a final increase to normal diffusion. This behavior may be connected to the visco-elastic properties of the complex cellular environment. The identification of fBm as stochastic mechanism is consistent with conclusions in [207, 226].

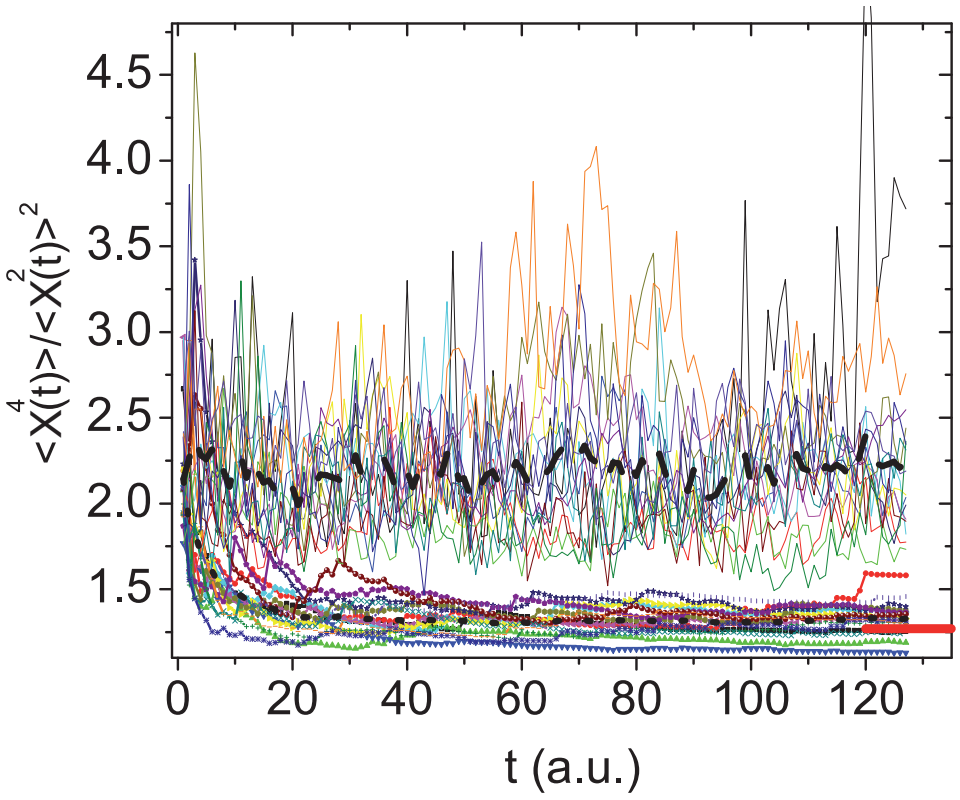

Figure 2.24: Displacement (upper curves) and maximal excursion (lower curves) ratio for lipid granules diffusing in *S. pombe* recorded by video tracking. Each curve correspond to a given lipid granule, and the overlaid black thick line stand for the average over the different lipid granules. The expected Brownian values are 2 for the displacement ratio, and 1.49 for the maximal excursion ratio. The experimental displacement ratio seems to be above the Brownian value, while the experimental maximal excursion ratio is under the expected value for a Brownian motion.

Quick summary

- At short time, the lipid granules share a lot of CTRW characteristics.
- At longer time, the lipid granules seem to evolve according to a fBm.
- Neither model explain completely experimental observations: none of the model of anomalous diffusion studied here cannot explain all features observed for lipid granules.

Conclusion

In this section we were looking for realistic observables to discriminate between several sub-diffusive mechanisms, namely diffusion on fractals, fBm and CTRW. We showed with simulated data sets that we were able to discriminate – "blindly" – those three mechanisms.

We also showed that the maximal excursion statistic was really useful for limited data sets. The fit accuracy to extract the anomalous diffusion exponent α or the diffusion coefficient K_α is better using second maximal excursion moment than using mean square displacement.

At last, when we applied our method to an experimental data set, we observed first that with a limited data set, we could think that lipid granules were evolving according to a fBm model. But after extensive analysis, over a large time span, we were forced to admit that none of our three simple model could explained the variety of behavior observed. The models considered so far do not allow to explain all the features observed in experimental data.

We still gathered some informations:

- We clearly observe a sub-diffusive behavior for lipid granules in *S. pombe* cells ;
- Several sub-diffusive regimes have been observed through video tracking, indicating that the geometry could have two levels (a cytoskeleton "cage", and a "cage" network) ;
- the ratio analysis show a clear antipersistent behavior, since the maximal excursion is clearly below the expected value for a Brownian motion.

2.3 Further models: starting with microscopical facts

During the previous section, we analyzed extensively an experimental data set of lipid granules diffusing in yeast cell. We observed that the lipid granules motion shared some features with CTRW (short time linear time-average MSD), and other with fBm (anti-persistent effects on the maximal excursion ration), with in particular distincts regimes at short times and at long times. This is in strong contrast with the three sub-diffusive models used so far: we always assumed that the process was scale-invariant.

So far, we studied classical diffusion model, and we tried to see if experimental data were in agreement with one of those models. We will now make the way back: we have identified key features of an experimental data set of lipid granules, we will in this section study successively three other mechanisms that illustrate at least one of this prominent features. Since we have also shown in the first section of this chapter that first-passage observables were very good candidates to discriminate between several sub-diffusive mechanism, we will focus on the MFPT as well as on the MSD.

The first sub-section will focus on persistent random walks: we will analyze how a very short memory (one step only) can affect a discrete random walk. We will perform an extensive theoretical analysis of this model, in order to extract the different regimes of the MFPT, the pseudo-Green functions as well as the mean square displacement. We will see that this model could lead, at times shorter than the memory length, to an apparent sub-diffusion, while normal diffusion is recovered at longer times.

The second sub-section will be dedicated to the statical cage model: we will assume that the random walker evolves in an ordered porous media. This model aim to mimic a non-fractal actin network: only two scales exist, the cage, and the cage network. The links between two neighbor cages are small compared to the cage size, and the cages are nodes of a regular lattice. For a very big diffusive particle, like lipid granules in yeast cells, we could indeed assume that the cytoskeleton is only seen on few scales, and is not fractal anymore. For this model, we will mainly focus on the mean square displacement.

At last, the third sub-section will assess the impact of crowding on MFPT. If we consider a random walker evolving in a medium where several hard-core obstacles move, we could expect to

find back a CTRW with a "natural" cut-off. Indeed, the obstacles surround the walker forming a dynamical cage. The walker have to wait that the obstacles leave a free space before moving. We could thus have our two scales here: a first behavior when the random walker is surrounded by obstacles, and the second behavior when he finally manage to exit a given dynamical cage. We will see that this model also leads to an anti-persistent random walk.

2.3.1 Persistent random walks

We observed in the previous section that lipid granules had a maximal excursion ratio under the expected value for a Brownian motion. This feature, shared by fBm, means that the longest excursion is somehow less dispersed for lipid granules than for a classical Brownian motion. We could attribute this effect to an anti-persistence behavior, like for sub-diffusive fBm. We will in this sub-section study the effect of anti-persistence, modelized by a short-term memory.

The results on first-passage properties we have obtained in the first chapter are valid for Markovian scale-invariant random walks [66]. At larger scales however, most examples of walkers – even if random – have at least short range memory skills and show persistent motions, as is the case for bacteria [32] or larger organisms [17], which cannot be described as Markovian scale-invariant processes. The study of persistent random walks has actually proved to be important in various fields such as neutron or light scattering [35, 139, 243]. In this context exact results have been derived that characterize the diffusion properties of persistent walks in infinite space [85, 95, 243], or mean return times in bounded domains [21, 35, 63, 139]. The question of determining first-passage properties of persistent walks has however remained unanswered so far.

"Persistence" will here mean that at each step, all orientations are not equivalent for the random walker: the probability to choose a given orientation depends on the past steps. For a fractional Brownian motion, the correlation run over all previous steps since $t = -\infty$. We will here study a simpler model, where the probability to choose an orientation only depends on the last step.

We will first show how to compute the MFPT in a confined media for a persistent random walker, namely for a walker that recall the last jump direction. Since the analytical result is rather complex, we will propose some approximations for the useful regimes: small and large persistence parameter and large distances. We will then go a step further and derive analytically the pseudo-Green functions for this problem. These functions can be used to obtain other first-passage observables using the result of the first chapter. One has to take care that the random walk alone is non Markovian here: only the joint process (position, last jump direction) is Markovian. At last, we will focus on the mean square displacement for a persistent random walk, in order to compare somehow this model with the results obtained for lipid granules.

MFPT of a persistent random walk

We will assume that the persistent random walk occurs in an usual Euclidian space. For the simplicity sake, we will use a discrete description with an Euclidian lattice.

We consider a cubic lattice in d dimensions, of size $(X_1, X_2, \ldots, X_d)$, with periodic boundary conditions. $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d)$ is an orthogonal base of this lattice. The random walker starts at $\mathbf{r}_0$ toward a given direction $\mathbf{e}_i$, and at each step, he has a probability p_1 to continue in the same direction, p_2 to go in the opposite one, and $p_3 = (1 - p_1 - p_2)/(2d - 2)$ to choose one of the orthogonal directions. The Brownian random walk is equivalent to $p_1 = p_2 = p_3 = 1/(2d)$. Following reference [85], we note $p_1 = p_3 + \epsilon$ and $p_2 = p_3 - \delta$.

If we consider the ensemble-averaged mean first-passage time $\langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}, \mathbf{e}_i)$ to reach a target located in $\mathbf{r}_T$, starting from $\mathbf{r}$ with an initial orientation along the $\mathbf{e}_i$ direction, this quantity satisfies a backward Kolmogorov equation:

$$\langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}, \mathbf{e}_i) = p_1 \langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}+\mathbf{e}_i, \mathbf{e}_i) + p_2 \langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}-\mathbf{e}_i, -\mathbf{e}_i) + p_3 \sum_{\pm \mathbf{e}_j \in \mathcal{B}, j \neq i} \langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}+\mathbf{e}_j, \mathbf{e}_j) + 1 \quad (2.67)$$

The definition of the "last" step direction is a bit delicate for the starting site $\mathbf{r}$, but for a persistent random walker, we need to know the initial random walker orientation to determine the next step. We will afterward be able to adapt our result to any convention chosen for the initial orientation, by combining the results for different $\mathbf{e}_i$.

After some computations detailed in Appendix 2, and a simplification in the Fourier space, we can obtain a simple expression. One can note that so far, the MFPT is defined up to a constant. Indeed, equation (2.67) is satisfied by $\forall \lambda \in \mathbb{R}/\ \langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}, \mathbf{e}_i) + \lambda$. We can set this constant by choosing λ so that $\langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}_T, \mathbf{e}_i) = 0$:

$$\langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}_S, \mathbf{e}_S) = \sum_{\mathbf{q} \neq \mathbf{0}} \left(\frac{1 - (\epsilon + \delta) e^{-2\imath\pi \mathbf{q}.\mathbf{e}_S}}{1 + \epsilon^2 - 2\epsilon \cos(2\pi \mathbf{q}.\mathbf{e}_S) - \delta^2} \frac{1 - e^{2\imath\pi \mathbf{q}.(\mathbf{r}_S - \mathbf{r}_T)}}{1 - 2p_3 \sum_{\mathbf{e}_j \in \mathcal{B}} \frac{\cos(2\pi \mathbf{q}.\mathbf{e}_j) - (\epsilon + \delta)}{1 + \epsilon^2 - 2\epsilon \cos(2\pi \mathbf{q}.\mathbf{e}_j) - \delta^2}} \right) \quad (2.68)$$

We can check that for a Brownian motion ($\epsilon = \delta = 0$ and $p_3 = 1/(2d)$), we retrieve the classical result [111], where $\mathbf{e}_S$ does not play any role:

$$\langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}_S, \mathbf{e}_S) = \sum_{\mathbf{q} \neq \mathbf{0}} \frac{1 - e^{2\imath\pi \mathbf{q}.(\mathbf{r}_S - \mathbf{r}_T)}}{1 - \frac{1}{d} \sum_{\mathbf{e}_j \in \mathcal{B}} \cos(2\pi \mathbf{q}.\mathbf{e}_i)} \quad (2.69)$$

If $d = 1$, equation (2.69) can be simplified. We will note $p = \epsilon + 1/2$ the probability to continue in the same direction, and $1 - p = 1/2 - \delta$ the probability to go back:

$$\begin{aligned} \langle \mathbf{T}(r) \rangle &= \frac{1}{2p} \sum_{k=1}^{X-1} \frac{1 - e^{2\imath\pi k r/X}}{1 - 2(1-p) \frac{\cos(2\pi k/X)}{1 - (2p-1)\cos(2\pi k/X)}} \\ &= \frac{2p-1}{2p} X(1 - \delta(r)) + \frac{1-p}{p} r(X - r), \end{aligned} \quad (2.70)$$

where $r = \|\mathbf{r}_T - \mathbf{r}_S\|$, and $\delta(r)$ is the Kronecker delta function ($\delta(r) = 1$ if $r = 0$, 0 else). We can check that the condition $\langle \mathbf{T}(r = 0) \rangle = 0$ is satisfied.

Some comments are in order:

(i) First of all, we retrieve the Brownian motion ($p = 1/2$):

$$\langle \mathbf{T}(r) \rangle = r(X - r) \quad (2.71)$$

(ii) For large volume ($X \gg 1$), and for large distances ($r \gg 1$), we retrieve a Brownian motion, with an effective diffusion coefficient $D_{\text{eff}} = (1-p)/p$, as obtained by reference [96]:

$$\langle \mathbf{T}(r) \rangle \underset{\substack{X \to \infty \\ r \to \infty}}{\sim} \frac{1-p}{p} r(X - r) \quad (2.72)$$

(iii) For a highly persistent random walk ($1-p=\epsilon \ll 1$), we obtain $\langle \mathbf{T}(r)\rangle \sim X+\epsilon r(X-r)$. This almost constant result ($\epsilon \ll 1$) is in fact due to the average over the initial orientation. If we fix this orientation to $-\mathbf{e}_x$, we obtain, using (2.68):

$$\langle \mathbf{T}(r|-\mathbf{e}_x)\rangle = \frac{2p-1}{p}r + \frac{1-p}{p}r(X-r) \simeq r + \frac{1-p}{p}r(X-r) \tag{2.73}$$

For $p \gg X/(X+1)$, we have a ballistic behavior $\langle \mathbf{T}(r|-\mathbf{e}_x)\rangle \simeq r$. But as soon as X is great enough, this regime disappear since $p<1$. For $1-p \ll 1$, we can consider that we have a Brownian behavior with an effective diffusion coefficient $D_{\text{eff}}=(1-p)/p$, and an effective volume $X_{\text{eff}}=X-1+1/(1-p)$. For $r \ll X_{\text{eff}}$, we then have a simili-ballistic behavior, but with a speed $D_{\text{eff}}X_{\text{eff}}>1$: $\langle \mathbf{T}(r|-\mathbf{e}_x)\rangle \simeq D_{\text{eff}}X_{\text{eff}}r$.

(iv) For a highly antipersistent random walk, the previous result still apply, but now $p \ll 1$. We always have a Brownian behavior with $D_{\text{eff}} \simeq 1/p$.

For $d \geq 2$, equation (2.68) cannot be simplified anymore. We can test this equation numerically: as shown in Figure 2.25, this result fits perfectly numerical simulation (here for $d=2$). Here, $\mathbf{r}_s=(x_S,0)$, $\mathbf{r}_T=(0,0)$, $\mathbf{e}_S=(1,0)$, with $x_S \in [0,9]$. The theoretical result of equation (2.68) (red line) fits exactly the simulation (black circles).

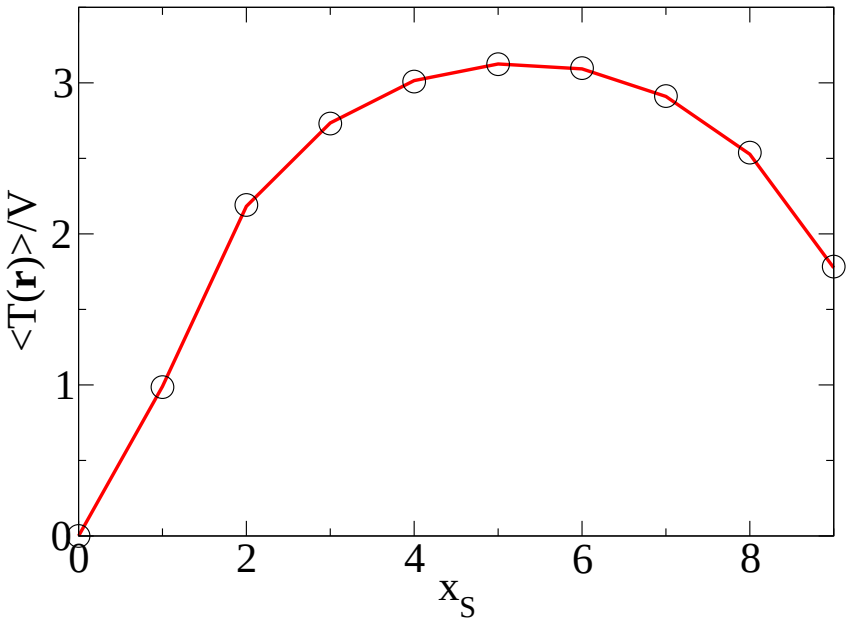

Figure 2.25: Mean first-passage time for a 10×10-large periodic network, where $\mathbf{r}_s=(x_S,0)$, $\mathbf{r}_T=(0,0)$, $\mathbf{e}_S=(1,0)$, with $x_S \in [0,9]$. Here $\epsilon=-0.04$ and $\delta=0.32$. The black circles stand for the numerical simulation, the red line for the theoretical result of equation (2.68).

This approach gives a quite simple (and exact) result, and will be extrapolated for reflective boundaries below, using persistent pseudo-Green functions (see Appendix 2 for details).

Some MFPT approximations

Equation (2.68) is exact, but not easy to deal with. We can propose some approximations of this mean first-passage time, in order to have more tractable expressions.

First of all, if we consider that $\epsilon \ll 1$, namely that the persistence is very low, we get, after averaging over all $\mathbf{e}_S$:

$$\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S) \underset{\epsilon \ll 1}{\simeq} A(\epsilon,\delta)V + \frac{1-\delta-\epsilon}{1+\delta+\epsilon}\,\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)|_{\epsilon=0,\delta=0}\,, \tag{2.74}$$

where $A(\epsilon,\delta)$ is a constant defined by:

$$A(\epsilon,\delta) = \frac{\delta-\epsilon}{1-\epsilon+\delta} + \frac{2\epsilon}{(1+\epsilon+\delta)(1-\epsilon+\delta)}B_d, \tag{2.75}$$

and where B_d is the limit of:

$$B_d = \lim_{V\to\infty}\frac{1}{V}\sum_{\mathbf{q}\neq\mathbf{0}}\frac{\frac{1}{d}\sum_{\mathbf{e}_j\in\mathcal{B}}\left(1-\cos(2\pi\mathbf{q}.\mathbf{e}_j)\right)^2}{\left(\frac{1}{d}\sum_{\mathbf{e}_j\in\mathcal{B}}1-\cos(2\pi\mathbf{q}.\mathbf{e}_j)\right)^2} \tag{2.76}$$

For $d=2$, $B_2 \simeq 1.36$.

We note that we have here an effective diffusion coefficient $D_{\text{eff}} = (1+\delta+\epsilon)/(1-\epsilon-\delta)$, that will later be related to the mean field approximation. We retrieve in equation (2.74) the expected result, namely the Brownian behavior multiplied by $1/D_{\text{eff}}$, but up to a constant. This constant being proportional to V, it does not disappears when $V\to\infty$. This could be related to the "residual" mean first-passage time described in reference [208]: when doing a persistent random walk, the walker "miss" some neighbors in his exploration, by keeping the same direction for several steps.

As shown in Figure 2.26, this approximation works really well for ϵ up to 0.3. The continuous lines here stand for, from up to down, $\epsilon = 0$ (red), 0.1 (green), 0.2 (blue) and 0.3 (magenta). The black dashed lines stand for the approximation of equation (2.74), from up to down, for $\epsilon = 0$, 0.1, 0.2 and 0.3.

What happens if $\epsilon \to 1$? If $1-\epsilon \ll 1$, we can approximate equation (2.68) after a small reorganization:

$$\langle\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle = \frac{V(\delta-\epsilon)}{1-\epsilon+\delta} + \frac{2\epsilon}{(1+\epsilon+\delta)(1-\epsilon+\delta)}\sum_{\mathbf{q}\neq\mathbf{0}}\frac{1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}}{\frac{1}{d}\sum_{\mathbf{e}_j\in\mathcal{B}}\frac{1}{1+\frac{(\epsilon-1)^2-\delta^2}{2\epsilon(1-\cos(2\pi\mathbf{q}.\mathbf{e}_j))}}} \tag{2.77}$$

We here focus on a $X\times X$ square ($d=2$), and obtain, after some math detailed in Appendix 2:

$$\begin{aligned}\langle\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle \underset{\epsilon\to 1}{\simeq}\ & \frac{X^2(\delta-\epsilon)}{1-\epsilon+\delta} + \frac{2\epsilon}{(1+\epsilon+\delta)(1-\epsilon+\delta)} \\ & \times\Big((X-1)(X+3) - \left(X\delta_{(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_x}+1\right)\left(X\delta_{(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_y}+1\right)+4\Big) \\ & +\frac{1-\epsilon-\delta}{1+\epsilon+\delta}\Big((X-1)\frac{X^2-1}{6} + (\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_x\,(X-(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_x) \\ & +(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_y\,(X-(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_y)\Big)\end{aligned} \tag{2.78}$$

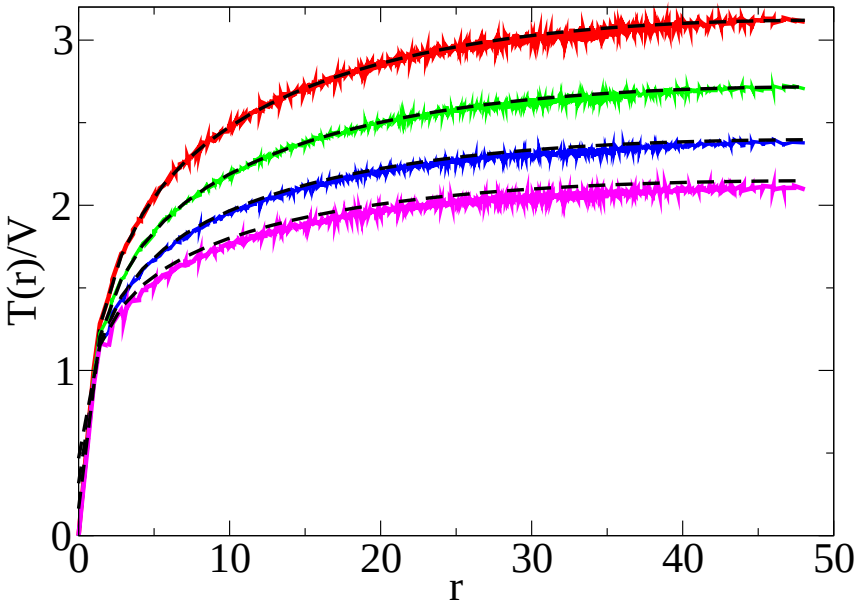

Figure 2.26: Mean first-passage time for a 70×70-large periodic network, with $\delta = 0$, and for $\epsilon = 0$ (red curve), 0.1 (green curve), 0.2 (blue curve) and 0.3 (magenta curve). The black dashed lines stand for the approximation of equation (2.74), from up to down, for $\epsilon = 0$, 0.1, 0.2 and 0.3.

This expression is simply polynomial in X, and shows that the mean first-passage time diverge when $\epsilon \to 1 + \delta$. The divergence is due to the network considered: when $\epsilon = 1 + \delta$, the random walker can't reorientate, and stay always on the same line. On an Euclidian network with periodic boundaries, closed orbits exists. If the random walker does not start on an orbit comprising the target, it stays on this orbit forever, and can never reach the target, leading to an infinite mean first-passage time. If the target is in the starting orbit, we find back an unidimensional problem, treated above.

At last, we will focus on what happens at large distances, for $d \geq 2$. We can approximate the discrete sum of equation (2.69) with an integral when all $X_i \to \infty$. We consider the case where $\mathbf{r} = \mathbf{r}_T - \mathbf{r}_S$ is great enough, namely when $r = \|\mathbf{r}\| = \sqrt{\sum_i (\mathbf{r}.\mathbf{e}_i)^2} \gg 1$. After some computations detailed in Appendix 2, we obtain:

$$\langle \mathbf{T}(r) \rangle \simeq V \left(\left(\frac{\delta - \epsilon}{1 - \epsilon + \delta} + \Delta \right) (1 - \delta(r)) + \frac{1 - \epsilon - \delta}{1 + \epsilon + \delta} \left(G_{0,d}(0) - G_{0,d}(r) \right) \right), \tag{2.79}$$

where $G_{0,d}$ is the infinite Green function in d dimensions. Δ is a correction depending on ϵ and δ. We can note that the second term is the one predicted in the mean field approximation: we obtain the first-passage time for a random walker diffusing with an effective diffusion coefficient:

$$D_{\text{eff}} = \frac{1 + \epsilon + \delta}{1 - \epsilon - \delta} \tag{2.80}$$

We will show later on that this effective coefficient is indeed the one predicted by a mean field approximation. We retrieve in equation (2.79) the expected result, namely the Brownian behavior multiplied by $1/D_{\text{eff}}$, but up to a constant.

We can estimate the Δ correction, as shown in Appendix 2:

$$\Delta \simeq \frac{(1+\epsilon^2-\delta^2)}{(1+\epsilon+\delta)(1-\epsilon+\delta)} \int_{[0,2\pi]^d} \frac{d\mathbf{q}}{(2\pi)^d} \times \left(\frac{1}{1-\frac{(\epsilon-1)^2-\delta^2}{d}\sum_{\mathbf{e}_j\in\mathcal{B}} \frac{\cos(\mathbf{q}.\mathbf{e}_j)}{1+\epsilon^2-2\epsilon\cos(\mathbf{q}.\mathbf{e}_j)-\delta^2}} - \frac{(\epsilon-1)^2-\delta^2}{1+\epsilon^2-\delta^2} \frac{1}{1-\frac{1}{d}\sum_{\mathbf{e}_j\in\mathcal{B}}\cos(\mathbf{q}.\mathbf{e}_j)} \right) \tag{2.81}$$

The simulation being discrete, we also obtain in equation (2.79) a "residual" MFPT, the first term, that do not vanish even in the large volume limit. This is quite a surprise, since the random walk is here discrete, but also a nearest neighbor walk. This residual MFPT vanishes for $\delta = \epsilon = 0$ (Brownian limit).

Persistent pseudo-Green functions

To get more insight on first-passage properties, we can compute directly pseudo-Green functions for persistent random walks. Those pseudo-Green functions can afterward be used to get splitting probabilities, mean occupation times or first-passage time densities, as shown in the first chapter.

To do so, we first have to get the propagator, and then deduce pseudo-Green functions, for a discrete network with periodic boundary conditions.

We note $P_{\mathbf{e}_i}(\mathbf{r},t|\mathbf{r}_S,\mathbf{e}_S)$ the conditional propagator, namely the probability to be at the position $\mathbf{r}$ at time t, when the last step have been following the vector $\mathbf{e}_i$, starting from $\mathbf{r}_S$ with orientation $\mathbf{e}_S$. For simplicity sake, we will skip $\mathbf{r}_S$ and $\mathbf{e}_S$ in the notation. To obtain the usual propagator, *i.e.* the probability to be in $\mathbf{r}$ at time t, without information on the last step, one has to sum those conditional propagators over all vectors :

$$P(\mathbf{r},t) = \sum_{\mathbf{e}_i\in\mathcal{B}} \left(P_{\mathbf{e}_i}(\mathbf{r},t) + P_{-\mathbf{e}_i}(\mathbf{r},t)\right) \tag{2.82}$$

The conditional propagator satisfies the master Chapman-Kolmogorov equation:

$$P_{\mathbf{e}_i}(\mathbf{r},t+1) = p_1 P_{\mathbf{e}_i}(\mathbf{r}-\mathbf{e}_i,t) + p_2 P_{-\mathbf{e}_i}(\mathbf{r}-\mathbf{e}_i,t) + p_3 \sum_{\mathbf{e}_j\in\mathcal{B},j\neq i} \left(P_{\mathbf{e}_j}(\mathbf{r}-\mathbf{e}_i,t) + P_{-\mathbf{e}_j}(\mathbf{r}-\mathbf{e}_i,t)\right) \tag{2.83}$$

After a passage in the Laplace-space (Z-transform), and a Fourier transform, we obtain (see Appendix 2 for more details):

$$\widehat{P}_{\mathbf{e}_i}(\mathbf{q},z) = \frac{(1-\epsilon z e^{2\imath\pi\mathbf{q}.\mathbf{e}_i})\widetilde{P}_{\mathbf{e}_i}(\mathbf{q},0) - \delta z e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}\widetilde{P}_{-\mathbf{e}_i}(\mathbf{q},0) + p_3 z(e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i} - (\epsilon+\delta)z)\widehat{P}(\mathbf{q},z)}{1+\epsilon^2z^2-\delta^2z^2-2\epsilon z\cos(2\pi\mathbf{q}.\mathbf{e}_i)} \tag{2.84}$$

We then can determine the Fourier transform of the propagator, $\widehat{P}(\mathbf{q},z)$, by summation over

all $\mathbf{e}_i$:

$$\widehat{P}(\mathbf{q},z)=\frac{\displaystyle\sum_{\mathbf{e}_i\in\mathcal{B}}\frac{\left(1-(\epsilon+\delta)ze^{2\imath\pi\mathbf{q}.\mathbf{e}_i}\right)\widetilde{P}_{\mathbf{e}_i}(\mathbf{q},0)+\left(1-(\epsilon+\delta)ze^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}\right)\widetilde{P}_{-\mathbf{e}_i}(\mathbf{q},0)}{1+\epsilon^2z^2-\delta^2z^2-2\epsilon z\cos(2\pi\mathbf{q}.\mathbf{e}_i)}}{1-2p_3z\displaystyle\sum_{\mathbf{e}_i\in\mathcal{B}}\frac{\cos(2\pi\mathbf{q}.\mathbf{e}_i)-(\epsilon+\delta)z}{1+\epsilon^2z^2-\delta^2z^2-2\epsilon z\cos(2\pi\mathbf{q}.\mathbf{e}_i)}} \tag{2.85}$$

This leads to an explicit expression of $\widehat{P}_{\mathbf{e}_i}(\mathbf{q},z)$.

Knowing the conditional propagator, we can obtain the conditional pseudo-Green functions, defined as:

$$H(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S)=\sum_{t=0}^{\infty}\left(P_{\mathbf{e}_i}(\mathbf{r},t|\mathbf{r}_S,\mathbf{e}_S)-P_{\text{stat}}(\mathbf{r},\mathbf{e}_i)\right), \tag{2.86}$$

where $\mathbf{r}_S$ is the initial position, $\mathbf{e}_S$ the initial "last" direction, and $P_{\text{stat}}(\mathbf{r},\mathbf{e}_i)=\prod_i X_i^{-1}/(2d)$ is the (conditional) stationary probability. We choose for initial condition:

$$\widetilde{P}_{\mathbf{e}_i}(\mathbf{q},0)=\delta(\mathbf{e}_i,\mathbf{e}_S)\exp(-2\imath\pi\mathbf{q}.\mathbf{r}_S) \tag{2.87}$$

, where $\delta(\mathbf{x},\mathbf{y})=1$ if $\mathbf{x}=\mathbf{y}$, and 0 else. We obtain:

$$\begin{aligned}H(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S) \;=\; &\frac{1}{\displaystyle\prod_i X_i}\sum_{\mathbf{q}\neq\mathbf{0}}e^{2\imath\pi\mathbf{r}.\mathbf{q}}\left(\frac{\left((1-\epsilon e^{2\imath\pi\mathbf{q}.\mathbf{e}_i})\delta(\mathbf{e}_i,\mathbf{e}_S)-\delta e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}\delta(-\mathbf{e}_i,\mathbf{e}_S)\right)e^{-2\imath\pi\mathbf{q}.\mathbf{r}_S}}{1+\epsilon^2-\delta^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_i)}\right.\\ &\left.+\frac{p_3(e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}-(\epsilon+\delta))\widehat{P}(\mathbf{q},z=1)}{1+\epsilon^2-\delta^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_i)}\right)\end{aligned} \tag{2.88}$$

where

$$\widehat{P}(\mathbf{q},z=1)=\frac{\dfrac{\left(1-(\epsilon+\delta)e^{2\imath\pi\mathbf{q}.\mathbf{e}_S}\right)}{1+\epsilon^2-\delta^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_S)}e^{-2\imath\pi\mathbf{q}.\mathbf{r}_S}}{1-2p_3\displaystyle\sum_{\mathbf{e}_i\in\mathcal{B}}\frac{\cos(2\pi\mathbf{q}.\mathbf{e}_i)-(\epsilon+\delta)}{1+\epsilon^2-\delta^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_i)}} \tag{2.89}$$

The $\mathbf{q}=\mathbf{0}$ term has been removed in the summation, to take in account the $-P_{\text{stat}}(\mathbf{r},\mathbf{e}_i)$ term of equation (2.86).

We can check that we retrieve Brownian motion if we choose $\epsilon=\delta=0$:

$$\sum_{\pm\mathbf{e}_i\in\mathcal{B}}H(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S)=\frac{1}{\displaystyle\prod_i X_i}\sum_{\mathbf{q}\neq\mathbf{0}}\frac{e^{2\imath\pi\mathbf{q}.(\mathbf{r}-\mathbf{r}_S)}}{1-\dfrac{1}{d}\displaystyle\sum_{\mathbf{e}_i\in\mathcal{B}}\cos(2\pi\mathbf{q}.\mathbf{e}_i)} \tag{2.90}$$

As shown in Appendix 2, we can use the pseudo-Green function given by equation (2.88) to obtain an equivalent expression of the exact expression (2.68) of the MFPT.

All the results we have presented previously that use pseudo-Green functions can be transposed here. In particular, we can use them to treat the reflective boundary conditions case. The calculus are presented in Appendix 2. Figure 2.27 shows that we can obtain an exact expression for the mean first-passage time, here for a 10×10 network with reflective boundaries. We have chosen

$\mathbf{r}_s = (x_S, 0)$, $\mathbf{r}_T = (0,0)$, $\mathbf{e}_S = (1,0)$, with $x_S \in [0,9]$. The theoretical result, given in Appendix 2, of equation (164) with pseudo-Green function of equation (170) (red line) fits exactly the simulations (black circles). The exact result for reflective boundary conditions is rather formal: using the approximations developed for periodic boundary conditions would be a way to obtain a simple, even if not analytical, expression.

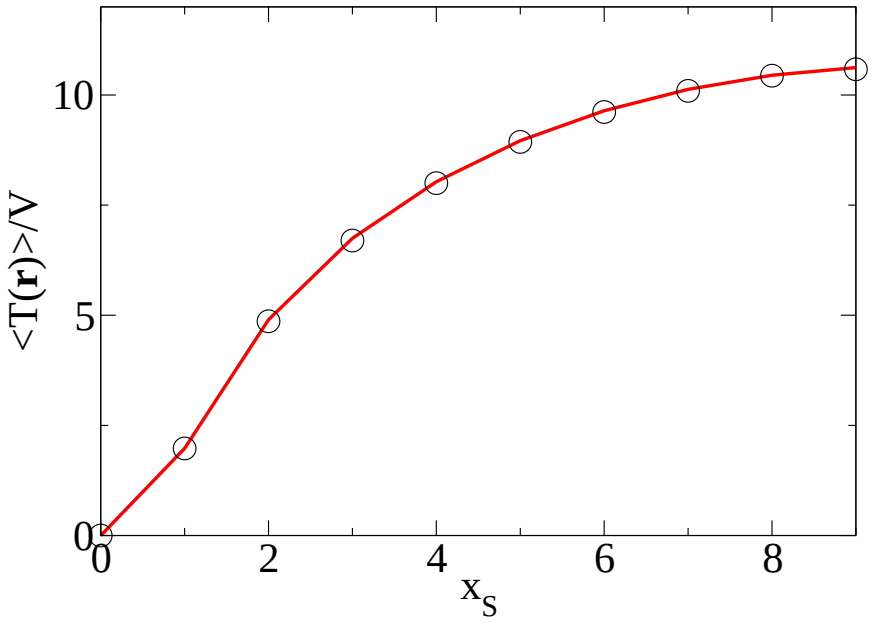

Figure 2.27: Mean first-passage time for a 10×10-large network with reflective boundaries, where $\mathbf{r}_s = (x_S, 0)$, $\mathbf{r}_T = (0,0)$, $\mathbf{e}_S = (1,0)$, with $x_S \in [0,9]$. Here $\epsilon = -0.04$ and $\delta = 0.32$. The black circles stand for numerical simulations, the red line for theoretical result of equation (164).

Mean square displacement

In this chapter, we focused mainly on the mean square displacement, which was a way to detect anomalous diffusion. The persistent random walk is not really anomalous, but can look like anomalous diffusion at "short" time (short will be quantified at the end of this paragraph). To compare this persistent model with the usual experimental observations, we can compute an approximation of the mean square displacement $\langle r^2(t) \rangle$, using the infinite space propagator [85]:

$$\widehat{P}(\mathbf{q}, z) = \frac{\sum\limits_{\mathbf{e}_i \in \mathcal{B}} \frac{1 - (\epsilon + \delta) z \cos(\mathbf{q}.\mathbf{e}_i)}{1 + (\epsilon^2 - \delta^2) z^2} - \epsilon z \cos(2\pi \mathbf{q}.\mathbf{e}_i)}{\sum\limits_{\mathbf{e}_i \in \mathcal{B}} \frac{1 + (\epsilon + \delta) z^2 - (1 + \epsilon + \delta) z \cos(2\pi \mathbf{q}.\mathbf{e}_i)}{1 + (\epsilon^2 - \delta^2) z^2} - \epsilon z \cos(2\pi \mathbf{q}.\mathbf{e}_i)} \tag{2.91}$$

The Z-transform of the mean square displacement is then given by $-\partial^2 \widehat{P}(\mathbf{q}, z)/\partial \mathbf{q}^2$ with $\mathbf{q} = \mathbf{0}$:

$$\langle r^2(z) \rangle = \frac{z}{(1-z)^2} \frac{1+(\epsilon+\delta)z}{1-(\epsilon+\delta)z} \Rightarrow \langle r^2(t) \rangle = \frac{1+\epsilon+\delta}{1-(\epsilon+\delta)} t - \frac{2(\epsilon+\delta)}{(1-\epsilon-\delta)^2} \left(1-(\epsilon+\delta)^t\right) \text{ if } t \geq 1 \tag{2.92}$$

After a while, we obtain a mean square displacement linear in time, as expected for a Brownian motion, but with an effective diffusion coefficient given by:

$$D_{\text{eff}} = \frac{1+(\epsilon+\delta)}{1-(\epsilon+\delta)} \tag{2.93}$$

A mean field approach would lead to the same result. Indeed, we would expect the effective diffusion coefficient to be in average [43, 111]:

$$D_{\text{eff}} = \frac{1+\langle \cos(\theta) \rangle}{1-\langle \cos(\theta) \rangle} = \frac{1+(\epsilon+\delta)}{1-(\epsilon+\delta)}, \tag{2.94}$$

where $\langle \cos(\theta) \rangle$ is the average cosine of the angle θ between to successive steps. In our case, $\theta = 0$ for a step backward, $\theta = \pi$ for a step forward, and $\theta = \pi/2$ for an orthogonal reorientation.

Equation (2.92) is valid even for short time: the approximation we made to obtain this result was to assume that the propagator to be used was the infinite space one. We then observe at short time either a ballistic regime ($\epsilon + \delta \geq 0$) or a static oscillating motion ($\epsilon + \delta \leq 0$). At short time, namely when $t \leq 2|\epsilon + \delta|$, the persistence effect is dominant. If the persistence is positive, we mainly see ballistic motions for the first steps. If the persistence is negative, the random walker oscillates around its initial position for a discrete random walk, and should stay almost stationary in the continuous limit of this discrete process.

Figure 2.28 shows the mean square displacement $\langle r^2 \rangle$ for persistent random walks for a highly anti-persistent random walk with $\epsilon = 0$ and $\delta = -0.9$ (black dots), and a highly persistent one with $\epsilon = 0.9$ and $\delta = 0$ (red dots). The continuous lines stand for equation (2.92), and the dashed line for the mean field approximation ($\langle r^2 \rangle = D_{\text{eff}} t$). The random walk takes place on a 400×400 periodic Euclidian network. Several comments are in order:

(i) One can see first that the approximation of equation (2.92) fits perfectly the simulations. Approximate the confined propagator by the infinite propagator is exact as long as the random walker does not hit a boundary.

(ii) After a while, we still retrieve a linear MSD, as expected with the mean field approach. But this mean field approach does not gives the constant correction of equation (2.92):

$$\langle r^2(t) \rangle \underset{t \to \infty}{\simeq} D_{\text{eff}} t - \frac{2(\epsilon+\delta)}{(1-\epsilon-\delta)^2} \tag{2.95}$$

This "residual" MSD is a first correction to the mean field approach.

(iii) The initial behavior for an anti-persistent random walk could lead to the erroneous conclusion that the random walk is sub-diffusive. One has to take care that in such model, at longer time, we retrieve a trivial Brownian behavior.

Anti-persistent random walk is thus a process that leads after a while to a Brownian diffusion, as observed in the extensive analysis of lipid granule trajectories, and can lead to a sub-diffusive like behavior at short time. Such short-term memory effects could be related to the medium

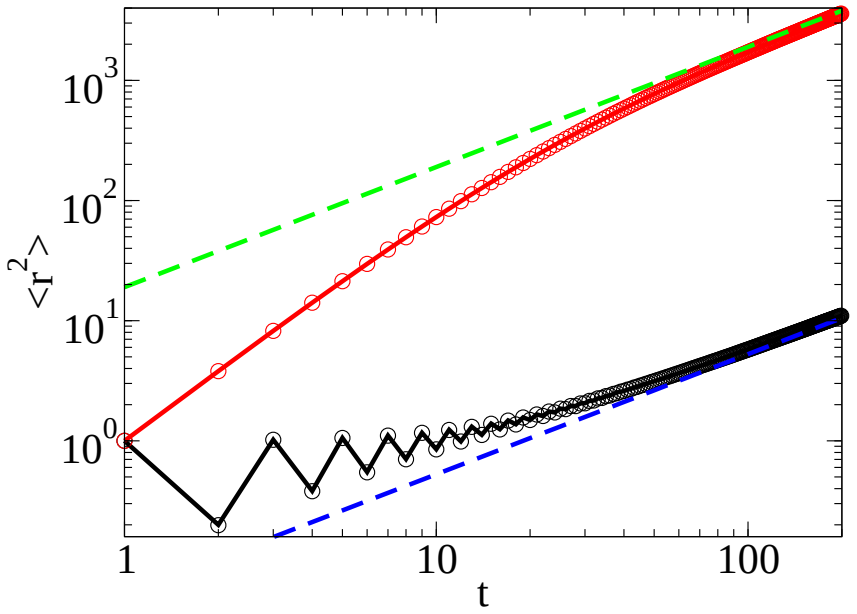

Figure 2.28: Mean square displacement for persistent random walkers on periodic 400×400 Euclidian lattice. The black dots stand for a numerical simulation of a highly anti-persistent random walk with $\epsilon = 0$ and $\delta = -0.9$, the red dots stand for a highly persistent one with $\epsilon = 0.9$ and $\delta = 0$. The continuous lines stand for equation (2.92), and the dashed line for the mean field approximation ($\langle r^2 \rangle = D_{\text{eff}} t$).

influence, like the solvent velocity field, or the obstacles position. Like for the fBm, we introduce correlation in the trajectory, but only this correlation is now finite.

We will see in the third chapter that this process can be used in another context, namely the search process optimization, when the target is hidden.

QUICK SUMMARY

- We obtained persistent pseudo-Green functions, and we can subsequently calculate analytically a wide range of first-passage observables.
- Concerning the MSD for an anti-persistent random walker, we observe an apparent initial sub-diffusive behavior, followed by a classical Brownian behavior, with a "residual" MSD.

2.3.2 Static cages: porosity

Experimental analysis showed that lipid granules had two diffusion modes: one at short time, and a different one at longer time. We cannot assume anymore that the global diffusion process is scale-invariant. We will mimic this effect using a model with two geometrical levels, with a

distinct diffusion mechanism at each level. We will assume that the random walker evolve in an ordered porous media: we consider a regular lattice of cages, each cage being connected to the neighboring ones through small holes. Those static cages could be linked to the microscopical cell organization (microtubules and F-actin network for instance): for a big random walker like lipid granules, the walk cannot explore the small bottlenecks or dead ends. The walker can just see a succession of confined volumes, linked together with small holes, the volume of each cage being approximately of the same magnitude order.

If we focus on the cage network, we have a diffusion with a quasi-CTRW behavior: the waiting time is the time needed for the random walker to enter the cage by one point, and exit through another one. This time is finite, but can be widely distributed. This introduces almost naturally a cut-off in the waiting time distribution, and could explain why the lipid granules behave like CTRW at short times, and differently at longer times.

The model considered here is thus a periodic network of cages. Diffusion in periodic porous networks has been extensively studied during the last decades. The periodic network can be an artificial system [105, 150], a crystallographic network [58, 162], or a porous materia l[145]. The diffusion in such system has been proposed as a way to differentiate cells [70], to control gas flux [145] or to sort DNA molecules of different size [60].

Several models have been developed to explain the diffusion observed in periodic porous structures. The first one is to consider the periodic network as a porous material: the diffusion coefficient is then a function of the fraction of accessible space, the porosity $\epsilon \in [0,1]$. Several formulae exist, depending on the geometry and on ϵ values: for $1-\epsilon \ll 1$, Maxwell's formula gives [138]

$$D(\epsilon) = D(0)\frac{\epsilon}{1+\frac{1}{2}(1-\epsilon)} \tag{2.96}$$

For isotropic 2D networks, the formula becomes [141]

$$D(\epsilon) = D(0)\left(1-(1-\epsilon)^{1/2}\right) \tag{2.97}$$

Those porous models only take ϵ as parameter: with only one parameter, they can be applied to random media, but no to ordered media, where the geometry can be more important than ϵ: for our model, we expect that the hole size govern the diffusion coefficient, and not the porosity ϵ.

Other models consider that if the holes between two cages are small enough, one can use a simple continuous time random walk (CTRW): the walker waits a random time in each cavity, and diffuses on the cavity network. The mean waiting time is approximately the time needed on average to escape the cavity through any of the holes, the random walker being initially uniformly distributed (noted $\langle\overline{T_{\text{esc}}}\rangle$). This approach works somehow for small holes in circular geometry [159] or for cubic networks [80]:

$$D_{CTRW} = D(0)\frac{1}{2\langle\overline{T_{\text{esc}}}\rangle} \tag{2.98}$$

The coefficient 2 in this formula appears somehow magically: authors claim in reference [159] that it is related to the probability 1/2 to cross the hole between to cages when the random walker hits this hole. This could be the case if each time the random walker hits a boundary, he has a probability 1/2 to be reflected. But for a continuous medium, where the holes are not semi-reflective, this probability is already present in the $\langle\overline{T_{\text{esc}}}\rangle$.

We will in our model take in account more precisely the microscopical details: we consider a random walker evolving on a network of identical cavities, as shown in Figure 2.29. As for the

experimental data, we will introduce two levels: for a given cavity, we have to compute the mean waiting time and the probability to choose a given neighbor, on the cavity network we will use these microscopical quantities to determine the (global) diffusion coefficient.

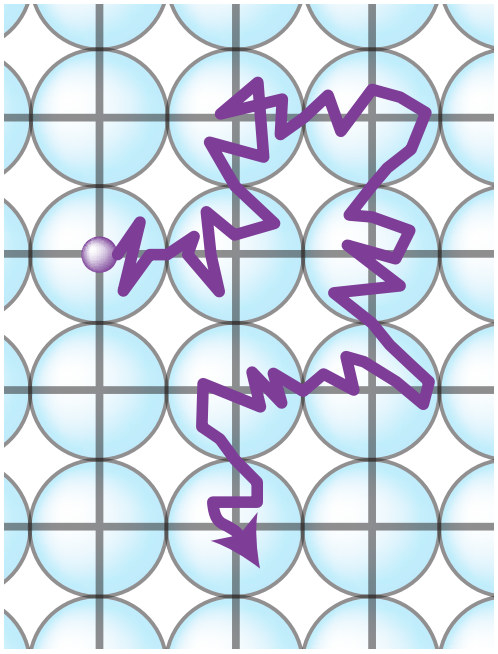

Figure 2.29: Scheme of the diffusion on a cage network.

If we consider the cage' network, we find back a CTRW, but one has to take care at transition probabilities between cages. This sub-section is structured as follows: we will first focus on the cage level, and we will find analytically the probability transition between two cages, and the related waiting times. Since those probabilities are not all equivalent, we will then focus on the cage level network, and compute the global diffusion coefficient, as a function of the cage characteristics, and of the network geometry. We will only consider the mean square displacement in this model, but once the transition probabilities and the waiting times are known, we could obtain somehow a propagator, and the related pseudo-Green functions. We will compare the result we obtain here with the existing approach, mainly the "simple" CTRW described in reference [159], for two cages geometries, discrete cubic networks and continuous circular cages.

Diffusion within a cavity

We first focus on the microscopical level, namely the diffusion within a cavity. The random walker enters through a given hole, and we need to compute two observables: the probability to choose a given hole, and the time needed to do so.

We will start by a problem simplification, using the network symmetry. As shown in Figure 2.30, starting from a given hole S, we have two kinds of holes accessible: the nearest neighbors (holes of index 3) and the opposite ones (holes of index 2). If we consider a single cavity, and look for the first hole reached, we will find the starting hole S, since the random walker starts on it. If we consider, as shown in Figure 2.30 two adjacent cavities, we can symmetrize the problem, and simply search the first hole reached, the starting hole being reflective. Once reached, we can consider that with a probability $1/2$, we reached this hole in the left cavity (in the original problem), and with a probability $1/2$, the symmetric hole in the right cavity.

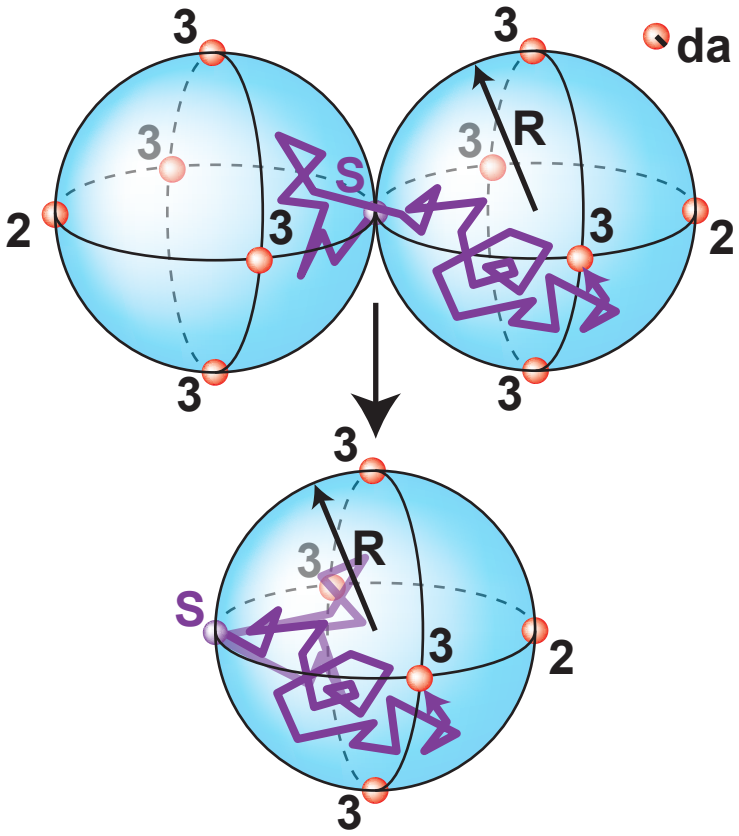

Figure 2.30: Scheme of the diffusion starting from a given hole, between two adjacent spheres.

By doing so, we avoid any problem in the definition of starting conditions: we do not have to introduce a cut-off distance where the random walker land after leaving the hole S.

After this simplification, we have three quantities to compute: the splitting probabilities p_2, p_3, and the mean escape time $\langle T_{\text{esc}} \rangle$ through any of the holes, starting from hole S. We will treat both continuous and discrete cavities with our formalism: for continuous cavities, the holes will be extended and not point-like, and the corresponding pseudo-Green functions will sometimes be only approximations. We will treat the problem with a discrete approach (point-like holes), and we will transpose the results obtained to continuous cavities using pseudo-Green function of extended targets in the final formula.

The pseudo-Green functions H_{ij} are defined as usual as

$$H_{ij} = \int_0^\infty (\text{Prob}(\mathbf{r}_j, t|\mathbf{r}_i, t=0) - P_{\text{stat}}(\mathbf{r}_j))\, dt, \tag{2.99}$$

where $\text{Prob}(\mathbf{r}_j, t|\mathbf{r}_i, t=0)$ is the probability to be at $\mathbf{r}_j$ at time t starting from $\mathbf{r}_i$ at time $t=0$, and P_{stat} the stationary probability.

We note S the initial hole, $\mathbf{r}_S$ its position, 2 the opposite hole and 3 any of the orthogonal hole. We note p_2 the probability for the random walker to exit through the hole 2 and p_3 the probability to hit one of the orthogonal holes first. The $(2d-2)$ orthogonal holes 3 are equivalent. Using the electrical analogy formalism of [65], that we applied in chapter 1 to compute the splitting

probabilities, we have:

$$\begin{cases} 0 = & \rho_0 + JH_{2S} - Jp_2H_{22} - J\dfrac{p_3}{2d-2}(2d-2)H_{23} \\ 0 = & \rho_0 + JH_{3S} - Jp_2H_{23} - J\dfrac{p_3}{2d-2}\left(H_{33} + H_{2S} + (2d-4)H_{23}\right) \end{cases} \tag{2.100}$$

We have to take care that if all orthogonal holes are equivalent when we start from S or 2, it is not the case anymore when we start from a given hole 3. We will illustrate this point with Figure 2.30: we note 3 the upper orthogonal hole, 3′ the lower orthogonal hole and 3" the two holes in the middle. The second equation use here the overall cavity symmetry: starting from hole 3, one can either choose the same hole 3, giving the H_{33} term, the opposite hole 3′, leading to $H_{33'} = H_{2S}$, or one of the middle hole 3", leading $2d-4$ times to $H_{33"} = H_{23}$.

Using the normalization of splitting probabilities, we deduce

$$\begin{pmatrix} p_2 \\ p_3 \end{pmatrix} = A^{-1}.\begin{pmatrix} H_{2S} - H_{23} \\ 1 \end{pmatrix} \tag{2.101}$$

with

$$A = \begin{pmatrix} H_{22} - H_{23} & \dfrac{2H_{23} - H_{22} - H_{2S}}{2(d-2)} \\ 1 & 1 \end{pmatrix} \tag{2.102}$$

We already used some symmetries, for instance $H_{3S} = H_{23}$ or $H_{33} = H_{22}$.

We obtain finally:

$$p_2 = 1 - \frac{(2d-2)(H_{22} - H_{2S})}{(2d-2)(H_{22} - H_{23}) - 2H_{23} + H_{22} + H_{2S}} \tag{2.103}$$

$$p_3 = \frac{(2d-2)(H_{22} - H_{2S})}{(2d-2)(H_{22} - H_{23}) - 2H_{23} + H_{22} + H_{2S}} \tag{2.104}$$

Knowing the splitting probabilities p_2 and p_3, we can deduce the mean waiting time to exit a given cavity through any of the holes, using once again the formalism of [65]:

$$\langle T_{\text{esc}} \rangle = V\left(p_2H_{22} + p_3H_{23} - H_{2S}\right), \tag{2.105}$$

V being the cavity volume. This lead to:

$$\langle T_{\text{esc}} \rangle = V(H_{22} - H_{2S})\frac{H_{22} + H_{2S} - 2H_{23}}{(2d-2)(H_{22} - H_{23}) - 2H_{23} + H_{22} + H_{2S}} \tag{2.106}$$

Diffusion on the cavity network

For simplicity sake, we will consider that the cavity network is a simple Euclidian network of d dimension. To avoid persistence effect, we will more precisely consider the lattice connecting the holes together, as shown in Figure 2.31.

Figure 2.31 shows that the holes have two kinds of links: between two neighboring holes, and between two opposite holes. Among each kind, all the links are equivalent. We thus have two probabilities, p_3 for the neighboring holes, and p_2 for opposite ones, associated to two distinct time distributions ψ_2 and ψ_3. Those time distributions are the distribution of the conditional time needed to reach for the first time a hole of a given kind, without touching any of the other

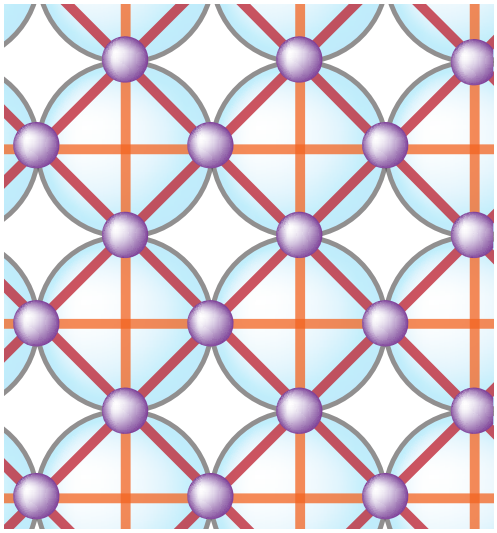

Figure 2.31: Scheme of hole network.

holes. The average time to reach any of the accessible holes from a given hole is, as defined previously:

$$\langle T_{\text{esc}} \rangle = p_2 \langle \psi_2 \rangle + p_3 \langle \psi_3 \rangle \tag{2.107}$$

Using a mean field approach, if we can escape either through a hole 2 with a probability p_2 and at a distance 1, or through a hole 3, with a probability $p_3 = 1 - p_2$ and at a distance $1/\sqrt{2}$, and that we wait in average $\langle T_{\text{esc}} \rangle$, we obtain, after a while:

$$\langle r^2 \rangle(t) \underset{t \gg 1}{\simeq} \left(p_2 (1)^2 + p_3 \left(\frac{1}{\sqrt{2}} \right)^2 \right) \frac{1}{\langle T_{\text{esc}} \rangle} Dt = \frac{1 + p_2}{2 \langle T_{\text{esc}} \rangle} Dt, \tag{2.108}$$

where D is the diffusion coefficient of the random walker within the cavity. If the lattice step is not $l = 1$, we just have to multiply this result by l^2.

Some comments are in order:

(i) We get for long time a classical Brownian behavior ($\langle r^2 \rangle \propto t$), as observed for the lipid granules in the previous section.

(ii) The diffusion coefficient is a product of three terms: D, usually $1/(2d)$, is the diffusion coefficient within a cavity, $(1+p_2)/2$ is a correction related to the fact that the two kinds of hole are not equivalent ($p_2 \neq p_3$), $1/\langle T_{\text{esc}} \rangle$ is the correction expected for a CTRW starting from a hole. Those three terms are uncorrelated, which is really a surprise: one could have expected a correlation between the probability and the escape time, where the whole time distribution ψ_2 and ψ_3 would have been involved.

(iii) We don't have to compute the mean conditional first-passage times, $\langle \psi_2 \rangle$ and $\langle \psi_3 \rangle$, but only the mean escape time $\langle T_{\text{esc}} \rangle$. This leads to a huge simplification of the problem. This time is not the one of the simple CTRW approach described in reference [159], since we assume that in this approach we start evenly distributed in the cavity: $\langle T_{\text{esc}} \rangle$ is the average time needed to exit the cavity, starting from a given hole (considered as reflective after), while $\langle \overline{T_{\text{esc}}} \rangle$ is the average time needed to exit the cavity, after averaging on all possible starting positions within the cavity.

Using expressions of p_2 and $\langle T_{\text{esc}} \rangle$ obtained previously, we can compute the effective diffusion coefficient, defined as:

$$D_{\text{eff}} = \frac{1}{2d} \frac{1+p_2}{2\langle T_{\text{esc}} \rangle} = \frac{1}{2V(H_{22} - H_{2S})} \tag{2.109}$$

The pseudo-Green functions are here the ones of a single cavity, where the hole 2 is absorbing, and the starting hole S is reflective. V is the cavity volume. This result is astonishingly simple: the effective diffusion coefficient is the half of the inverse of the mean first-passage time to hole 2, starting from hole S (with reflective boundary conditions in S), without considering any other hole. This MFPT is related to the "narrow escape" problem, problem that have been extensively studied during the last few years [29, 108, 188, 196, 197]. We thus have analytical results for many geometries.

Alternative approach

We do not always have the exact pseudo-Green functions H for the problem with one absorbing hole and a reflective one (the starting hole). If one only has the H functions with absorbing holes, we can retrieve the previous results for D_{eff} using a small approximation.

We will consider that when the random walker starts, he does not start exactly on the starting hole, but at a very near position $\mathbf{r}_S$. We thus note 1 the starting hole, $\mathbf{r}_1$ its position, with $\|\mathbf{r}_S - \mathbf{r}_1\| \ll 1$. This problem becomes a persistent random walk: it is no more equivalent to reach a given hole from the right side or from the left side, since one is teleported at a given distance $\|\mathbf{r}_S - \mathbf{r}_1\|$ to the hole, but not in the same direction. We can use the formalism developed for a persistent random walk, with $\widetilde{p_2}$ the probability to cross straight a cavity, $\widetilde{p_1}$ to exit back through the initial hole, $\widetilde{p_3}$ to reorientate orthogonally and $\widetilde{\langle T_{\text{esc}} \rangle}$ the average escape time of a cavity starting from $\mathbf{r}_S$. The diffusion coefficient on the cavity network is, as shown in the last sub-section:

$$D_{\text{eff}} = \frac{1}{2d} \frac{1+(\widetilde{p_1} - \widetilde{p_2})}{1-(\widetilde{p_1} - \widetilde{p_2})} \frac{1}{\widetilde{\langle T_{\text{esc}} \rangle}} \tag{2.110}$$

When $\|\mathbf{r}_S - \mathbf{r}_1\| \to 0$, $\widetilde{p_1} \to 1$, $\widetilde{p_2}$ and $\widetilde{p_3} \to 0$ and $\widetilde{\langle T_{\text{esc}} \rangle} \to 0$. The effective diffusion coefficient D_{eff} is then undetermined.

We can tackle this problem by considering a slightly different situation: we will consider as previously two adjacent cavities, the random walker being initially in the hole linking those two cavities. We have once again a situation where the random walker can exit this ensemble of two cavities, either by choosing one of the two extreme holes (with a probability p_2) or through one of the $4(d-1)$ orthogonal holes (with a probability p_3). Since both cavities are the same, this problem is equivalent to a problem where the random walker start in $\mathbf{r}_1$, and where $\mathbf{r}_1$ is not a hole anymore, but a part of the reflecting boundary. For this problem, we can compute as previously p_2 and p_3.

Considering the two cavities together, we can link $\widetilde{p_1}$, $\widetilde{p_2}$ and p_2. Indeed, the probability to choose one of the opposite holes, is the probability to choose it directly, or to come back in $\mathbf{r}_1$ once or several times, and then choose the opposite hole rather than one of the orthogonal holes:

$$p_2 = \sum_{i=0}^{\infty} \widetilde{p_2} \widetilde{p_1}^{\,i} = \frac{\widetilde{p_2}}{1-\widetilde{p_1}} \tag{2.111}$$

The mean escape time of the two cavities ensemble, noted $\langle T_{\text{esc}} \rangle$, is the time necessary to exit the cavity through a hole 2 or a hole 3, or to go back in $\mathbf{r}_1$ and to exit through a hole 2 or a hole

3 and so on. For our problem, the time needed to come back in $\mathbf{r}_1$ is $\langle T_{1S}\rangle \to \langle T_{11}\rangle = 0$. We thus have:

$$\langle T_{\text{esc}}\rangle = \sum_{i=0}^{\infty} \left(\widetilde{\langle T_{\text{esc}}\rangle} + i\langle T_{1S}\rangle\right) \widetilde{p_1}^{\,i} \to \frac{\widetilde{\langle T_{\text{esc}}\rangle}}{1-\widetilde{p_1}} \tag{2.112}$$

We can retrieve the effective diffusion coefficient D_{eff}, recalling the $\widetilde{p_1} \to 1$ and $\widetilde{p_2} \to 0$:

$$D_{\text{eff}} = \frac{1}{2d}\frac{1-(\widetilde{p_1}-\widetilde{p_2})}{1+(\widetilde{p_1}-\widetilde{p_2})}\frac{1}{\widetilde{\langle T_{\text{esc}}\rangle}} = \frac{1+p_2}{2}\frac{1}{\langle T_{\text{esc}}\rangle} \tag{2.113}$$

If we know the behavior of $\widetilde{p_1}$, $\widetilde{p_2}$ and $\widetilde{\langle T_{\text{esc}}\rangle}$ with $\epsilon = \|\mathbf{r}_S - \mathbf{r}_1\|$, we can take the limit $\epsilon \to 0$ to get back the initial result. This approach is useful if the ϵ dependence is known, and if the problem with a reflective starting hole cannot be treated exactly. An example is a cavity where the random walker can stick to the boundary and diffuse on the external surface: the introduction of a reflective hole is rather cumbersome, while the ϵ dependency can be estimated.

Comparison with classical CTRW approach

We can compare this result to the simple CTRW approach of reference [159]. With this model we expect, for an Euclidian cage network in d dimensions, that

$$\langle r^2\rangle \underset{t\gg 1}{\simeq} \frac{1}{2d}\frac{1}{2\langle T_{\text{esc}}\rangle}t \tag{2.114}$$

where $\langle T_{\text{esc}}\rangle$ is the mean escape time of a given cavity averaged over all starting points possible within this cavity.

We will use a simplification based on the assumption that, for a large volume

$$H_{ij} \underset{V\to\infty}{\to} G_0(\|\mathbf{r}_i - \mathbf{r}_j\|) = G_0(r_{ij}) \tag{2.115}$$

where G_0 is the infinite volume Green function. We also assume that $G_0(r) \to \infty$ when $r \to 0$, and in particular $G_0(1) \ll G_0(0)$. $G_0(0)$ is the average number of times that a random walker return in the starting site for an infinite random walk in an infinite environment: it is finite for a transient random walk, and infinite for a recurrent one. The above approximation thus holds for recurrent random walk (compact exploration). The effective diffusion coefficient for our model becomes, with this approximation:

$$D_{\text{eff}} = \frac{1}{2V(H_{22}-H_{2S})} \simeq \frac{1}{2V(G_0(0)-G_0(1))} \simeq \frac{1}{2VG_0(0)} \tag{2.116}$$

We can approximate $\langle T_{\text{esc}}\rangle$ by considering, using symmetries, a cavity fraction $V/(2d)$ with only one target, and reflective boundary conditions. For such situation, if we consider the large volume limit, where $H \simeq G_0$

$$\langle T_{\text{esc}}\rangle = \frac{V}{2d}H_{22} \underset{V\to\infty}{\to} \frac{V}{2d}G_0(0) \tag{2.117}$$

We thus obtain the same limit for both models:

$$\langle r^2\rangle(t) \underset{V\to\infty}{\to} \frac{1}{2VG_0(0)}t \tag{2.118}$$

We note that as soon as $G_0(0) - G_0(1)$ is significantly different from $G_0(0)$, the two models do not lead to the same result. We will see that this difference can be observed for discrete cavities (where $G_0(0)$ is finite), or for marginal explorations of the cavity (where $G_0(0)$ is infinite, but where $G(1) = \mathcal{O}(G_0(0))$).

Discrete euclidian cavities

We consider that the cavities are hypercubic discrete lattices of dimension d, with reflective boundary conditions, of size L^d, where a hole is placed on the center of each face, as shown in Figure 2.32 for $d = 2$. The pseudo-Green function are for this case exactly known for any dimension d.

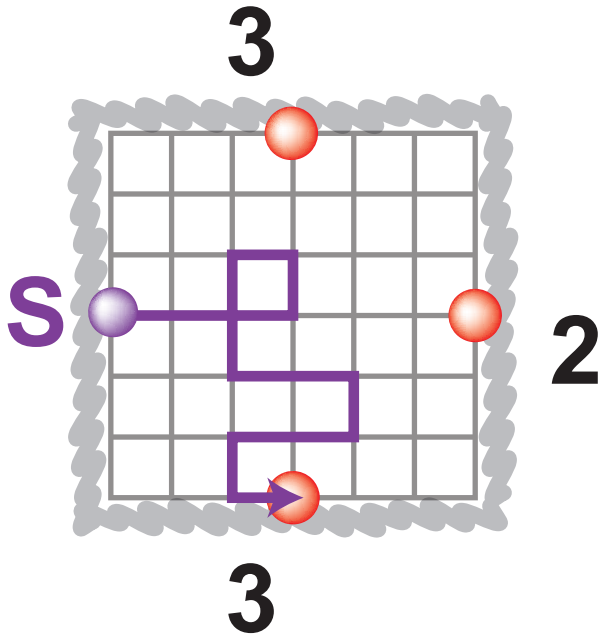

Figure 2.32: Scheme of the diffusion within an hypercubic discrete lattice (here for $d = 2$).

We thus can compute analytically p_2 and $\langle T_{\text{esc}} \rangle$, and compare our result to numerical simulations. Figure 2.35 shows the mean square displacement for a random walker evolving on a 3D cubic network of cage, the cages being cubic discrete lattices. Circles stand for numerical simulation results, for several cage size: black for 11^3-cages, red for 21^3-cages and green for 31^3-cages. The continuous line are obtained by computing analytically the diffusion coefficient, using equation (2.109) and the exact expression of pseudo-Green function, analogous to equation (1.90) given in the first chapter. The dashed line are the result of the simple CTRW model proposed in reference [159], where the mean exit time has been computed numerically.

Figure 2.35 shows an excellent agreement between simulations and our theory, and also a clear discrepancy between simulations and the simple CTRW approach. This is due to the discrete nature of the random walk: for a 3D euclidian lattice, $G_0(0) = 1.5163\ldots$ does not diverge, and $G_0(1) = 3/(2\pi)$. The classical CTRW approach gives:

$$\langle r^2 \rangle(t) \underset{V\to\infty}{\to} \frac{1}{2VG_0(0)} t \simeq \frac{0.3297\ldots}{V} t \tag{2.119}$$

Our approach gives:

$$\langle r^2 \rangle(t) \underset{V\to\infty}{\to} \frac{1}{2V(G_0(0) - G_0(1))} t \simeq \frac{0.4813\ldots}{V} t \tag{2.120}$$

As soon as $G_0(1) = \mathcal{O}(G_0(0))$, the simple CTRW approach does not work well: Discrete networks are an example of such situation.

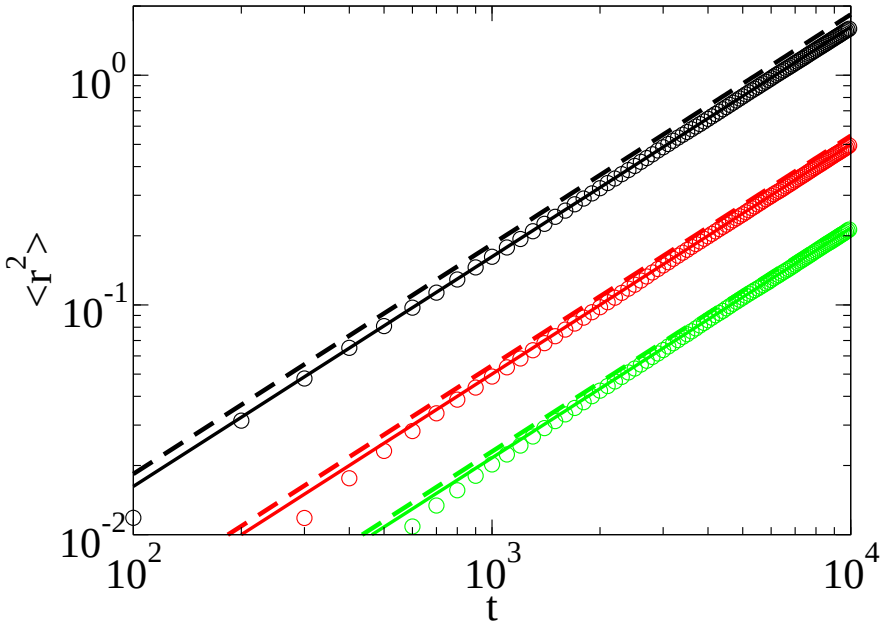

Figure 2.33: $\langle r^2 \rangle$ for a diffusion on a 3D-network, where cages are discrete cubic lattices. Black circles stand for numerical simulation where the cages are 11^3 cubes, red circles for 21^3 cubes, and green circles for 31^3 cubes. The continuous line stand for equation (2.109) using exact pseudo-Green functions, the dashed line for the simple CTRW model, where the mean exit time has been computed numerically.

Continuous compact exploration also lead to $G_0(1) = \mathcal{O}(G_0(0))$, and thus to the same discrepancy between numerical results and simple CTRW approach is expected. We will study the case of critical exploration, neither compact nor non-compact, with the $2D$ continuous exploration.

Continuous spherical cavities

We will now assume that the cages are spherical, with $2d$ holes, two on each direction. The sphere of radius R is centered on $\mathbf{r} = \mathbf{0}$, the random walker start at $-R\mathbf{e}_2$ and can exit if he hit any of the sphere portion of radius $da \ll R$ centered on $\pm R\mathbf{e}_i$.

We now have a continuous space exploration, in continuous time. For this problem, the pseudo-Green function H are not known exactly, but very good approximations exist [59]. For instance, $d = 2$ (2D circle of radius $R = 1$) gives

$$H_{22} - H_{2S} = \frac{2\ln(2) - \ln(da)}{2\pi} \tag{2.121}$$

Using this expression, we obtain a very good approximation of the simulated diffusion coefficient. Figure 2.34 shows the mean square displacement as a function of time for a 2D Euclidian lattice of $R = 1/2$ disks. The circles stand for the numerical simulation, from up to down, with a window of size $da = 0.2$ (green), 0.01 (red) and 0.05 (black). The straight lines stand for the prediction of equation (2.109), where the pseudo-Green function are approximated by the

formula of equation (2.121). We observe a very good fit between predictions and numerical simulations, even for "large" holes. The dashed lines stand for the simple CTRW approach [31], the theoretical prediction being detailed below.

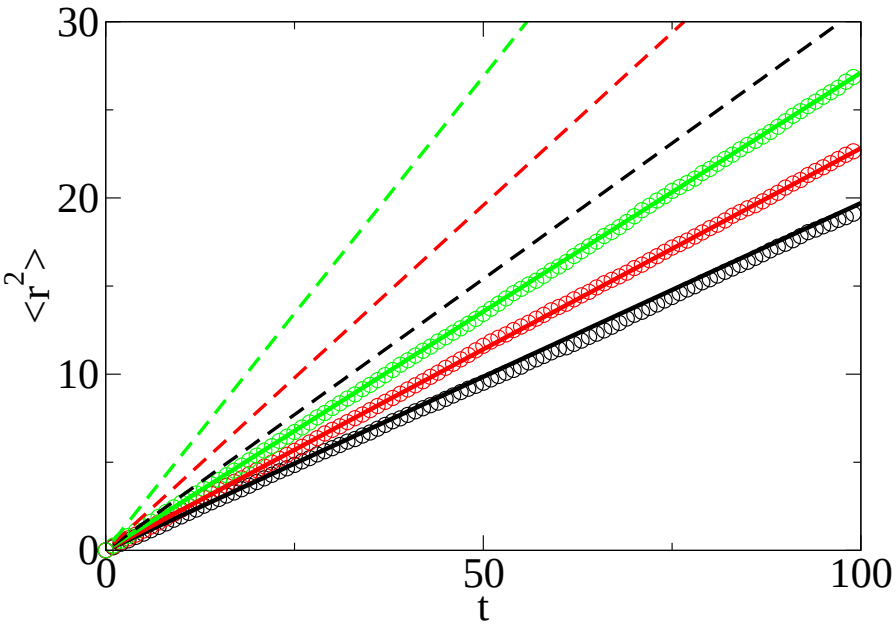

Figure 2.34: $\langle r^2 \rangle$ for a diffusion on a 2D-network, where cages are continuous disks of radius $R = 1/2$. The circles stand for the numerical simulation, from up to down, with a hole of size $da = 0.2$ (green), 0.01 (red) and 0.05 (black). The straight lines stand for the prediction of equation (2.109), where the pseudo-Green function are approximated by the formula of equation (2.121). The dashed lines stand for the simple CTRW approach [31], where the mean exit time is computed through equation (2.122).

$\langle \overline{T_{\rm esc}} \rangle$ can here be computed analytically using the formula of reference [158]: if we have a disk of radius $R = 1$, and $N = 4$ holes of size da, then

$$\langle \overline{T_{\rm esc}} \rangle = \frac{1}{2} \left(\frac{1}{2} - \ln(2da) \right) \tag{2.122}$$

The diffusion coefficient would then be with this approach

$$D_{CTRW} = \frac{D(0)}{2\langle \overline{T_{\rm esc}} \rangle} = \frac{D(0)}{\frac{1}{4} - \ln(2da)} \tag{2.123}$$

instead of (here $V = \pi R^2 = \pi$)

$$D_{\rm eff} = \frac{D(0)}{2V(H_{22} - H_{2S}} = \frac{D(0)}{2\ln(2) - \ln(da)} \tag{2.124}$$

Once again, we retrieve the influence of H_{2S}, missing in the simple CTRW approach. For infinitely small holes, $da \to 0$, both models lead to $D \to -D(0)/\ln(da)$. But for small holes $da \sim R/100$, as shown in Figure 2.34, the simple CTRW approach gives only a rough estimate of the exact value, while our approach lead to a very good fit.

We have developed in this sub-section a model to compute the global diffusion constant for a random walker evolving on an ordered porous network. The usual approach either considers very small holes (simple CTRW approach) or a random media with a given porosity (porous approach). Here, we show how to explicitly take in account the cavity geometry, and we obtain a clear improvement of the estimated diffusion coefficient.

This approach can be widely used for regular artificial systems, crystallographic networks or ordered porous material. If we manage to integrate a slight variability in the cavity size or in the hole diameter, this model could be transposed to biological problems, like diffusion in entangled actin networks, or in the extracellular matrix.

If we go back to the original problem, namely the lipid granules diffusion in yeast cell, we can here integrate easily the two levels observed experimentally: we can consider that the diffusion within a cavity is anomalous, while we will find back a Brownian behavior after a while if we consider the cavity network. Once again, we could at short times see only the sub-diffusion within a cavity, without seeing that at longer times, we find back a Brownian behavior. Introducing an anti-persistence effect related to the solvent velocity field within a cavity would make sense for a lipid granule in an actin network.

Quick summary

- We developed an analytical model to estimate the effective diffusion coefficient in regular porous networks.
- This result has been checked through numerical simulation, and is in contradiction with previous works, like [58, 159].
- This two-levels model makes sense for a big random walker evolving in complex medium, where only independent connected cavities are seen.

2.3.3 Dynamical cages: crowding effect

In this last sub-section, we will focus on crowding effect and its impact on first-passage properties. We observed in lipid granules trajectories that, at short time, the diffusive behavior shared a lot of common features with CTRW. CTRW could be related to a crowding effect [181, 182] within a cell: lipid granules cannot move before the crowding particles leave some space. Those crowding particles form a cage that disappears after some time, releasing the lipid granules.

We could use a classical CTRW formalism, with a cut-off time that can be fitted with experimental data. But we chose here to model as simply as possible this crowding problem, in order to extract the general behavior, and to understand the crowding influence on first-passage processes.

The observable we will focus on is the MFPT. We will not study here the MSD, but our approach could be used to lead to a mean field approximation of the MSD. To obtain the MFPT in our model, we will first consider the problem with only one vacancy (over-crowded case). We

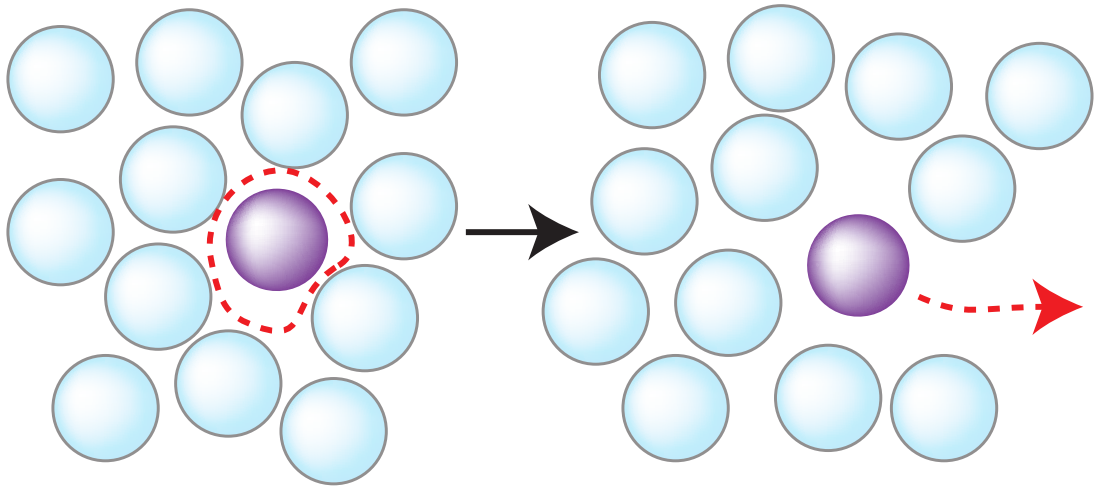

Figure 2.35: Crowding and dynamical cages: surrounding particles can form, for a given time, a "cage" around the diffusing molecule.

will compute the conditional probabilities for a tracer step, knowing the initial vacancy position, as well as the mean waiting time between two tracer steps. We will then be able to get analytically the MFPT for the problem with a single vacancy. We will treat similarly the case of an arbitrary number of vacancy: we will approximate the conditional probabilities and the mean waiting time in order to get an approximation of the MFPT. This approximation will be compared with numerical simulation results.

Model

We will consider a crowding problem on a discrete network. This allow somehow a simplification of the problem, since the crowding effect can be simply defined as follows: two particles cannot occupy the same node.

On the considered network, we will mix one or several obstacles, each occupying one node, and a tracer particle, also occupying one node. At each time step, all obstacle will successively choose a neighboring site, and move toward it if the site is free. Since the obstacles are all the same, the movement order does not really matter. After the obstacle diffusion, the tracer chooses a neighboring site, and move toward it if there is no obstacle already occupying this node. Else, the tracer stays in the same position.

We can see the problem symmetrically by considering the vacancies: instead of considering n_o obstacles, we can assume that the network contains $n = V - (n_o + 1)$ vacancies, where V is the network volume, *i.e.* the number of site, and the $+1$ term comes from the tracer. Each vacancy is an independent random walker, and two vacancies cannot occupy the same site. In this vision, the tracer moves as soon as a vacancy hit him: the tracer position and the vacancy position are simply exchanged. As previously, the vacancies move first, and if a vacancy hit the tracer, the tracer moves. At last, we impose that the tracer can only move once at each time step.

To allow an analytical treatment, the network considered will be a periodic Euclidian lattice in d dimensions, the lattice being X_i large in the i direction.

Problem for a single vacancy

We will start with a single vacancy, $n = 1$. This problem can be solved exactly for certain geometries. Our goal will be to obtain the mean first-passage time $\langle \mathbf{T}_t \rangle(\mathbf{r}_T, \mathbf{r}_S)$ for the tracer between an initial position $\mathbf{r}_S$, and a target site $\mathbf{r}_T$.

At each step, the tracer has to wait that the vacancy comes back, and thus performs a continuous time random walk. The waiting time is the time needed for the vacancy to hit again the tracer. Since the tracer motion is related to the path chosen by the vacancy to hit the tracer, the latter performs an anti-persistent random walk: it is more likely that the vacancy hits the tracer from the opposite direction of the last step than in the same direction.

To obtain the MFPT, we will use the MFPT of a persistent random walk: the tracer performs indeed an anti-persistent CTRW. We will first compute the persistence probabilities of the random walk for the single vacancy problem (a), and then the mean waiting time between two steps (b). Assembling those two results, we will deduce the mean first-passage time of a tracer in a discrete network with a single vacancy (c).

a) Conditional step probabilities

As shown previously, to obtain the MFPT for a persistent random walker, one has first to compute the conditional step probabilities, namely the probability to perform a step in a given direction, knowing what was the last direction chosen.

We will consider an Euclidian lattice of dimension d, where $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ is an orthonormal base. If we note $\mathbf{r}(t)$ the position of the tracer at time t, $\mathbf{r}_V(t)$ the position of the vacancy, making a step in the direction $\mathbf{e}_i$ at time t means that $\mathbf{r}(t+1) = \mathbf{r}(t) + \mathbf{e}_i$ and $\mathbf{r}_V(t+1) = \mathbf{r}(t)$. The vacancy then starts from $\mathbf{r}(t)$, and the next step occurs at $t' > t$ when $\mathbf{r}_V(t') = \mathbf{r}(t+1)$. The tracer will go in $\mathbf{r}_V(t'-1)$: to compute the probability to go in the direction $\mathbf{e}_j$, we have to compute the probability that $\mathbf{r}_V(t'-1) = \mathbf{r}(t+1) + \mathbf{e}_j$ if $\mathbf{r}_V(t') = \mathbf{r}(t+1)$ for the first time since t, knowing that $\mathbf{r}_V(t+1) = \mathbf{r}(t+1) - \mathbf{e}_i$. We have here a random walk with a short memory: the probability to make a step in direction $\mathbf{e}_j$ depends on the position $\mathbf{r}(t)$, but also on the direction $\mathbf{e}_i$ of the last step. We can define the conditional probability to perform a step in direction $\mathbf{e}_j$ after a step in direction $\mathbf{e}_i$ starting from $\mathbf{r}(t)$:

$$\mathrm{Prob}(\mathbf{e}_j|\mathbf{e}_i, \mathbf{r}(t)) = \mathrm{Prob}\Big(\mathbf{r}_V(t'-1) = \mathbf{r}(t) + \mathbf{e}_j | \mathbf{r}_V(t) = \mathbf{r}(t) - \mathbf{e}_i\Big), \tag{2.125}$$

where t' is defined as:

$$t' = \min\big(t_1 \in \mathbb{N} / t_1 > t \ \&\ \mathbf{r}_V(t_1) = \mathbf{r}(t)\big) \tag{2.126}$$

t' is here the first-passage time of the vacancy to the initial tracer position.

To compute this conditional probability, we need to know the last position of the vacancy before hitting the tracer, knowing that the vacancy start in $\mathbf{r}(t) - \mathbf{e}_i$. This problem can be formalized using an electrical analogy [65]: we consider that we inject a flux J on the network in $\mathbf{r}(t) - \mathbf{e}_i$, and we remove a flux J in $\mathbf{r}(t)$. Our problem is to determine the flux $J_{\pm j}$ on each edge $\mathbf{r}(t) \pm \mathbf{e}_j \to \mathbf{r}(t)$. Using the pseudo-Green function, we can write that the stationary density (for the flux problem) fulfills:

$$\rho(\mathbf{r}_i) = J\Big(H\left(\mathbf{r}(t)|\mathbf{r}(t)\right) - H\left(\mathbf{r}(t)|\mathbf{r}(t) - \mathbf{e}_i\right) + H\left(\mathbf{r}_i|\mathbf{r}(t) - \mathbf{e}_i\right) - H\left(\mathbf{r}_i|\mathbf{r}(t)\right)\Big), \tag{2.127}$$

where $H(\mathbf{r}_i|\mathbf{r}_j)$ is the pseudo-Green function. As previously, this function is defined using the propagator $P\left(\mathbf{r}_i|\mathbf{r}_j, t\right)$, probability to be in $\mathbf{r}_i$ at time t, starting from $\mathbf{r}_j$ at time $t = 0$, and

$P_{\text{stat}}(\mathbf{r}_i)$, the stationary probability (for the original network, without flux) to be in $\mathbf{r}_i$:

$$H\left(\mathbf{r}_i|\mathbf{r}_j\right) = \sum_{t=0}^{\infty}\Big(P\left(\mathbf{r}_i|\mathbf{r}_j,t\right) - P_{\text{stat}}(\mathbf{r}_i)\Big) \tag{2.128}$$

We can thus compute the stationary probability in $\mathbf{r}(t)\pm\mathbf{e}_j$, and deduce the flux $J_{\pm j}$:

$$\begin{aligned}\rho(\mathbf{r}(t)\pm\mathbf{e}_j) &= J\Big(H\left(\mathbf{r}(t)|\mathbf{r}(t)\right) - H\left(\mathbf{r}(t)|\mathbf{r}(t)-\mathbf{e}_i\right) \\ &\qquad + H\left(\mathbf{r}(t)\pm\mathbf{e}_j|\mathbf{r}(t)-\mathbf{e}_i\right) - H\left(\mathbf{r}(t)\pm\mathbf{e}_j|\mathbf{r}(t)\right)\Big), \qquad (2.129)\\ J_{\pm j} &= \omega_{\mathbf{r}(t)\pm\mathbf{e}_j\to\mathbf{r}(t)}\rho(\mathbf{r}(t)\pm\mathbf{e}_j), \qquad (2.130)\end{aligned}$$

where $\omega_{\mathbf{r}(t)\pm\mathbf{e}_j\to\mathbf{r}(t)}$ is the transition probability from $\mathbf{r}(t)\pm\mathbf{e}_j$ to $\mathbf{r}(t)$. We finally obtain the probability:

$$\begin{aligned}\text{Prob}\left(\mathbf{e}_j|\mathbf{e}_i,\mathbf{r}(t)\right) &= \frac{J_j}{J} \qquad (2.131)\\ &= \omega_{\mathbf{r}(t)+\mathbf{e}_j\to\mathbf{r}(t)}\Big(H\left(\mathbf{r}(t)|\mathbf{r}(t)\right) - H\left(\mathbf{r}(t)|\mathbf{r}(t)-\mathbf{e}_i\right) \\ &\qquad + H\left(\mathbf{r}(t)+\mathbf{e}_j|\mathbf{r}(t)-\mathbf{e}_i\right) - H\left(\mathbf{r}(t)+\mathbf{e}_j|\mathbf{r}(t)\right)\Big) \qquad (2.132)\end{aligned}$$

Some comments are in order:

(i) Those probabilities are entirely determined by the vacancy motion, which is so far a classical Brownian motion. In particular, the pseudo-Green functions can be known exactly for some geometries.

(ii) The tracer motion is persistent for the general case. One can easily understand that it is easier for the vacancy to come back using the same path than to circle around the tracer to arrive in the opposite direction: the tracer performs an anti-persistent motion.

(iii) The transition probabilities for the tracer depend on the position $\mathbf{r}(t)$ in the general case.

We can slightly simplify the problem if we consider a tracer moving on a periodic Euclidian lattice in d dimensions. The network is parallepipedic, being X_i long on direction $\mathbf{e}_i$. For such network, the transition probability does not depends anymore on the tracer position, and the pseudo-Green function is exactly known:

$$H(\mathbf{r}|\mathbf{r}') = \frac{1}{\prod_{i=1}^{d} X_i}\sum_{\mathbf{q}\neq\mathbf{0}}\frac{e^{2\imath\pi\mathbf{q}.(\mathbf{r}-\mathbf{r}')}}{1-\frac{1}{d}\sum_{i=1}^{d}\cos\left(2\pi\mathbf{q}.\mathbf{e}_i\right)}, \tag{2.133}$$

where $\mathbf{q} = (1/X_1,\ldots,1/X_d)$. We can note that the pseudo-Green function only depends on the difference $\mathbf{r}-\mathbf{r}'$: $H(\mathbf{r}|\mathbf{r}') = H(\mathbf{r}-\mathbf{r}') = H(\mathbf{r}'-\mathbf{r})$. The tracer transition probability becomes:

$$\begin{aligned}\text{Prob}\left(\mathbf{e}_j|\mathbf{e}_i\right) &= \frac{1}{2d}\Big(H\left(\mathbf{0}\right) - H\left(\mathbf{e}_i\right) + H\left(\mathbf{e}_j+\mathbf{e}_i\right) - H\left(\mathbf{e}_j\right)\Big) \\ &= \frac{1}{2d\prod_{i=1}^{d}X_i}\sum_{\mathbf{q}\neq\mathbf{0}}\frac{1-e^{2\imath\pi\mathbf{q}.\mathbf{e}_i}+e^{2\imath\pi\mathbf{q}.(\mathbf{e}_j-\mathbf{e}_i)}-e^{2\imath\pi\mathbf{q}.\mathbf{e}_j}}{1-\frac{1}{d}\sum_{i=1}^{d}\cos\left(2\pi\mathbf{q}.\mathbf{e}_i\right)} \qquad (2.134)\end{aligned}$$

One can note that if $X = X_1 = X_2 = \ldots = X_d$ (hypercubic domain) we have only three probability, as soon as $d \geq 2$: $\text{Prob}(\mathbf{e}_i,\mathbf{e}_i)$, $\text{Prob}(\mathbf{e}_i,-\mathbf{e}_i)$ and $\text{Prob}(\mathbf{e}_i,\mathbf{e}_{j\neq i})$. We will note them respectively $\text{Prob}^{\rightarrow\rightarrow}$, $\text{Prob}^{\rightarrow\leftarrow}$ and $\text{Prob}^{\rightarrow\uparrow}$.

b) Waiting time

The second quantity we will compute is the waiting time between two tracer steps. Knowing that the tracer performs an anti-persistent random walk, we can indeed compute the mean number of steps necessary to reach a given target, using the formalism developed for persistent random walks. But we do not know how many time the tracer have to wait between two steps.

One could argue that the first return time is not the same for each couple $(\mathbf{e}_i, \mathbf{e}_j)$. The correct mean first-passage time $\langle \mathbf{T}_t \rangle$ for the tracer should be:

$$\langle \mathbf{T}_t(\mathbf{r}_S, \mathbf{r}_T) \rangle = \sum_{i=1}^{d} \sum_{j=1}^{d} \langle N(\mathbf{r}_S, \mathbf{r}_T) \rangle (\pm\mathbf{e}_i, \pm\mathbf{e}_j) \langle \psi \rangle (\pm\mathbf{e}_i, \pm\mathbf{e}_j), \tag{2.135}$$

where $\langle N \rangle(\mathbf{e}_i, \mathbf{e}_j)$ is the average number of reorientation from $\mathbf{e}_j$ to $\mathbf{e}_i$ during the tracer walk, and $\langle \psi \rangle(\mathbf{e}_i, \mathbf{e}_j)$ the mean waiting time between a step along $\mathbf{e}_j$ and a step along $\mathbf{e}_i$. The two random variables N and ψ are uncorrelated, we can thus multiply the means to obtain the mean of the product. $\langle \psi \rangle(\mathbf{e}_i, \mathbf{e}_j)$ is here a conditional mean first-passage time for the vacancy: starting from $\mathbf{r} - \mathbf{e}_j$, this is the mean first-passage time to $\mathbf{r}$ when the last step occurs along $\mathbf{r} - \mathbf{e}_i \to \mathbf{r}$. We will compute successively $\langle N \rangle(\mathbf{e}_i, \mathbf{e}_j)$ and $\langle \psi \rangle(\mathbf{e}_i, \mathbf{e}_j)$, after some simplifications.

Assuming once again that the problem occurs in a hypercube of size X in d dimensions, we can slightly simplify the problem:

$$\langle \mathbf{T}_t(\mathbf{r}_S, \mathbf{r}_T) \rangle = \langle N^{\to\to}(\mathbf{r}_S, \mathbf{r}_T) \rangle \langle \psi^{\to\to} \rangle + \langle N^{\to\leftarrow}(\mathbf{r}_S, \mathbf{r}_T) \rangle \langle \psi^{\to\leftarrow} \rangle + \langle N^{\to\uparrow}(\mathbf{r}_S, \mathbf{r}_T) \rangle \langle \psi^{\to\uparrow} \rangle. \tag{2.136}$$

The problem can be further simplified. We note $N^{\to\uparrow}(\mathbf{r}_T | \mathbf{r}_S, \mathbf{e}_i)$ the number of reorientation $\to\uparrow$ starting from $\mathbf{r}$, the previous step being $\mathbf{e}_i$ and the target $\mathbf{r}_T$. This quantity satisfies, for $\mathbf{r} \neq \mathbf{r}_T$:

$$\begin{aligned} \langle N^{\to\uparrow} \rangle(\mathbf{r}_T | \mathbf{r}, \mathbf{e}_i) \;=\; & \mathrm{Prob}^{\to\to} \langle N^{\to\uparrow} \rangle(\mathbf{r}_T | \mathbf{r} - \mathbf{e}_i, \mathbf{e}_i) + \mathrm{Prob}^{\to\leftarrow} \langle N^{\to\uparrow} \rangle(\mathbf{r}_T | \mathbf{r} - \mathbf{e}_i, -\mathbf{e}_i) \\ & + \mathrm{Prob}^{\to\uparrow} \sum_{\pm\mathbf{e}_j, j \neq i} \langle N^{\to\uparrow} \rangle(\mathbf{r}_T | \mathbf{r} - \mathbf{e}_i, \mathbf{e}_j) + \mathrm{Prob}^{\to\uparrow}. \end{aligned} \tag{2.137}$$

We can compare it with the equation satisfied by the mean first-passage time for a "classical" persistent random walker, namely a persistent random walker waiting one time step between two jumps, starting from $\mathbf{r}$ with orientation $\mathbf{e}_i$, when $\mathbf{r}_T$ is the target, $\langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r}, \mathbf{e}_i)$, once again for $\mathbf{r} \neq \mathbf{r}_T$:

$$\begin{aligned} \langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r}, \mathbf{e}_i) \;=\; & \mathrm{Prob}^{\to\to} \langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r} - \mathbf{e}_i, \mathbf{e}_i) + \mathrm{Prob}^{\to\leftarrow} \langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r} - \mathbf{e}_i, -\mathbf{e}_i) \\ & + \mathrm{Prob}^{\to\uparrow} \sum_{\pm\mathbf{e}_j, j \neq i} \langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r} - \mathbf{e}_i, \mathbf{e}_j) + 1. \end{aligned} \tag{2.138}$$

Both quantities satisfy the same equation, except for the source term. We can then deduce that:

$$\langle N^{\to\to} \rangle(\mathbf{r}_T | \mathbf{r}_S, \mathbf{e}_S) = \mathrm{Prob}^{\to\to} \langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r}_S, \mathbf{e}_S) \tag{2.139}$$

$$\langle N^{\to\leftarrow} \rangle(\mathbf{r}_T | \mathbf{r}_S, \mathbf{e}_S) = \mathrm{Prob}^{\to\leftarrow} \langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r}_S, \mathbf{e}_S) \tag{2.140}$$

$$\langle N^{\to\uparrow} \rangle(\mathbf{r}_T | \mathbf{r}_S, \mathbf{e}_S) = \mathrm{Prob}^{\to\uparrow} \langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r}_S, \mathbf{e}_S) \tag{2.141}$$

Those equations are still valid after an average over $\mathbf{e}_S$. $\langle \mathbf{T} \rangle(\mathbf{r}_T | \mathbf{r}_S, \mathbf{e}_S)$ is the mean first-passage time for a persistent random walker waiting a time unit between two steps, and is analytically known, using equation (2.68).

The mean first-passage time for our problem becomes:

$$\langle \mathbf{T}_t(\mathbf{r}_S, \mathbf{r}_T)\rangle = \langle \mathbf{T}\rangle(\mathbf{r}_T|\mathbf{r}_S)\left(\text{Prob}^{\rightarrow\rightarrow}\langle\psi^{\rightarrow\rightarrow}\rangle + \text{Prob}^{\rightarrow\leftarrow}\langle\psi^{\rightarrow\leftarrow}\rangle + \text{Prob}^{\rightarrow\uparrow}\langle\psi^{\rightarrow\uparrow}\rangle\right) \tag{2.142}$$

The last term can be simplified:

$$\langle\psi\rangle = \text{Prob}^{\rightarrow\rightarrow}\langle\psi^{\rightarrow\rightarrow}\rangle + \text{Prob}^{\rightarrow\leftarrow}\langle\psi^{\rightarrow\leftarrow}\rangle + \text{Prob}^{\rightarrow\uparrow}\langle\psi^{\rightarrow\uparrow}\rangle, \tag{2.143}$$

where $\langle\psi\rangle$ is the mean first-passage time for the vacancy, starting from $\mathbf{r} - \mathbf{e}_i$ to $\mathbf{r}$. Using the Kac' formula, we simply obtain $\langle\psi\rangle = 1/P_{\text{stat}} = X^d$.

The important point is here that we can only consider the mean first-passage time of the vacancy, without computing the conditional mean first-passage times for each possible jump $\mathbf{e}_j$. This simplification can be done as soon as the probabilities $\text{Prob}^{\rightarrow\rightarrow}$ and the conditional mean waiting times $\langle\psi^{\rightarrow\rightarrow}\rangle$ do not depend on the position $\mathbf{r}$.

c) Mean first-passage Time

We can combine the previous result to obtain the mean first-passage time for a tracer evolving in a crowded periodic euclidian lattice with $n = 1$ vacancy:

$$\langle \mathbf{T}_t(\mathbf{r}_S, \mathbf{r}_T)\rangle = X^d\langle \mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle \tag{2.144}$$

The mean first-passage time for a tracer in a crowded environment with $n = 1$ vacancy is simply the mean first-passage time of a persistent random walker multiplied by the mean first return time of the vacancy. $\langle \mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle$ is here the average first-passage time for a persistent random walk, already obtained previously, that we briefly recall here:

$$\langle \mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle = \frac{1}{d}\sum_{\mathbf{q}\neq\mathbf{0}}\left(\sum_{\mathbf{e}_j\in\mathcal{B}}\frac{1-(\epsilon+\delta)\cos(2\pi\mathbf{q}.\mathbf{e}_i)}{1+\epsilon^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_i)-\delta^2}\frac{1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}}{1-2p_3\sum_{\mathbf{e}_j\in\mathcal{B}}\frac{\cos(2\pi\mathbf{q}.\mathbf{e}_j)-(\epsilon+\delta)}{1+\epsilon^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_j)-\delta^2}}\right), \tag{2.145}$$

where $\epsilon = \text{Prob}^{\rightarrow\rightarrow} - \text{Prob}^{\rightarrow\uparrow}$ and $\delta = \text{Prob}^{\uparrow\rightarrow} - \text{Prob}^{\rightarrow\leftarrow}$. Those conditional probabilities are obtained using equation (2.134).

Figure 2.36 shows the mean first-passage time of a tracer diffusing on a crowded 2D periodic Euclidian lattice, as a function of the initial (Euclidian) distance between the tracer and the target ($r = \sqrt{\|\mathbf{r}_S - \mathbf{r}_T\|^2}$). The black circles strand for the numerical simulation result on a 10×10 network, red circles for a 20×20 network. The continuous line corresponds to the analytical result of equation (2.144). The analytical result fits perfectly the numerical results.

We obtained, for a periodic Euclidian network, an analytical result for the mean first-passage time of a tracer in a crowded environment, with $n = 1$ vacancy. Using the result of equation (2.144), we can obtain the GMFPT, with the same formalism as for the persistent random walker.

Extension for $n \geq 1$

We will now propose an approximation for the mean first-passage time of a random walker evolving on a crowded periodic euclidean lattice with $n \geq 1$.

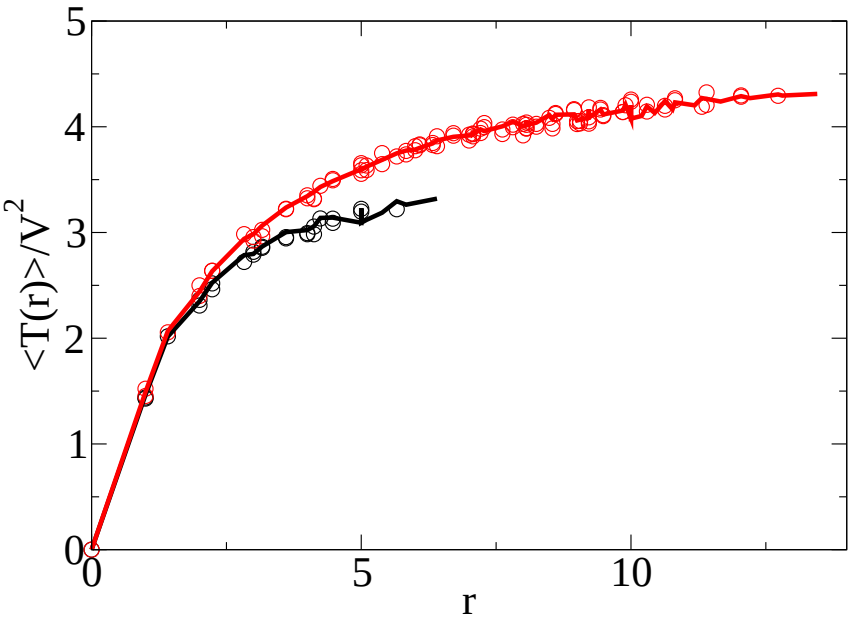

Figure 2.36: Mean first-passage time for a tracer evolving on a crowded 2D-periodic Euclidian lattice, with a single vacancy. The black circles stand for the simulation results on a 10×10 network, the red circles for a 20×20 network. The continuous line corresponds to the analytical result of equation (2.144).

The basic idea is to consider that we still have a $n = 1$ vacancy problem, that we can solve analytically, but where the conditional jump probabilities (ϵ, δ) depend on n. We will apply the result of equation (2.144), with (ϵ_n, δ_n), now depending on n (a). The mean waiting time between two tracer steps $\langle\psi\rangle$ will also depends on n (b). Combining those two results, we will be able to propose an approximation of the mean first-passage time of a tracer in a crowded periodic Euclidian lattice for an arbitrary vacancy number n (c).

a) Approximation of ϵ_n and δ_n

When the tracer evolves with n vacancies, we will consider that the tracer chooses with conditional probability (ϵ, δ) the next jump if he is hit by the *same* vacancy that the one that hits him during the last step. We will note 0 this vacancy. Else, the tracer chooses randomly a neighbor with an uniform probability.

This approximation is somehow a short memory one: only the last vacancy has a memory of the last step, the $n-1$ other are assumed to by uniformly distributed when the last step occurs. What we will be looking for is the probability that the vacancy hitting the tracer at a given step, vacancy 0, is the first vacancy to hit him at the next step. We will write Prob_{tw} this probability that the same vacancy hits the tracer twice in a row.

With our approximation:

$$\begin{aligned} \epsilon_n &\simeq \text{Prob}_{\text{tw}}(n)\epsilon \\ \delta_n &\simeq \text{Prob}_{\text{tw}}(n)\delta \end{aligned} \tag{2.146}$$

ϵ and δ are the conditional probabilities obtained for a single vacancy, using equation (2.134).

To compute $\text{Prob}_{\text{tw}}(n)$, we need to know the first-passage density to the tracer for each of the vacancy. We note ψ_0 the first-passage density of the vacancy that hits the tracer during the last step, and ψ_b the first-passage density of a vacancy uniformly distributed in the "bulk". Both kind of vacancy perform a simple nearest neighbor random walk on a periodic Euclidian network. Since the network is periodic, this first-passage density does not depend on the position $\mathbf{r}$ considered.

We can first approximate those first-passage density using the result of reference [151], assuming that $d \geq 2$:

$$\begin{aligned}\psi_0(t) &= \left(1 - \frac{X^d}{\text{GMFPT}}\right)\delta\left(\frac{t-1}{\text{GMFPT}}\right) + \frac{X^d}{\text{GMFPT}^2}\exp\left(-\frac{t}{\text{GMFPT}}\right)\\ \psi_b(t) &= \frac{1}{\text{GMFPT}}\exp\left(-\frac{t}{\text{GMFPT}}\right)\end{aligned} \tag{2.147}$$

$\delta(x)$ is here a Dirac delta function. The X^d term come from the fact that the first-passage time for the vacancy 0 is X^d, and the GMFPT is the global mean first-passage time to a given point of the network (they are all equal for a periodic network). The $t-1$ for $\psi_0(t)$ means that the vacancy 0 need at least one step to hit the tracer.

The first-passage densities in equation (2.147) work for vacancies freely diffusing. If we can neglect the interactions between the $n-1$ "bulk" vacancies since we do not care which one of them arrive first, we have to take into account interactions between vacancy 0 and the $n-1$ "bulk" vacancies, because the arrival order changes the tracer motion (one case is persistent, the other is not). We will assume that the general behavior of bulk vacancies is not influenced by vacancy 0. But the first-passage density of vacancy 0 depend on the bulk vacancy number $n-1$.

We will take in account interactions between vacancy 0 and the bulk vacancies in equation (2.147) with a mean field approach. We will treat successively the short time and the long time behavior of $\psi_0(t)$.

Since two vacancies cannot occupy the same site at the same time, vacancy 0 has to wait a given time, that we note $\langle T_0\rangle$, before moving. A first approximation of $\langle T_0\rangle$ is to consider that at each time step, the vacancy 0 can move only if the next site is not occupied by one of the $n-1$ bulk vacancies. The next site is occupied with a flat probability $(n-1)/(X^d-1)$ at each time step. $n-1$ because there is $n-1$ bulk vacancies, X^d-1 is the volume accessible minus the site occupied by the tracer. We then obtain for $\langle T_0\rangle$:

$$\langle T_0\rangle = \frac{\displaystyle\sum_{t=1}^{\infty} t\left(\frac{n-1}{X^d-1}\right)^t}{\displaystyle\sum_{t=1}^{\infty}\left(\frac{n-1}{X^d-1}\right)^t} = \frac{1}{1-\dfrac{n-1}{X^d-1}} = \frac{X^d-1}{X^d-n} \tag{2.148}$$

We can check that if $n=1$, $\langle T_0\rangle = 1$, and if $n \to X^d$, $\langle T_0\rangle \to \infty$.

In average, vacancy 0 has to perform X^d steps before hitting the tracer again. We can consider that the interaction with bulk vacancies is, in average, equivalent to wait a time $\langle T_0\rangle$ at each step. We can then approximate the long time behavior of $\psi_0(t)$ as follow

$$\psi_0(t) \underset{t\gg 1}{\simeq} \frac{X^d}{\text{GMFPT}^2}\exp\left(-\frac{\langle T_0\rangle}{\text{GMFPT}}t\right) \tag{2.149}$$

At short time, we will consider that the vacancy 0 can arrive at $t = 1$ only if the target site (here the tracer) is free. We will indeed consider that the persistence effect works only if vacancy 0 is the only one that hit the tracer. We thus multiply the Dirac delta of equation (2.147) by the probability that no bulk vacancy hit the tracer at this time. In the end, we have:

$$\psi_0(t) = \left(1 - \frac{X^d}{\text{GMFPT}}\right)\left(1 - \frac{n-1}{X^d - 1}\right)\delta\left(\frac{t-1}{\text{GMFPT}}\right) + \frac{X^d}{\text{GMFPT}^2}\exp\left(-\frac{\langle T_0\rangle}{\text{GMFPT}}t\right) \quad (2.150)$$

Knowing both first-passage densities, we can compute Prob_{tw}:

$$\begin{aligned}
\text{Prob}_{\text{tw}}(n) &= \int_0^\infty \psi_0(t)\left(\int_t^\infty \psi_b(t')dt'\right)^{n-1} dt \\
&= \left(1 - \frac{X^d}{\text{GMFPT}}\right)\left(1 - \frac{n-1}{X^d-1}\right) \qquad (2.151)\\
&\quad + \int_0^\infty \frac{X^d}{\text{GMFPT}^2}\exp\left(-\frac{\langle T_0\rangle}{\text{GMFPT}}t\right)\exp\left(-\frac{(n-1)t}{\text{GMFPT}}\right)dt \\
&= \left(1 - \frac{X^d}{\text{GMFPT}}\right)\frac{X^d - n}{X^d - 1} + \frac{X^d}{\text{GMFPT}}\frac{1}{\langle T_0\rangle + n - 1} \\
&= \left(1 - \frac{X^d}{\text{GMFPT}}\right)\frac{X^d - n}{X^d - 1} + \frac{X^d}{\text{GMFPT}}\frac{X^d - n}{n(X^d - n + 1)} \qquad (2.152)
\end{aligned}$$

Once again, $\text{Prob}_{\text{tw}}(1) = 1$, and $\text{Prob}_{\text{tw}} \to 0$ when $n \to X^d$, .

Using this approximation, we can obtain the conditional probabilities ϵ_n and δ_n.

b) Approximation of $\langle\psi\rangle$

We now have to obtain $\langle\psi\rangle$, the time between two tracer steps. It does not matter anymore wether the vacancy 0 or any bulk vacancy hit the tracer first: as shown for the single vacancy case, we just need to have the mean waiting time between two tracer steps. This time can be approximated by the first-passage time of n particles uniformly distributed. We will assume that we have n vacancies having the first-passage density $\psi_b(t)$:

$$\begin{aligned}
\langle\psi\rangle &= \int_0^\infty t\psi_b(t)\left(\int_t^\infty \psi_b(t')dt'\right)^{n-1} dt' \qquad (2.153)\\
&= \int_0^\infty \frac{t}{\text{GMFPT}}\exp\left(-\frac{t}{\text{GMFPT}}\right)\exp\left(-\frac{(n-1)t}{\text{GMFPT}}\right)dt \qquad (2.154)\\
\langle\psi\rangle &= \frac{\text{GMFPT}}{n} \qquad (2.155)
\end{aligned}$$

This time is simply the minimum of n random variables exponentially distributed.

c) First-passage time with n vacancies

At last, the first-passage time with n vacancies is the one of a persistent random walk, with conditional probabilities (ϵ_n, δ_n) and a mean waiting time $\langle\psi\rangle = \mathrm{GMFPT}/n$:

$$\begin{aligned}\langle \mathbf{T}_t(\mathbf{r}_S, \mathbf{r}_T, n)\rangle &= \frac{\mathrm{GMFPT}}{n}\left(\frac{X^d(\delta_n-\epsilon_n)}{1-\epsilon_n+\delta_n}+\frac{1+\epsilon_n^2-\delta_n^2}{(1+\epsilon_n+\delta_n)(1-\epsilon_n+\delta_n)}\right.\\ &\left.\times\sum_{\mathbf{q}\neq\mathbf{0}}\frac{1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}}{1-\frac{(\epsilon_n-1)^2-\delta_n^2}{d}\sum_{\mathbf{e}_j\in\mathcal{B}}\frac{\cos(2\pi\mathbf{q}.\mathbf{e}_j)}{1+\epsilon_n^2-2\epsilon_n\cos(2\pi\mathbf{q}.\mathbf{e}_j)-\delta_n^2}}\right)\end{aligned} \tag{2.156}$$

with

$$\begin{aligned}\epsilon_n &\simeq \left(1-\frac{X^d}{\mathrm{GMFPT}}\right)\frac{X^d-n}{X^d-1}+\frac{X^d}{\mathrm{GMFPT}}\frac{X^d-n}{n(X^d-n+1)}\epsilon\\ \delta_n &\simeq \left(1-\frac{X^d}{\mathrm{GMFPT}}\right)\frac{X^d-n}{X^d-1}+\frac{X^d}{\mathrm{GMFPT}}\frac{X^d-n}{n(X^d-n+1)}\delta\end{aligned} \tag{2.157}$$

Figure 2.37 shows the mean first-passage time for a tracer evolving on a 20×20 periodic Euclidian lattice, as a function of the initial distance between the tracer and the target. The circles stand for numerical simulations, for several crowding conditions: from down to up, 10 particles (including the tracer, in black), 20 (red), 80 (green) and 120 (blue). This corresponds to an obstacle density varying between 2.5 to 30 %. The continuous lines stand for the approximation of equation (2.156): one can clearly see that this approximation fits very well the simulation results.

Figure 2.38 shows the mean first-passage time for a tracer evolving on a 20×20 periodic Euclidian lattice, as a function of the initial distance between the tracer and the target. The circles stand for numerical simulations, for higher crowding conditions: from down to up, 240 particles (including the tracer, in black), 280 (red), 3200 (green). The obstacle density is now varying between 60 to 80 %. The continuous lines stand for the approximation of equation (2.156): even for a very crowded network, we still have an excellent approximation of simulation results. We also know that for an overcrowded environment, with a single vacancy, approximation of equation (2.156) gives back the exact result of equation (2.144).

We have developed here a quite simple approximation of hard-core crowding on discrete network. This model was developed to see what would be the effect of crowding on first-passage processes: we surprisingly retrieve a "simple" persistent random walk with a waiting time at each step. This process shares some features with CTRW at short time, in particular a scattering of the diffusion coefficient due to the variability of the vacancy first return time, and become a persistent random walk with an effective diffusion coefficient at longer time.

The final approximation is astonishingly good: we consider a tracer diffusing on a network with many hard-core obstacles interacting, and we see that the mean first-passage time can be approximated very decently through a rather simple expression. Once again, the absence of correlation between conditional probabilities and conditional mean first-passage time is surprising, and simplifies greatly the formalism.

We had introduced in the first sub-section of this chapter a model of persistent random walk in order to mimic some characteristics of the lipid granules trajectories. This crowding model, with its dynamical cages, gives a possible explanation of the anti-persistence observed experimentally.

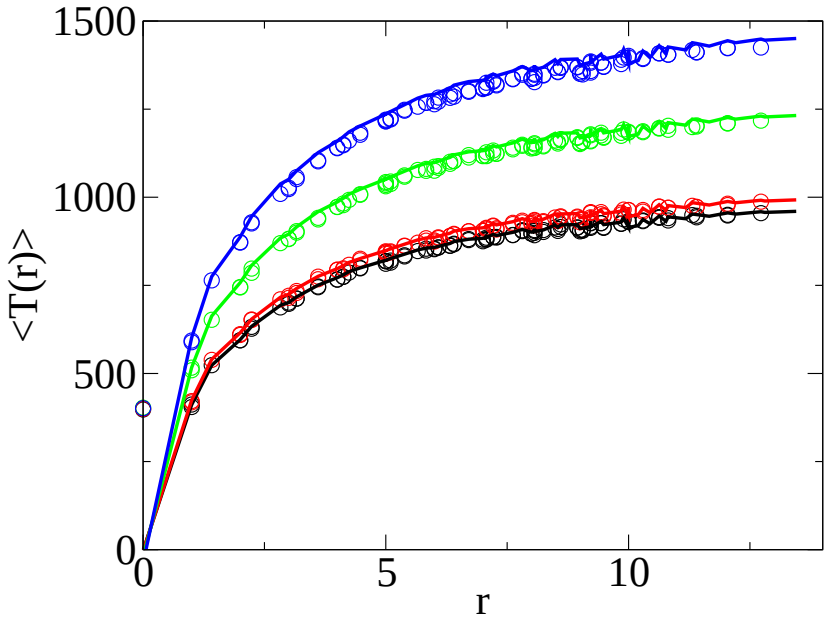

Figure 2.37: Mean first-passage time for a tracer evolving on a 20x20 periodic euclidian lattice, with a low obstacle density. The circles stand for numerical simulations, for higher crowding conditions: from down to up, 240 particles (including the tracer, in black), 280 (red), 3200 (green). The obstacle density is now varying between 60 to 80 %. The continuous lines stand for the approximation of equation (2.156).

To get the MSD, we can use the result found for persistent random walker, and divide the effective diffusion coefficient by the mean waiting time $\langle\psi\rangle$ between two steps.

Quick summary

- We have developed a simple approximation of the mean first-passage time of a tracer evolving on a crowded discrete Euclidian lattice.
- We can generate a "natural" cut-off on a CTRW model, as well as persistence, using hard-core interaction on a crowded network.

The three models developed in this section aim to give some alternatives to the three sub-diffusive models studied previously, namely diffusion on fractals, CTRW and fBm. We started with a vision of the random walk that was "infinite": fractals are infinitely scale-invariant, CTRW can have infinite mean waiting times, fBm are infinitely correlated. Since the real systems are finite, we tried to introduce alternative models that had the same behavior as the initial models at short time, but that led to a "classical" behavior at longer time. With persistent random walks, we were looking for a finite version of fBm, with only one step of correlation instead of an

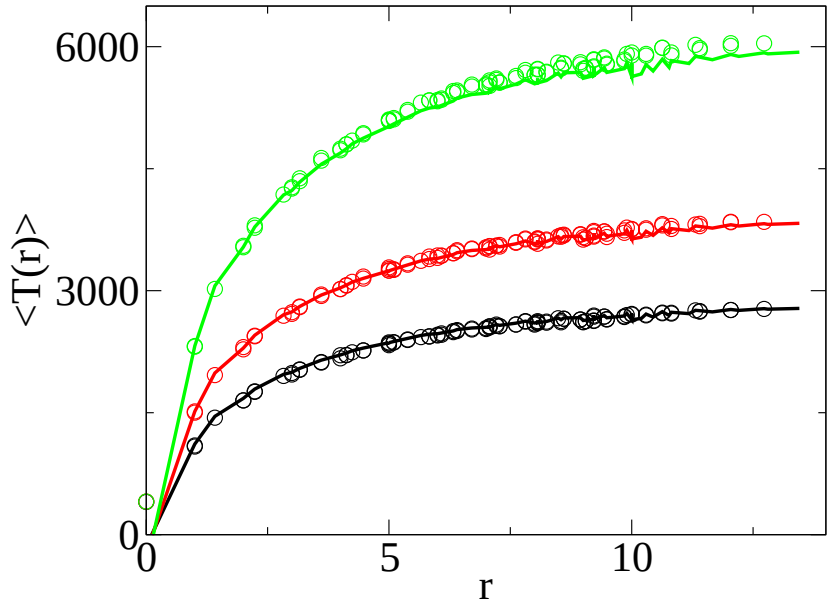

Figure 2.38: Mean first-passage time for a tracer evolving on a 20x20 periodic euclidian lattice, with a high obstacle density. The circles stand for numerical simulations, for higher crowding conditions: from down to up, 240 particles (including the tracer, in black), 280 (red), 3200 (green). The obstacle density is now varying between 60 to 80 %. The continuous lines stand for the approximation of equation (2.156).

infinity. We showed that we could obtained an apparent anomalous behavior at short time with such model. With the static cages, we proposed a process with two spatial scales, and not an infinity like for a fractal. We focused only on the long time behavior, and found that the simple CTRW approach used in the litterature was incorrect in some cases. At last, the crowding effect was a way to combine a CTRW model (with a natural cut-off) with persistence effects (fBm like). This model is quite basic, with the same size and the same diffusion coefficient for both tracer and obstacles, but could be a realistic way to understand the lipid granules motion in fission yeast.

Conclusion

This second chapter was dedicated to a way to extract information from experimental data exhibiting anomalous diffusion, namely to find the microscopical cause of anomalous diffusion. The goal was to obtain more than just a power-law fit on ensemble-average or time-average MSD plot leading to the anomalous diffusion exponent α.

The first section was a preliminary work done with a theoretical point of view: if we have a trajectory, how can we define a test that allow to distinguish between diffusion on fractals and CTRW? We showed that first-passage properties, among them the mean first-passage time, the first-passage density, the occupation time and the splitting probabilities were tools allowing to

discriminate the sub-diffusive mechanisms. We also imagined an experimental set-up that could be used to measure such observables. However, after sharing this set-up with experimentalists, it appeared that the data amount we could get was, for the time being, insufficient for our tests.

The second section took the experimental constraints in mind to adapt the initial idea: find a way, using somehow first-passage observables, to discriminate between the sub-diffusive mechanisms. We showed that maximal excursion was a way to improve the estimation of the anomalous diffusion exponent α with the same data set, and that second moment, ratio and growing shell tests were efficient tool for our problem. We applied it on simulated data sets of realistic size, then on an experimental Brownian data set, before trying it on an experimental sub-diffusive data set. We observed that the lipid granule motion in fission yeast cell was sub-diffusive, with an anti-persistence behavior like fBm at long time, and a more CTRW-like behavior at short time. For very long time, we find back a Brownian behavior. None of the three model introduced could explain entirely the behavior observed.

The third section was somehow a way back to the theory: since the motion of the lipid granules was somehow better understood, could we propose an unified model that explained the observed features, and was understandable from the biological point of view? The models studied here, persistence random walk, cage networks and hard-core crowding, were somehow finite versions of the usual anomalous diffusion mechanisms. We saw that each of those model could explain several features observed experimentally. They can all lead to an apparent sub-diffusive behavior at short time, that becomes a Brownian behavior after a while: experiments that show anomalous diffusion at short time could lead to a Brownian one after a while if those models apply. At last, we emphasize that if we used those models in the context of lipid granule motion, they can be applied in many other fields, like in biology, chemistry or condensed matter...

3 How to optimize random search processes?

Contents

During the second chapter, the main question was to determine the microscopical mechanism underlying a stochastic process. We will during this chapter assume that the diffusion mechanism is known, and we will try to optimize the random search process, namely the time needed to find a target. This random search process can model at the microscopical level a protein looking for a DNA site [25, 27] in the context of diffusion–controlled reactions. This random search process is also relevant at the macroscopical scale, as in the case of an animal looking for food, mate or shelter [17, 22, 90, 120, 130, 153, 218, 219]. In all these example, the minimization of the search time can be a competitive advantage, since it is a limiting quantity.

We can start by assuming that the searcher does not know at the beginning where the target is. We also assume that the target and the searcher are confined in a finite volume, otherwise the searcher could die ten times before ever seeing a glimpse of the target. The goal is to define a generic strategy that minimizes the search time: what will be the searcher strategy that minimizes the first-passage time to the target?

After a brief introduction to the usual search strategies, we will introduce in the first section an alternative way to minimize the search time. To do so, we will use a model developed in the previous chapter, namely persistent random walks. Under some assumptions, we will show that such process can be optimized. The quantity that will be minimized is the global mean first-passage time, namely the mean first-passage time averaged over all possible starting sites, for a given target. This quantity is to be used as soon as the random walker does not know its relative distance to the target at the beginning of the walk.

During the second section, we will focus to the mean first-passage time when the target is moving. This case could model a predator hunting down a moving prey or two chemical reactants diffusing in a solution: if both the searcher and the target move, what do first-passage properties become? This problem is very cumbersome from the mathematical point of view, and we will thus only focus on a 1D sub-domain.

The last section will think the random search problem differently: if we assume that the walker diffuses, and that we know where the target is, how can we change the environment to minimize the searching time? This question is essential to design lab-on-a-chip, or for biological cells: proteins do diffuse, and have to find their target. How can we optimize the search process, without using energy through molecular motors? We will see that the network topology have a certain influence on the MFPT, for a given size N and a given initial distance source-target r.

The first sub-section will focus on general bounds of the GMFPT. The target connectivity will appears to be an important parameter for the GMFPT, and we will continue our investigations in the second sub-section by assessing the target connectivity effect directly on the MFPT. The last sub-section will focus on network perturbations, namely adding or removing a link between two nodes somewhere in the network. Once again, we will estimate the impact of such attack on the MFPT.

Introduction

The search strategy problem defined previously turns out to be a very general question, which arises at different scales and in various fields. The first theoretical studies on this subject date back to World War II, when the U.S. Navy was hunting submarines [56, 193]. The same algorithms have been rationalized to be used in rescue context [92, 171]: once a ship or a submarine is lost, it can be advantageous to scan the shores with an optimized strategy, in order to rescue as fast as possible the survivors. The same kind of search process applies to animals looking for food, mate or shelter, as mentioned above [17, 22, 90, 120, 130, 153, 218, 219].

At the microscopic scale, search processes are directly related to the kinetic of diffusion-controlled reactions. More generally, the first step of a chemical or biochemical reaction is the encounter between reactive molecules. The theory of diffusion-controlled reactions, initiated years ago by the work of Smoluchowski [222], is widely applied for biochemical reactions in cell. In a cell, the average number of reactive molecules can be very small, and the first encounter between two reactants governs the reaction kinetic. The confined volume of the cell, or of the cell compartment, makes it a good example of search process in a confined medium. We can cite for instance the search for specific DNA sequences by transcription factors [33, 39, 100, 221, 231].

We will shortly present the two major search strategies that have been developed to minimize search processes. The first one is the Lévy search model, the second is an intermittent strategy.

Lévy walks are random walks where the jump size is distributed according an α-stable distribution. The probability to perform a jump of size l is $P(l) \propto r^{-(\alpha+1)}$, with $\alpha \in [0, 2]$. Lévy walkers thus perform long ballistic jumps, and reorientate between two jumps. Since the jumps can be very long, the probability to explore twice the same region is lower than for a classical Brownian motion. If we assume that the random walker evolves in an infinite media where several targets are randomly distributed, we can optimize the average number of targets found per time unit. If once reached the target reappears at the same location, $\alpha \simeq 1$ optimizes the search process. If once reached the target is definitely destroyed, $\alpha \to 0$ (ballistic motion) is the optimal strategy. This model has led to a controversy [81, 218, 219]: Lévy walks were found in trajectories of foraging animals, like albatrosses, bumblebees and deers, but data artifacts and inappropriate data processing have been pointed as a potential origin heavy tails observed in the jump distribution. Besides, the optimization is somehow simple: if the target reappears at the same place, once the random walker finds the first target, he can just stay in the same position and wait for the target to reappear instead of making a small turn around. If the target does not regenerate, the straight line optimizes the territory explored.

Another problem in the previous model is that it implicitly assumes that the random walker can find the target as soon as he walks over it. But if the target is hidden, the velocity degrades the perception. The search behavior of "saltatory animals", often considered as Lévy walks, could be explained differently: the animal explores intensively a small area, and relocates after a while. During the relocation, the animal goes indeed too fast to find any target [126, 152].

Intermittent models have thus been developed: the search process could be made of ballistic relocations alternating with diffusive searching phases [22]. In such models, the time spent in both phases can be optimized to minimize the search time. Another model has been developed in the same spirit to model proteins searching for a site on DNA: the protein alternates 3D "fast" diffusive relocations and 1D diffusive searchs along DNA. The 3D excursions are a way to jump from one DNA site to another in a very short time. If we look at the linear distance along DNA, such excursions are very fast jumps. Once again, the time spent in both phase (1D and 3D) can be optimized to minimize the search time [26], namely to find the target faster than just diffusing along DNA or just diffusing in 3D and hitting from time to time the DNA. The interested reader can report to [27] for more details on those models.

3.1 Persistence as an optimal search strategy

We presented in the introduction two strategies that could optimize the search time of an hidden target, under some conditions: Lévy walks and intermittent strategies. We will focus here on a specific problem: what is the better search strategy of a hidden target, in a confined Euclidian space, where the search ends when the searcher finds the target? Can we do better than Lévy walks, often presented as the optimal strategy for such problems? We saw while analyzing lipid granules trajectories in the second chapter that the granules seemed to exhibit an anti-persistent behavior. We will here analyze how the persistence of a random walker influence the search efficiency.

As presented previously, Lévy walks, which are defined as randomly reoriented ballistic excursions whose length l is drawn from a power law distribution $P(l) \propto 1/l^{1+\mu}$ when $l \to \infty$, with $0 < \mu \leq 2$, have been suggested as potential candidates of optimal strategies [219]. In fact, Lévy walks have been shown – mostly numerically – to optimize the search efficiency, but only in the particular case where the targets are distributed in space according to a Poisson law, and are in addition assumed to regenerate at the same location after a finite time. Conversely, in the case of a destructive search where each target can be found only once the optimal strategy proposed in [219] is not anymore of Lévy type, but reduces to a trivial ballistic motion. Given these restrictive conditions of optimization, the potential selection by evolution of Lévy trajectories as optimal search strategies is disputable, and in fact the field observation of Lévy trajectories for foraging animals is still elusive and controversial [18, 81, 112]. We will thus compare persistent search strategies to Lévy walks, in order to show that under some conditions, persistent random walkers perform better than any Lévy walk.

In this section, we consider a minimal model of persistent search process – called persistent random walk model hereafter – with short range memory characterized by an exponential distribution of the length of its successive ballistic excursions $P(l) \propto e^{-\alpha l/l_p}$ when $l \to \infty$, where l_p is the persistence length of the walk and α a numerical factor. We already focused on this model in the previous chapter, where we extracted the MFPT for such model. Here, we will focus on a different observable: we will see how the jump distribution influences the search efficiency, measured by the GMFPT.

Our model shares some features with Lévy walks: at each step, the random walker reorientates and performs a ballistic excursion of length l. But in contrast to Lévy walks, the jump distribution has a finite variance and a finite mean. We calculate exactly for this kind of persistent random walker the MFPT to a target in a bounded domain and find that it admits a non trivial minimum as a function of l_p, thus revealing an optimal search strategy which is very different from the simple ballistic motion obtained in the case of Poisson distributed targets. In addition, we show

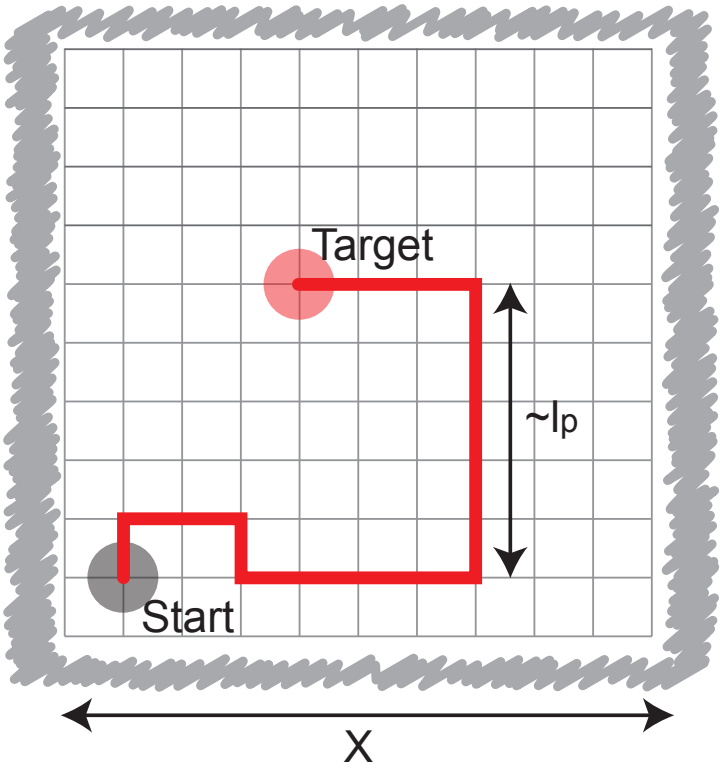

Figure 3.1: Example of search trajectory for a persistent random searcher in a bounded domain.

numerically that such optimal persistent random walk strategy is more efficient than *any* Lévy walk of parameter $\mu \in]0, 2[$. Together, our results show that the distribution of targets plays a crucial role in the random search problem. In particular, in the biologically relevant cases of either a single target or patterns of targets characterized by a peaked distribution of the target to target distance [17], we find that, in marked contrast with repeated statements in the literature, persistent random walks with exponential distribution of excursion lengths can minimize the search time, and in that sense performs better than any Lévy walk.

Model

We first recall the model that we have defined in the previous chapter (see Figure (3.1)). We consider a persistent random walker in discrete time and space, moving on a d–dimensional cubic lattice $\mathcal{L}$ of volume $V = X^d$, where a single target site is located. In practice we take $d = 2$ or $d = 3$, make use of periodic boundary conditions (reflecting boundary conditions lead to similar results) and consider the dilute regime $X \gg 1$. This geometry encompasses both cases of a single target in a confined domain, and of regularly spaced targets in infinite space with concentration $1/V$. The latter situation can be seen as a limiting case of target distribution with strong correlations, as opposed to the Poissonian case, and is biologically meaningful for example in the case of repulsive interactions between targets [17]. Note that here the lattice step size corresponds to the target size and is set to 1, which defines the unit length of the problem. At each time step, the random searcher has a probability p_1 to continue in the same direction, p_2 to go backward, and p_3 to choose an orthogonal direction, so that $p_3 = (1 - p_1 - p_2)/(2d - 2)$. Following [85], we denote $p_1 = p_3 + \epsilon$ and $p_2 = p_3 - \delta$, and set in what follows $\delta = 0$ for the

sake of simplicity. The probability of a ballistic excursion of l consecutive steps with unchanged direction is then $P(l) = (1-p_1)p_1^{l-1}$, and the persistence length of the walk can be defined as

$$l_p = \sum_{l=1}^{\infty} lP(l) = 1/(1-p_1), \tag{3.1}$$

where $p_1 = (1+(2d-1)\epsilon)/(2d)$. In chapter 2, we calculated analytically the search time $\langle \mathbf{T} \rangle$ between a given starting point and a given target, defined as previously as the MFPT. We will use this result to get the GMFPT, defined as the MFPT to a given target averaged over all possible starting positions *and over all initial velocities* of the searcher, and analyze its dependence on the persistence length l_p (or equivalently ϵ) and the volume X^d.

While the position process alone is non Markovian, the joint process of the position and velocity of the searcher is Markovian. We briefly recall that we can obtain an exact backward equation for the MFPT $\langle \mathbf{T} \rangle(\mathbf{r}, \mathbf{e}_i)$ to the target of position $\mathbf{r}_T$, for a random searcher starting from $\mathbf{r}$ with initial velocity $\mathbf{e}_i$, where $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ defines a basis of the lattice :

$$\langle \mathbf{T} \rangle(\mathbf{r}, \mathbf{e}_i) = p_1 \langle \mathbf{T} \rangle(\mathbf{r}+\mathbf{e}_i, \mathbf{e}_i) + p_2 \langle \mathbf{T} \rangle(\mathbf{r}-\mathbf{e}_i, -\mathbf{e}_i) + p_3 \sum_{\pm \mathbf{e}_j \in \mathcal{B}, j \neq i} \mathbf{T}(\mathbf{r} \pm \mathbf{e}_j, \pm \mathbf{e}_j) + 1. \tag{3.2}$$

After some transformations detailed in chapter 2, we can average the MFPT over all possible starting positions and velocities to obtain:

$$\text{GMFPT} = \frac{-\epsilon(V-1)}{1-\epsilon} + \frac{1+\epsilon^2}{1-\epsilon^2} \sum_{\mathbf{q} \neq \mathbf{0}} \frac{1}{1-h(\mathbf{q}, \epsilon)} \tag{3.3}$$

where

$$h(\mathbf{q}, \epsilon) = \frac{(\epsilon-1)^2}{d} \sum_{\mathbf{e}_j \in \mathcal{B}} \frac{\cos(\mathbf{q}.\mathbf{e}_j)}{1+\epsilon^2-2\epsilon\cos(\mathbf{q}.\mathbf{e}_j)} \tag{3.4}$$

and $\sum_{\mathbf{q} \neq \mathbf{0}}$ denotes the sum over all possible vectors $\mathbf{q} = (2\pi k_1/X_1, \dots, 2\pi k_d/X_d)$, $(k_1, \dots, k_d) \in [0, X_1 - 1] \times \dots [0, X_d - 1]$ except $\mathbf{q} = \mathbf{0}$. This exact expression of the search time for a non Markovian searcher constitutes the central result of this section. We discuss below its physical implications, based on two useful approximations, and see how we can minimize it.

We first consider the case where the persistence length is of the same order as the target size, that is $l_p = \mathcal{O}(1)$, or equivalently $\epsilon \ll 1$. In this regime the search time reads:

$$\text{GMFPT} \underset{\epsilon \ll 1}{=} A(\epsilon)(V-1) + \frac{\text{GMFPT}_0}{D(\epsilon)}, \tag{3.5}$$

where GMFPT_0 is the search time of a non persistent random walk ($\epsilon = 0$) which is known exactly [61]. $A(\epsilon)$ is a constant defined by

$$A(\epsilon) = (B_d - 1)\epsilon + \mathcal{O}(\epsilon^2) \tag{3.6}$$

where

$$B_d = \lim_{V \to \infty} \frac{2}{V} \sum_{\mathbf{q} \neq \mathbf{0}} \frac{\frac{1}{d} \sum_{\mathbf{e}_j \in \mathcal{B}} (1-\cos(2\pi \mathbf{q}.\mathbf{e}_j))^2}{\left(\frac{1}{d} \sum_{\mathbf{e}_j \in \mathcal{B}} 1-\cos(2\pi \mathbf{q}.\mathbf{e}_j) \right)^2} \tag{3.7}$$

is a dimension dependent numerical constant (for example $B_2 \simeq 2.72$). In turn, $D(\epsilon) = (1+\epsilon)/(1-\epsilon)$ is the diffusion coefficient of the persistent random walk normalized by the diffusion coefficient of the non persistent walk (case $\epsilon = 0$) [85]. Hence, in equation (3.5), GMFPT_0/D is the search time expected for a non persistent random searcher of same normalized diffusion coefficient D. Note that the persistence property yields a non trivial additive correction which scales linearly with the volume, and therefore should not be neglected; this could be related to the "residual" mean first passage time described in [208].

We next consider the case where the persistence length is much larger than the target size, that is $l_p \gg 1$, or equivalently $\epsilon \to 1$. In this regime the search time reads in the case $d = 2$:

$$\mathrm{GMFPT} = \frac{2(X-1)}{1-\epsilon} + \frac{(X-1)^2}{2} + (1-\epsilon)\frac{(X-1)(X+3)(X-2)}{12} + \mathcal{O}\left((1-\epsilon)^2\right). \qquad (3.8)$$

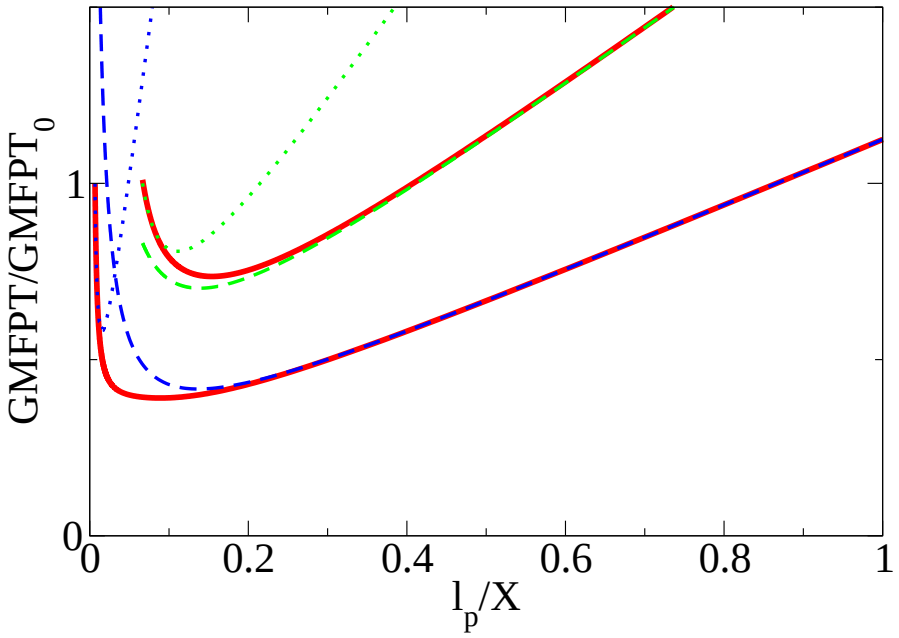

Figure 3.2: Search time for a persistent random walker GMFPT normalized by the search time for a non persistent walker (GMFPT_0) as a function of the rescaled persistence length, for $X = 10$ (upper set of curves) and $X = 100$ (lower set of curves). The red line stands for the exact result of equation (3.3), the dotted lines for the approximation $\epsilon \ll 1$ of equation (3.5), the dashed lines for the approximation $\epsilon \to 1$ of equation (3.8).

Figure 3.2 shows the search time for a persistent random walker (GMFPT) normalized by the search time for a non persistent walker (GMFPT_0) as a function of the rescaled persistence length, for $X = 10$ (upper set of curves) and $X = 100$ (lower set of curves). The red line stands for the exact result of equation (3.3), the dotted lines for the approximation $\epsilon \ll 1$ of equation (3.5), the dashed lines for the approximation $\epsilon \to 1$ of equation (3.8). The approximation of equation (3.5) is accurate as long as l_p is small. Equation (3.8) also provides a good approximation of the exact

result of equation (3.3) for $l_p \gg 1$. Note that the search time diverges for $l_p \to \infty$ because the searcher can then be trapped in extremely long unsuccessful ballistic excursions.

Both asymptotics $\epsilon \to 0$ and $\epsilon \to 1$ clearly show that the search time can be minimized as a function of ϵ or equivalently l_p, as seen in figure 3.2. The minimum can be obtained from the analysis of the exact expression (3.3), and reveals that the search time is minimized for $l_p = l_p^* \underset{X\to\infty}{\sim} \lambda X$ with $\lambda = 0.14...$ in 2D. Note that the asymptotic expression (3.8) yields a good analytical approximate of this minimum. This defines the optimal strategy for a persistent random searcher, which is realized when the persistence length has the same order of magnitude as the typical system size. We stress however that the numerical factor λ is non trivial and notably small. This optimal strategy can be understood as follows. In the regime $l_p \ll X$, the random walk behaves as a regular diffusion and is therefore recurrent for $d = 2$. The exploration of space is therefore redundant and yields a search time that scales in this regime as $V \ln V$ [209]. On the contrary for $l_p \gg 1$ exploration is transient at the scale of l_p and therefore less redundant. As soon as $l_p \sim X$ one therefore expects the search time to scale as V [209]. Taking l_p too large however becomes unfavorable since the searcher can be trapped in extremely long unsuccessful ballistic excursions, so that one indeed expects an optimum in the regime $l_p \sim X$. This argument suggests the following scaling of the optimal search time scaled by the non persistent case in the case $d = 2$:

$$\frac{\mathrm{GMFPT}_{l_p^*}}{\mathrm{GMFPT}_0} \propto \frac{1}{\ln(V)}, \tag{3.9}$$

which can indeed be derived from the asymptotic expression (3.8). This shows the efficiency of the optimal persistent search strategy in the large volume limit, as compared to the non persistent Brownian strategy.

Figure 3.3 shows the optimal search time scaled by the non persistent case as the function of the domain volume V. The black line stands for the numerical optimization of equation (3.3), the red line for the analytical optimization of equation (3.8), and the green dashed line for a fit $A/\ln(V)$. In the inset, the black line stands for the persistence length at the minimum, l_p^*, obtained by a numerical optimization of equation (3.3), as a function of X. The dashed red line is a linear fit of this curve ($l_p \simeq 0.14\ X + 3.6$). This figure shows that the scaling in $1/\ln(V)$ is indeed observed.

Note that for $d = 3$ a similar analysis applies. However, since in this case $\mathrm{GMFPT}_0 \propto V$ the scaled optimal search time tends to a constant.

Last, we compare the efficiency of the persistent and Lévy walk strategies. More precisely we consider a Lévy walker such that the distribution of the jump length of its successive ballistic excursions $P(l)$ follows a symmetric Lévy law of index μ and scale parameter c restricted to the positive axis, defined by the Fourrier transform $\widehat{P}(k) = e^{-c|k|^\mu}$, so that $P(l) \underset{l\to\infty}{\propto} 1/l^{1+\mu}$. For $0 < \mu \leq 1$, the persistence length is infinite, yielding in turn an infinite search time. We therefore focus on the regime $1 < \mu \leq 2$ and study numerically the dependence of the search time on both μ and l_p (which is set by c). Figure 3.4 shows the results of this numerical computation. The search time for a Lévy walk on a 2D lattice ($X = 50$) is plotted. Plots with circles stand for the search time for the following values of μ (from top to bottom): $\mu = 1.2$, 1.4, 1.6, 1.8 and 2. The black line stands for a persistent random walk for several values of ϵ. The abscissa stands for the persistence length l_p, function of c or ϵ.

Figure 3.4 shows that the search time can be minimized as a function of l_p for all $\mu \in]1, 2]$, and that this optimal value decreases when μ is increased. In particular the search time for the Lévy strategy is minimized when $\mu = 2$, *i.e.* when the length of the ballistic excursions has a finite

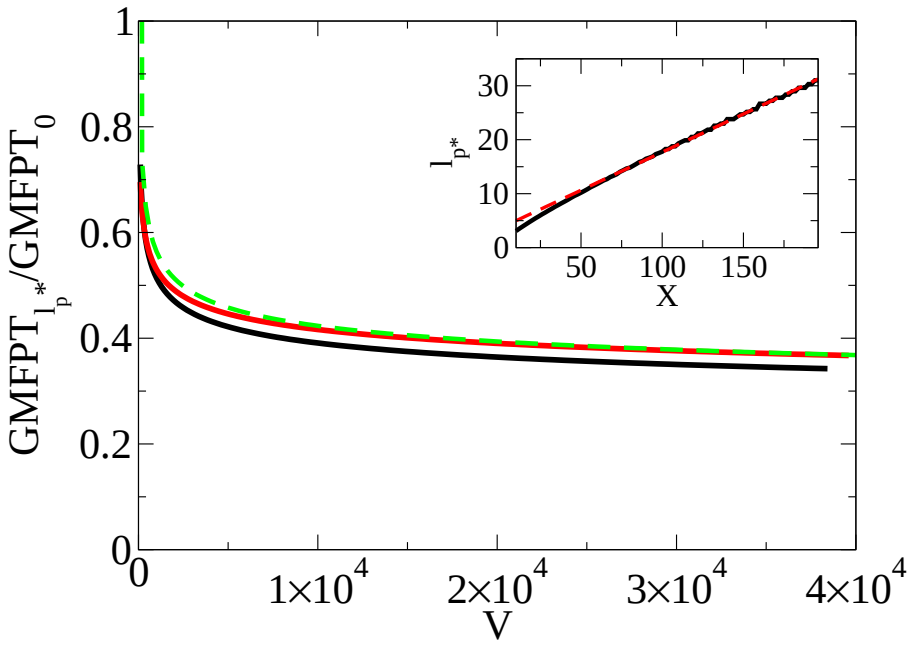

Figure 3.3: Optimal search time scaled by the non persistent case as the function of the domain volume V. The black line stands for the numerical optimization of equation (3.3), the red line for the analytical optimization of equation (3.8), and the green dashed line for a fit $A/\ln(V)$. In the inset, the black line stands for the persistence length at the minimum, l_p^*, obtained by a numerical optimization of equation (3.3), as a function of X. The dashed red line is a linear fit of this curve ($l_p \simeq 0.14\ X + 3.6$).

second moment so that the walk is no longer of Lévy type. As seen in Figure 3.4 the optimal persistent random walk strategy therefore yields to a search time shorter than any Lévy walk strategy.

This optimal persistent search strategy is in marked contrast with the simple ballistic motion obtained in the case of Poisson distributed targets, and shows that the distribution of targets plays a crucial role in the random search problem. In particular, in the biologically relevant cases of either a single target or patterns of targets characterized by a peaked distribution of the target to target distance, we find that, as opposed to repeated statements in the literature, persistent random walks with an exponential distribution of excursion lengths can minimize the search time, and in that sense perform better than any Lévy walk.

Quick summary

- The search process in a confined environment, or in a free environment with periodic targets, can be optimized by a persistent random walker.
- An optimized persistent walker finds an hidden target faster than any Lévy walk.

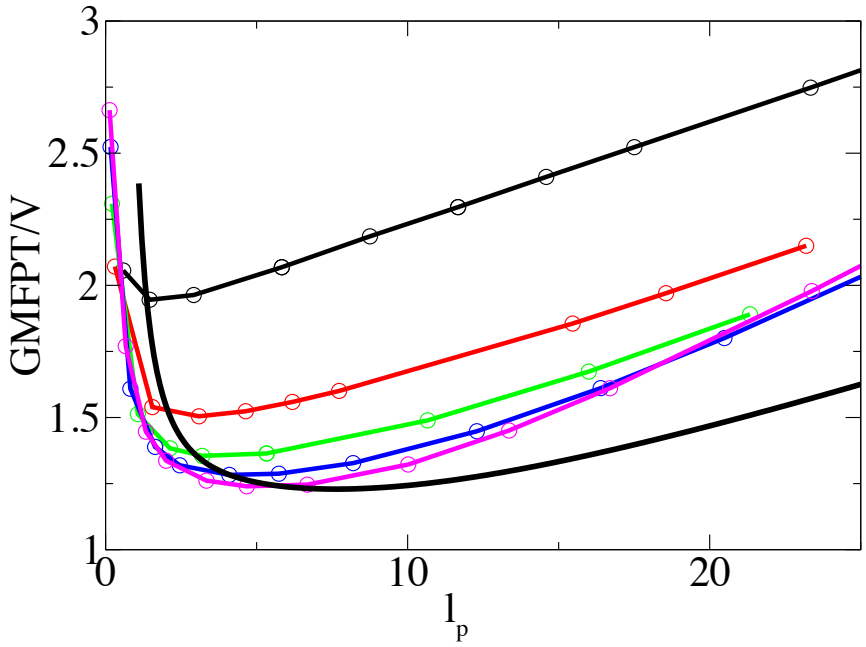

Figure 3.4: Numerical computation of the search time for a Lévy walk on a 2D lattice ($X = 50$). Plots with circles stand for the search time for the following values of μ (from top to bottom): $\mu = 1.2$, 1.4, 1.6, 1.8 and 2. The black line stands for a persistent random walk for several values of ϵ. The abscissa stands for the persistence length l_p, function of c or ϵ.

3.2 First-passage with moving target

Search problems can be a little more involved in the real life: in chemistry as well as in biology, the searching walker could look for a moving target. In such case, knowing where the target was initially is irrelevant. The formalism developed in the last section is not directly usable. We will see in this section how we can adapt our formalism to compute first-passage properties for moving targets. This case is really complex to deal with, we will thus consider a rather simple case, namely unidimensional systems, in order to extract usable expression, and to understand how physical parameters can optimize the search time.

Introduction

A classical example for the application of first-passage concepts with moving targets is the diffusion-regulation of bimolecular chemical reactions upon mutual diffusional encounter of the two molecular reactants in a three-dimensional liquid environment [215]. However, there are many cases in which the diffusional encounter of particles in a *one-dimensional* (or pseudo one-dimensional) environment becomes relevant. To name but a few examples, we mention the diffusional sliding motion of proteins or enzymes on DNA [163, 200], the diffusion of chemical reactants in the nanoconfinement of fluidic channels [114], and the relative motion of two aminoacids of a protein along the one-dimensional reaction coordinate [233].

In an unbounded environment, the encounter problem of two random walkers reduces to the consideration of the relative coordinate of the two walkers, with a diffusivity that equals the sum of the two individual diffusion constants. Similar to this unconfined case, for the diffusion of two particles on a finite domain with *periodic* boundary conditions, we may assume that one of the two walkers is fixed, and that the other diffuses with diffusivity $D' = 2D$, where D is the diffusion constant of a single walker. The problem is therefore equivalent to the first-passage problem for a single random walker, such that the mean first encounter time becomes

$$\langle \mathbf{T} \rangle = \frac{1}{2D'} r(L - r), \tag{3.10}$$

where r denotes the initial distance between the two random walkers and L the interval size.

If, however, we consider reflective or absorbing boundaries at the interval endpoints, the problem becomes more involved despite the apparent simplicity of this process. Indeed, we can no longer reduce the two-walkers problem to an effective one-walker scenario, because we now need two free parameters to characterize the system, for instance, the position of one walker and its distance to the second walker, instead of only the mutual distance in the unbounded case. We consider two cases: first, the probability that the random walkers meet *before* one of them is removed at the absorbing interval boundaries; and then, the typical encounter time between the two walkers when the boundaries are reflective. An analytic solution for the former problem has recently been presented [106],

$$P_M(x_1, x_2) = -\frac{2}{\pi} \Im \left\{ \log \left[\wp \left(\frac{\omega(x_2 + \imath x_1)}{L\sqrt{8}} \right) \right] \right\}. \tag{3.11}$$

Here, $\omega = \int_1^\infty (x(x-1))^{-3/4} dx \approx 5.244$, $\imath = \sqrt{-1}$, and the initial positions of the two walkers are x_1 and x_2 [106]. In equation (3.11), $\wp$ represents the Weierstrass elliptic function satisfying the differential equation

$$\wp'(x)^2 = 4\wp^3(x) - \wp(x). \tag{3.12}$$

The computation of the imaginary part of the logarithm of a complex number may become difficult and quite time consuming numerically. We will therefore try to find a simpler expression for this result. In the following we show that weak approximations lead to closed form expressions for the relevant quantities in terms of trigonometric, hyperbolic, and logarithmic functions. The accuracy of these results is corroborated by numerical simulations. We will first calculate the encounter probability of the two walkers before being annihilated by hitting the absorbing walls. We then proceed to the scenario of first encounter in the opposite case of reflecting boundaries, before concluding on relevant parameters for search problems with moving targets.

Computation being already quite cumbersome, we will restrict for this section to a confined 1D problem, the searcher and the target diffusing with the same diffusion coefficient D.

3.2.1 Encounter probability

We first consider the case of two absorbing boundaries and calculate the probability for encounter of the two walkers before either one of them becomes removed on hitting a boundary. This probability is simply a splitting probability: we compute the probability that the random walker hits first the moving target before hitting the absorbing wall.

To this end we rephrase the problem of two walkers in a one-dimensional domain by a single walker in a finite two-dimensional domain of size $L \times L$. We then seek the probability that,

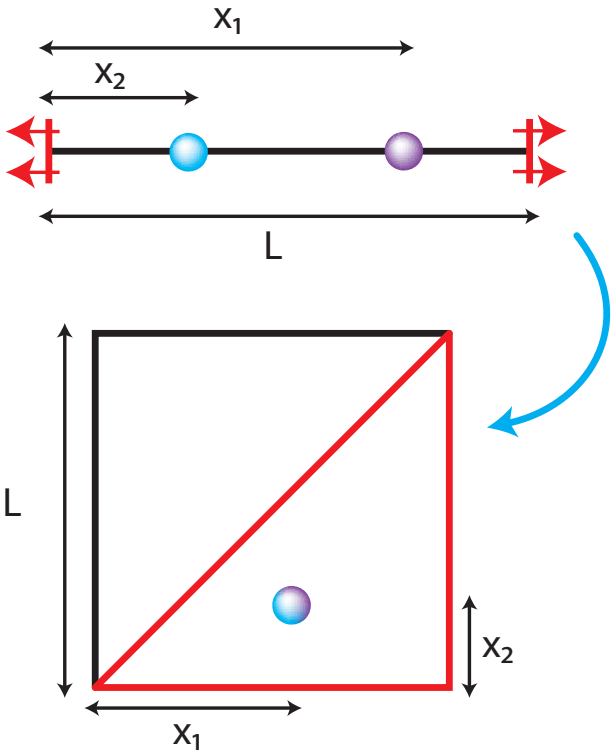

Figure 3.5: Scheme of the transformation 1D $\rightarrow$ 2D for a moving target with absorbing boundaries.

after starting from the point (x_1, x_2), this single random walker on the two-dimensional domain crosses the diagonal $x = y$, before touching the boundaries for the two coordinates, $x \in \{0, L\}$ or $y \in \{0, L\}$. Without loss of generality, we assume that $x_1 > x_2$. Figure 3.5 shows the transformation considered.

As the process is terminated when the two-dimensional random walker crosses the diagonal $x = y$, we use the method of images to determine the associated probability. First, we compute the probability that a two-dimensional random walker with diffusion coefficient D hits a given wall in an $L \times L$ square, whose boundaries are absorbing, without considering the absorbing diagonal. The propagator, namely the probability to be at position (x, y) at time t starting from (x_1, x_2) at $t = 0$, satisfy a standard diffusion equation, here in $2D$,

$$\frac{\partial P}{\partial t}(x, y, t|x_0, y_0) = D\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) P(x, y, t|x_0, y_0) \tag{3.13}$$

The solution for a square of size L with absorbing boundaries is

$$P_{\square}(x, y, t|x_1, x_2) = \frac{4}{L^2}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \sin\left(\frac{kx\pi}{L}\right) \sin\left(\frac{kx_1\pi}{L}\right) \sin\left(\frac{ly\pi}{L}\right) \sin\left(\frac{lx_2\pi}{L}\right) e^{-\frac{(k^2+l^2)D\pi^2 t}{L^2}}. \tag{3.14}$$

The probability to hit the wall at $x = L$ is then calculated using the Fick law. The probability to be absorbed at time t in a given position of the wall is given by the flux

$$\phi(x, y, t) = -D \int \vec{\nabla} P(x, y, t|x_1, x_2).\vec{dn} \tag{3.15}$$

where $\overrightarrow{dn}$ is a vector oriented normally to the surface considered, and $\overrightarrow{\bigtriangledown}$ the gradient operator. The probability to be absorbed by a given wall, that we will note $\text{Prob}_{\text{wall}}$, is the sum of this flux along the whole wall, integrated between $t = 0$ and ∞:

$$\text{Prob}_{\text{wall},\square}(x = L|x_1, x_2) = \int_{t=0}^{\infty} \int_{y=0}^{L} \phi(L, y, t) dy dt \tag{3.16}$$

After some maths detailed in Appendix 3, we obtain:

$$\text{Prob}_{\text{wall},\square}(x = L|x_1, x_2) = \frac{4}{\pi} \sum_{l=0}^{\infty} \frac{\sin\left(\frac{(2l+1)x_2\pi}{L}\right)}{2l+1} \frac{\sinh\left(\frac{(2l+1)x_1\pi}{L}\right)}{\sinh\left((2l+1)\pi\right)}. \tag{3.17}$$

We have here a "simple" infinite sum: we sum standard functions, with only real variables, over only one integer l. We will see that this equation allows a simpler result than the one of equation (3.11).

This expression simplifies when we introduce the approximation $\sinh((2l+1)\pi) \approx \exp((2l+1)\pi)/2$, leading to, after some computations detailed in Appendix 3:

$$\text{Prob}_{\text{wall},\square}(x = L|x_1, x_2) \simeq \frac{2}{\pi} \left(\arctan\left(\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(L-x_1)}{L}\right)} \right) - \arctan\left(\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(L+x_1)}{L}\right)} \right) \right) \tag{3.18}$$

whose maximal deviation from the exact expression (3.17) is around 0.2 %. Consistently, equation (3.18) vanishes if $x_1 = 0$, $x_2 = 0$, or $x_2 = L$. At $x_1 = L$, equation (3.18) gives approximately 1.

To link this result with the meeting probability of two random walkers in a one-dimensional domain with absorbing boundaries, we consider a random walker on the two-dimensional $L \times L$ square with diffusion coefficient D, starting from (x_1, x_2). Since the process terminates when $x = y$, the line $x = y$ is considered absorbing. We use the image method to set the probability on the line $x = y$ equal to 0. Namely, if we consider a $L \times L$ domain, and if we note $P_{\square}(x, y, t|x_1, x_2)$ the probability to be on (x, y) at time t starting from (x_1, x_2) at $t = 0$ on a square with absorbing boundaries, and $P(x, y, t|x_1, x_2)$ the same quantity when the line $x = y$ is absorbent, then

$$\forall (x, y) \in [0, L]^2 \ / \ x > y, \ P(x, y, t|x_1, x_2) = P_{\square}(x, y, t|x_1, x_2) - P_{\square}(x, y, t|x_2, x_1) \tag{3.19}$$

Figure 3.6 illustrate this image method: if we add a particle with a positive probability, and a fictive particle with a negative probability, we have an overall probability equal to 0 across the diagonal $x = y$, and along all the square boundary. This image method gives then a propagator $P(x, y, t|x_1, x_2)$ satisfying the diffusion equation as well as the boundary conditions. Since the solution is unique, we have the good one. One just has to take care that it only works for the lower triangle: the negative values in the upper triangle are not physical.

We can now compute the encounter probability. The probability to touch the $x = y$ line before the wall $y = 0$ or $x = L$ is 1 minus the probability to touch either the wall $y = 0$ or $x = L$. We can compute those probabilities to touch a given wall $\text{Prob}_{\text{wall}}$ using the expression found for an absorbing square $\text{Prob}_{\text{wall},\square}$, and the image method:

$$\text{Prob}_{\text{wall}}(x = L|x_1, x_2) = \text{Prob}_{\text{wall},\square}(x = L|x_1, x_2) - \text{Prob}_{\text{wall},\square}(x = L|x_2, x_1) \tag{3.20}$$

We thus obtain that the encounter probability for the original problem, noted her P_M, is

$$P_M(x_1, x_2) = 1 - \text{Prob}_{\text{wall}}(x = L|x_1, x_2) - \text{Prob}_{\text{wall}}(y = 0|x_1, x_2) \tag{3.21}$$

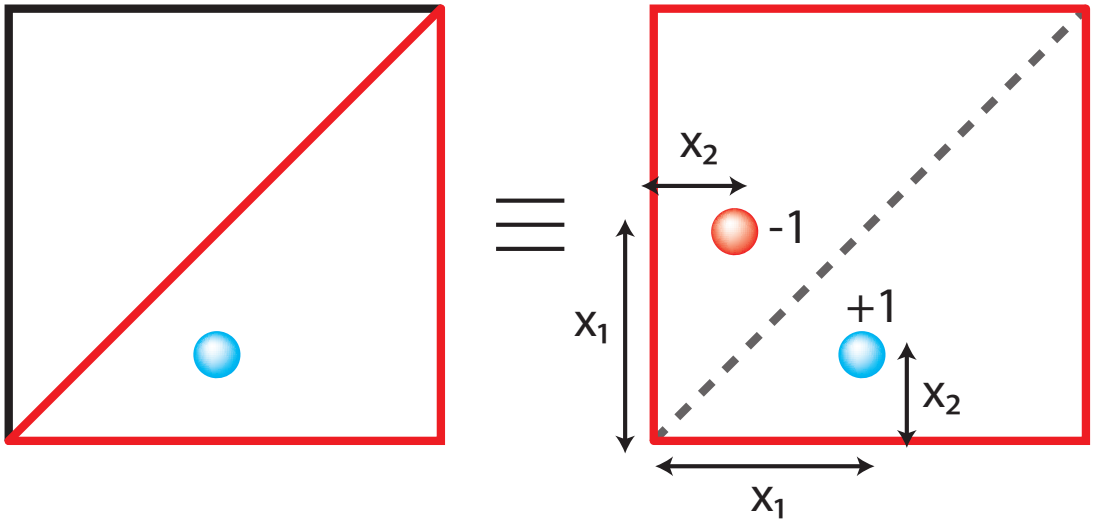

Figure 3.6: Scheme of the image method used to obtain $P(x, y, t|x_1, x_2)$ knowing $P_{\square}(x, y, t|x_1, x_2)$.

Using equation (3.17), we obtain an exact expression for the encounter probability (detailed computation are in Appendix 3)

$$P_M(x_1, x_2) = \frac{8}{\pi}\left(\sum_{l=0}^{\infty} \frac{\sin\left(\frac{(2l+1)x_2\pi}{L}\right)}{2l+1} \frac{\sinh\left(\frac{(2l+1)(L-x_1)\pi}{L}\right)}{\sinh\left((2l+1)\pi\right)} + \sum_{l=0}^{\infty} \frac{\sin\left(\frac{(2l+1)x_1\pi}{L}\right)}{2l+1} \frac{\sinh\left(\frac{(2l+1)x_2\pi}{L}\right)}{\sinh\left((2l+1)\pi\right)}\right) \tag{3.22}$$

As previously, we can approximate this expression

$$P_M(x_1, x_2) \simeq \frac{4}{\pi}\left(\arctan\left(\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi x_1}{L}\right)}\right) - \arctan\left(\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(2L-x_1)}{L}\right)}\right) + \arctan\left(\frac{\sin\left(\frac{\pi x_1}{L}\right)}{\sinh\left(\frac{\pi(L-x_2)}{L}\right)}\right) - \arctan\left(\frac{\sin\left(\frac{\pi x_1}{L}\right)}{\sinh\left(\frac{\pi(L+x_2)}{L}\right)}\right)\right) \tag{3.23}$$

We see that the encounter probability decreases with the distance $x_1 - x_2$, and that it decreases faster if an absorbing wall is in the nearby. We checked both exact et approximate results with numerical simulations.

Figure 3.7 shows the encounter probability P_M of two random walkers on the interval $[0, 1]$, initially placed at (x_1, x_2), where $x_1 = 0.5$ and x_2 varies between 0 and 0.5. We compare simulations results (black circles) with the approximate result given by equation (3.23) (red crosses) and the exact result given by equation (3.22) (green line). The inset shows the relative

difference between the approximate and the exact results. The relative error is of the same order of magnitude for all x_2 (below 0.2 %).

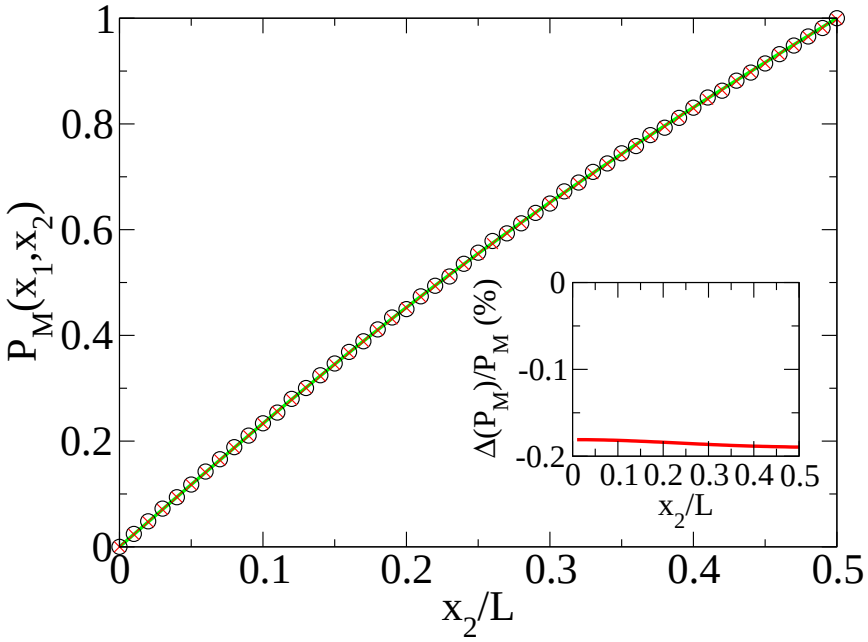

Figure 3.7: Encounter probability of two random walkers on the interval $[0, 1]$, starting in (x_1, x_2), where $x_1 = 0.5$, and $x_2 \in [0, 0.5]$. Black circles stand for simulations results, red crosses for the approximate expression of equation (3.23) and the green line for the exact result given by equation (3.22). (Inset) Relative difference between the approximate and the exact results in %. The relative error is of the same order of magnitude for all x_2 (below 0.2 %).

Figure 3.8 shows the encounter probability P_M of two random walkers initially placed at (x_1, x_2), where now $x_1 = 0.9$ and x_2 varies between 0 and 0.9. We still compare the simulations results (black circles) with the approximate result given by equation (3.23) (red crosses) and the exact result, equation (3.22) (green line). The inset shows the relative difference between the approximate and the exact results (red line). The relative error decreases as x_2 grows, but is always below 0.2 %.

Figures 3.7 and 3.8 show excellent agreement between the exact and the approximate formula of equations (3.22) and (3.23), as well as with numerical simulations of the encounter process. The approximate formula provided by equation (3.23) obviously provides an excellent approximation. Its numerical evaluation is significantly quicker than the exact result involving the Weierstrass elliptic function with complex argument. Moreover, the accuracy is sufficient for most purposes: the relative error is always smaller than 0.2 %. The approximate result is far easier to handle analytically than the exact expression (3.11) proposed in [106], and the dependence on the geometrical parameters is much more explicit.

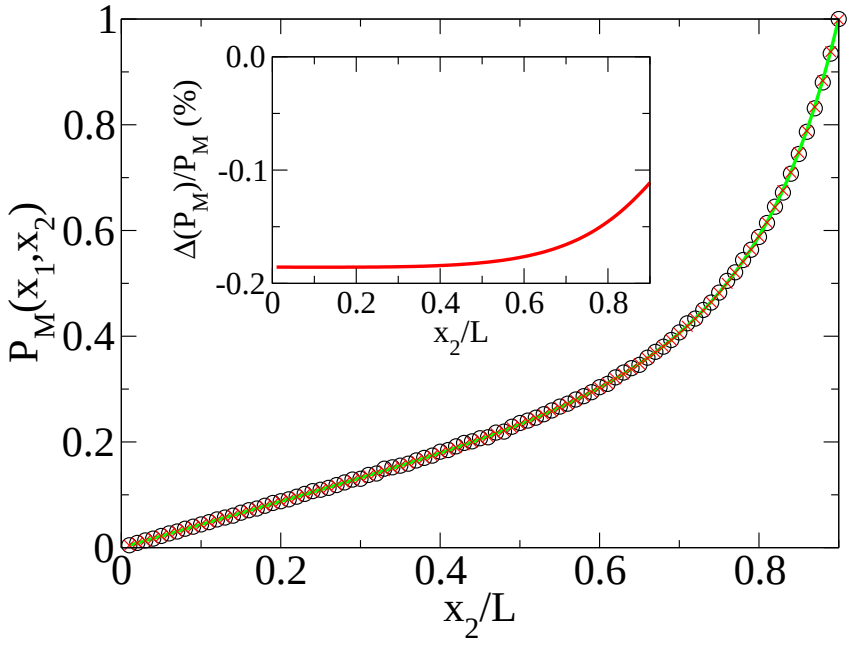

Figure 3.8: Encounter probability of two random walkers on the interval $[0, 1]$, starting in (x_1, x_2), where $x_1 = 0.9$, and $x_2 \in [0, 0.9]$. Black circles stand for simulations results, red crosses for the approximate expression of equation (3.23) and the green line for the exact result given by equation (3.22). (Inset) Relative difference between the approximate and the exact results in %. The relative error decreases as x_2 grows, and stays below 0.2 %.

3.2.2 First encounter time

We now consider the analogous problem with reflective boundaries at the endpoints of the one-dimensional interval. We determine the mean first encounter time, namely, the average of the first time when the two random walkers encounter each other. To compute this time, we again transform the problem of two random walkers in the one-dimensional domain into a two-dimensional single walker problem. This two-dimensional walker moves on one half of the square domain $L \times L$. In this half-square, the diagonal is absorbing while the two equilateral edges are reflecting. As shown in figure 3.9, we transform by symmetry this problem to the first exit time of a $\sqrt{2}L \times \sqrt{2}L$ square, where, if we assume $x_1 > x_2$, the initial coordinates are $(x_0, y_0) = ((x_1 - x_2)/\sqrt{2}, (x_1 + x_2)/\sqrt{2})$.

We start with the propagator on the $\sqrt{2}L \times \sqrt{2}L$ square, with absorbing boundaries. As previously, starting from (x_0, y_0), we have

$$P(x, y, t|x_0, y_0) = \frac{2}{L^2} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \sin\left(\frac{kx\pi}{\sqrt{2}L}\right) \sin\left(\frac{kx_0\pi}{\sqrt{2}L}\right) \sin\left(\frac{ly\pi}{\sqrt{2}L}\right) \sin\left(\frac{ly_0\pi}{\sqrt{2}L}\right) e^{-\frac{(k^2+l^2)D\pi^2 t}{2L^2}}.$$

To obtain the mean first exit time of this square, we will first compute the survival probability, $\mathscr{S}(t|x_0, y_0)$, namely the probability to still be in the square at time t, starting from (x_0, y_0). This

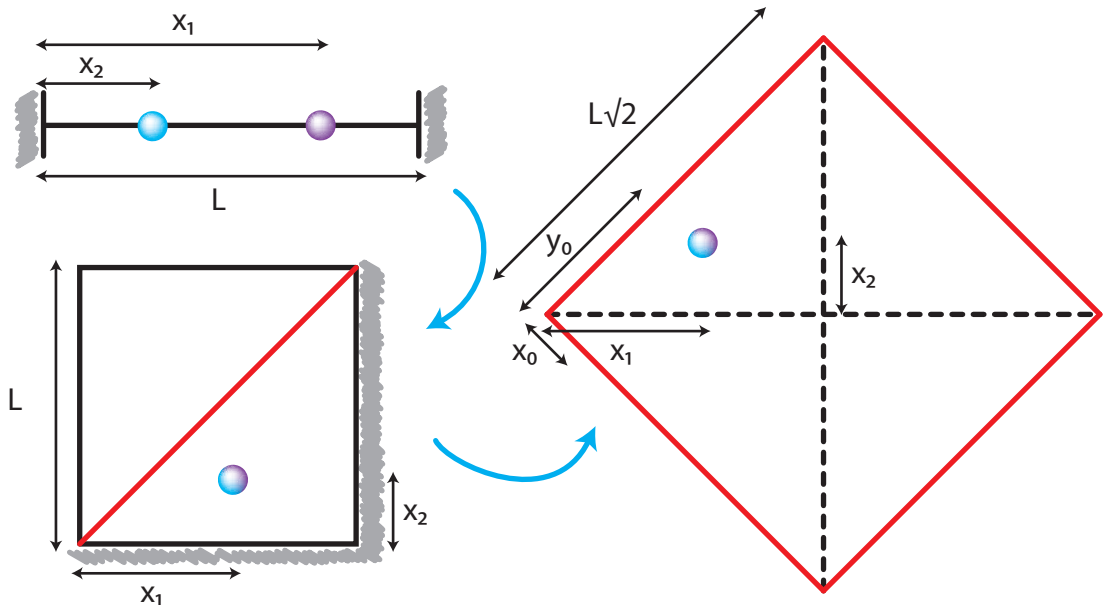

Figure 3.9: Scheme of the symmetry used to simplify the first encounter problem.

survival probability is simply the integrate over all square of the propagator

$$\begin{aligned}\mathscr{S}(t|x_0,y_0) &= \int_0^{\sqrt{2}L}\int_0^{\sqrt{2}L} P(x,y,t|x_0,y_0)dxdy \\ &= \frac{16}{\pi^2}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{2k+1}\frac{\sin\left(\frac{(2l+1)y_0\pi}{\sqrt{2}L}\right)}{2l+1}e^{-\frac{((2k+1)^2+(2l+1)^2)D\pi^2 t}{2L^2}}.\end{aligned}$$

We then deduce the mean first encounter time $\langle\mathbf{T}\rangle(x_0,y_0)$ as function of the initial positions x_0 and y_0 (computation details are in Appendix 3):

$$\begin{aligned}\langle\mathbf{T}\rangle(x_0,y_0) &= \int_0^{\infty}\mathscr{S}(t|x_0,y_0)dt \\ &= \frac{x_0}{2D}\left(\sqrt{2}L-x_0\right) \\ &\quad -\frac{8L^2}{D\pi^3}\sum_{k=0}^{\infty}\frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{(2k+1)^3}\frac{\sinh\left(\frac{(2k+1)y_0\pi}{\sqrt{2}L}\right)+\sinh\left(\frac{(2k+1)(\sqrt{2}L-y_0)\pi}{\sqrt{2}L}\right)}{\sinh\left((2k+1)\pi\right)}\end{aligned} \quad (3.24)$$

Figure 3.10 shows the mean first encounter time $\langle\mathbf{T}\rangle$ between two random walkers initially placed at (x_1,x_2) on a one-dimensional domain, where $x_1 = 0.5$, and x_2 varies between 0 and 0.5. We compare the simulations results (black circles) with the exact result, equation (3.24) (red line). The exact result fits perfectly with numerical simulations.

In equation (3.24) the first term is the mean first exit time of a one-dimensional random walker confined to a domain of size $\sqrt{2}L$, with diffusion coefficient D. The second term is the correction in a square domain. We approximate this second term in some limits. For instance, when the two particles are initially near a corner of the one-dimensional domain, i.e., $x_2 + x_1 \ll L$, such

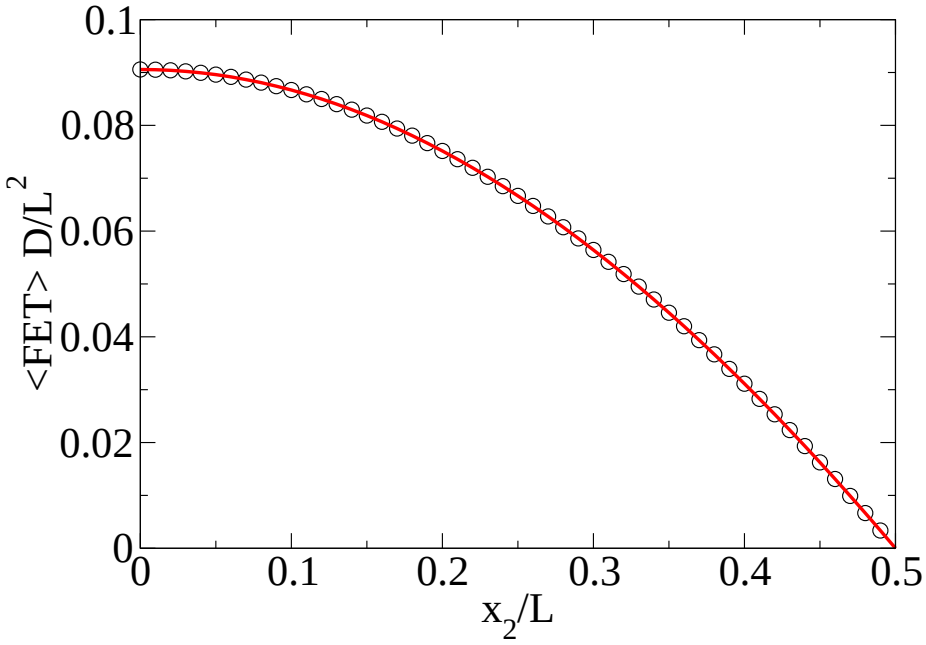

Figure 3.10: Mean first encounter time $\langle \mathbf{T} \rangle$ between two random walkers initially placed at (x_1, x_2), where $x_1 = 0.5$ and $x_2 \in [0, 0.5]$. The black circles stand for simulation results, the red line for the exact result of equation (3.24).

that $y_0/L \ll 1$ and $x_0/L \ll 1$, we have

$$\sum_{k=0}^{\infty} \frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{(2k+1)^3} \frac{\sinh\left(\frac{(2k+1)y_0\pi}{\sqrt{2}L}\right) + \sinh\left(\frac{(2k+1)(\sqrt{2}L-y_0)\pi}{\sqrt{2}L}\right)}{\sinh\left((2k+1)\pi\right)}$$
$$\approx \sum_{k=0}^{\infty} \frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{(2k+1)^3} \exp\left(-\frac{(2k+1)y_0\pi}{\sqrt{2}L}\right). \tag{3.25}$$

This expression can be simplified in the limit $x_0/L \ll 1$, using

$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{k^3} \exp(-ky) = x\mathrm{Li}_2(\exp(-y)) + \mathcal{O}(x^3), \tag{3.26}$$

where Li_2 is the dilogarithm defined as

$$\mathrm{Li}_2(z) = \sum_{k=0}^{\infty} \frac{z^k}{k^2}. \tag{3.27}$$

The series expansion of $\mathrm{Li}_2(z)$ around 1^- is

$$\mathrm{Li}_2(z) = \frac{\pi^2}{6} + (1 - \ln(1-z))(z-1) + \mathcal{O}\left((z-1)^2\right) \tag{3.28}$$

We thus obtain

$$\sum_{k=0}^{\infty} \frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{(2k+1)^3} \frac{\sinh\left(\frac{(2k+1)y_0\pi}{\sqrt{2}L}\right) + \sinh\left(\frac{(2k+1)(\sqrt{2}L-y_0)\pi}{\sqrt{2}L}\right)}{\sinh\left((2k+1)\pi\right)}$$
$$\simeq \frac{x_0\pi}{2\sqrt{2}L}\left(\frac{\pi^2}{4} + \left(\ln\left(\frac{y_0\pi}{\sqrt{2}L}\right) - 1 - \ln(2)\right)\frac{y_0\pi}{\sqrt{2}L}\right). \tag{3.29}$$

Thus, when both particles are initially close to an endpoint of the interval, we find

$$\langle \mathbf{T} \rangle(x_0, y_0) \simeq \frac{2x_0y_0}{D\pi}\left(1 + \ln(2) - \ln\left(\frac{y_0\pi}{\sqrt{2}L}\right)\right) - \frac{x_0^2}{2D}. \tag{3.30}$$

A similar approach leads to the same result when $\sqrt{2}L - y_0 \ll L$ and $x_0/L \ll 1$, if we replace y_0 by $\sqrt{2}L - y_0$. In the original variables we rewrite the previous expression as follows

$$\langle \mathbf{T} \rangle(x_1, x_2) \simeq \frac{x_1^2 - x_2^2}{D\pi}\left(1 + \ln(2) - \ln\left(\frac{(x_1+x_2)\pi}{2L}\right)\right) - \frac{(x_1-x_2)^2}{4D}. \tag{3.31}$$

Figure 3.11 shows the mean first encounter time $\langle \mathbf{T} \rangle$ between two random walkers initially placed at (x_1, x_2) on a one-dimensional domain, where $x_1 = 0.95$ and x_2 varies between 0 and 0.95. We compare the simulation results (black circles) with the exact result given by equation (3.24) (red line) and the approximation of equation (3.31) (green dashed line). The inset shows a zoom into the area around 0.95, where the approximation is valid ($\sqrt{2}L - y_0 \ll L$ and $x_0/L \ll 1$). We see that in the approximation validity area, approximation of equation (3.31) demonstrates excellent agreement with the simulations. This approximation gives a very good shape of the exact result, even if the precision decreases as the initial distance between the two random walkers grows.

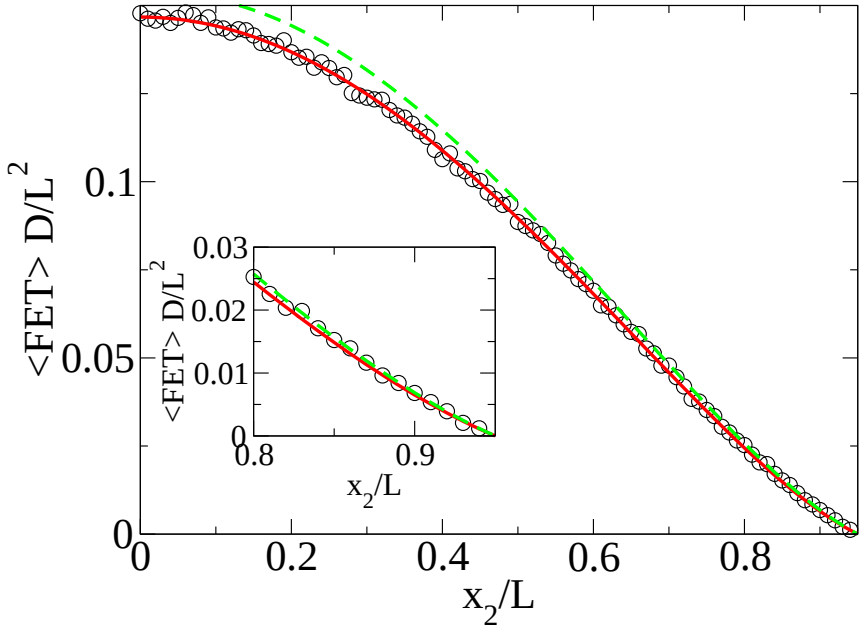

Figure 3.11: Mean first encounter time between two random walkers initially placed at (x_1, x_2), where $x_1 = 0.95$ and $x_2 \in [0, 0.95]$. The black circles stand for simulation results, the red line for the exact result of equation (3.24) and the dashed green line for the approximation of equation (3.31). (Inset) Zoom around the approximation validity area ($x_1 - x_2 \ll 1$), with the same data set.

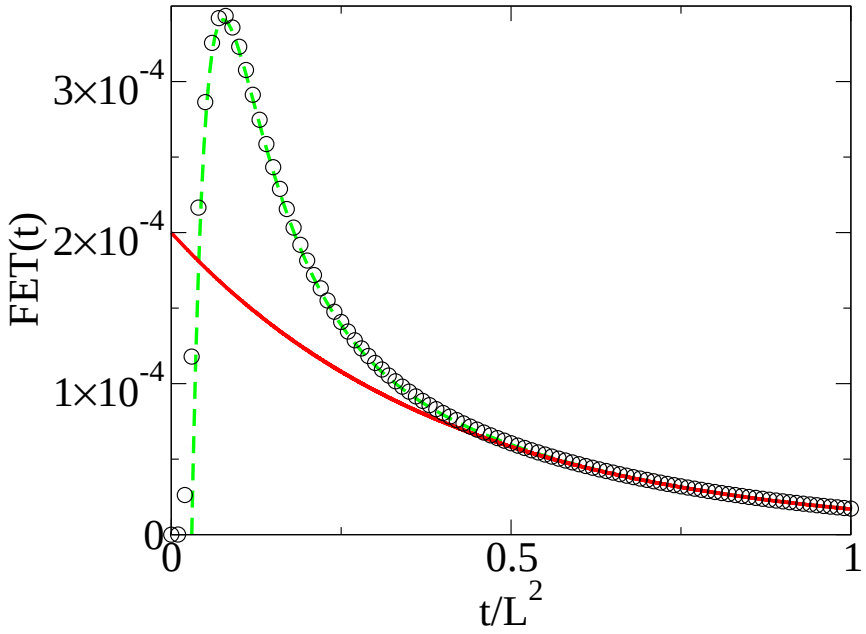

Figure 3.12: First encounter time distribution FET of two random walkers on a finite $1D$ domain, initially at (x_1, x_2), where $x_1 = 0$ and $x_2 = 0.5$, as a function of time t. Black circles stand for simulation results, red line for the leading exponential term of the exact result (3.32), and the dashed green line for the ten first leading terms of the same equation.

Finally, as a by-product, we can calculate the associated first passage density $\mathrm{FET}(x_0, y_0, t)$. This first passage density is in fact the first exit density of a square of size $L\sqrt{2} \times L\sqrt{2}$. It directly follows from the survival probability $\mathscr{S}(t|x_0, y_0)$ in equation (3.24), through

$$\begin{aligned} \mathrm{FET}(x_0, y_0, t) &= -\frac{\partial \mathscr{S}(t|x_0, y_0)}{\partial t} \\ &= \frac{8D}{L^2} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \left(\frac{2k+1}{2l+1} + \frac{2l+1}{2k+1} \right) \sin\left(\frac{(2k+1)x_0\pi}{L\sqrt{2}} \right) \sin\left(\frac{(2l+1)y_0\pi}{L\sqrt{2}} \right) \\ &\qquad \times \exp\left(-\frac{((2k+1)^2 + (2l+1)^2) D\pi^2 t}{2L^2} \right). \end{aligned} \tag{3.32}$$

Figure 3.12 shows the first encounter time distribution FET of two random walkers on a finite one-dimensional domain, initially positioned at (x_1, x_2), where $x_1 = 0$ and $x_2 = 0.5$, as a function of time t. We compare the simulations results (black circles) with the leading exponential of the exact result given by equation (3.32) (red line) and the first ten terms of the same expression (dashed green line).

Figure 3.12 shows that only the first few terms of this infinite sum are sufficient to reproduce

very accurately the simulations results. We observe that the leading term is given by

$$\mathrm{FET}(x_0, y_0, t) \sim \frac{16D}{L^2} \sin\left(\frac{x_0\pi}{\sqrt{2}L}\right) \sin\left(\frac{y_0\pi}{\sqrt{2}L}\right) \exp\left(-\frac{D\pi^2 t}{L^2}\right). \tag{3.33}$$

Conclusions

In term of search processes, we have here obtained analytically the impact of confinement on a search problem where both the searcher and the target move. We obtained a simple formula for the encounter probability of two random walkers on a finite one-dimensional domain with absorbing boundary conditions at the interval endpoints. The resulting infinite sum can be approximated by analytical functions to high accuracy. The obtained result is quite elegant mathematically, and facilitates significantly the analytical and numerical handling of this problem, compared to the rigorous mathematical result. For the first encounter time of two random walkers in the presence of reflective boundary conditions, analogous results are presented and approximations are obtained.

3.3 Search efficiency and network topology

We will now focus more specifically on random walks on networks, and try to identify the relevant parameters that influence the search efficiency. The networks are used here as discrete description of continuous random walk, either on Euclidian space (for Euclidian lattice) or on fractals (for scale-invariant networks). Random walks on networks are somehow simpler to deal with than continuous diffusion, when looking for first-passage observables. We will extensively use the pseudo-Green functions formalism introduced in the first chapter, as well as the pseudo-Green functions approximation. Networks allow to introduce some parameters that have no clear equivalent in continuous diffusion, such as the target connectivity, or the intensity of a link.

We will in the first sub-section study the Global Mean First Passage Time (GMFPT) as a measure of the search efficiency. We already saw with persistent random walkers that this observable could quantify, under some assumptions, the time needed to hit for the first time a hidden target. We will now analyze the influence of network topology on this GMFPT, and exhibit general boundaries on this observable. We will in particular see that this GMFPT is somehow proportional to $1/P_{\mathrm{stat}}(T)$, the stationary probability being proportional to the target connectivity.

The second sub-section will explore this dependence on the target connectivity directly on the MFPT, namely when the target position is known. To do so, we will have to assume a scaling form of the propagator that depends on the target connectivity, for a general scale invariant process. We will then extend the result of [66] presented in the first chapter. Once again, compact and non compact processes will behave completely differently.

The last sub-section will focus on the impact of a small network modification on the search time. If we want to minimize the search time of a random walker on a given network, knowing the initial and the target localization, we can modify directly the network topology. We will investigate the case where a link is removed, and where a link is added.

3.3.1 General bounds of GMFPT

If the target position is unknown, or symmetrically if the starting position is unknown, we can estimate the average time to reach for the first time the target by averaging the MFPT over

all possible starting points. In this first section, we will see how the network and the target characteristics influence the search time for a nearest-neighbor random walk.

If the walker starts at a random position, what can we say on the time necessary to find a hidden target while performing a nearest neighbor random walk? How the environment, for instance the space topology, does influence this search time? Many authors focused [3, 13, 78] on so-called "complex networks", in order to extract a link between the network topology and the averaged MFPT (we will define the average we use later). Some topological properties, introduced in chapter 1, have been found to be relevant: small-world property [13, 225], scale free property [3, 4], or fractal scalings [204]. For scale-invariant environments, we already presented in the first chapter how to link fractal dimensions d_f and d_w with mean first passage time from a given starting point toward a given target [66].

Following the seminal work of Montroll [142], many authors have focused on the MFPT averaged over the starting point of the walker [1, 2, 38, 103, 124, 206, 236, 238, 240], sometimes called the global mean first-passage time (GMFPT). The goal was to estimate the volume dependence of this GMFPT: if we double the available volume, how does the GMFPT evolve? The less it grows, the more efficient the search strategy is.

In 2009, a sub-linear dependence on the size N of the network of the GMFPT to the most connected node of a specific network was shown [238], and was interpreted as favorable for an efficient trapping. For such network, the search time through a simple nearest neighbor random walk is lower than expected: the network topology has been interpreted as allowing an efficient search process. It is not the random walk in itself that is changed, but the network construction that increases search efficiency, compared to a simple Euclidian lattice. This finding, in strong contrast with previously known results in the case of regular [142] or fractals [1, 38, 103, 124] lattices, has motivated an increasing number of works [2, 236, 239, 240, 242] that have tried to find examples of networks with high trapping efficiency, namely displaying weaker and weaker dependence on N of the GMFPT. Relying on these specific examples, the heterogeneity, and more precisely the scale-free property was put forward as advantageous [236, 238], whereas the fractal property was suggested to be unfavorable [240].

In this sub-section, we propose a general framework, applicable to a broad class of networks, which deciphers the dependence of the search efficiency on the network topology. More specifically, we will see how the GMFPT depends on the network size N, and how we can explain the recent results obtained on specific examples [2, 236, 238, 239, 240, 242]. We will first show on the example of a new set of networks that the GMFPT to the most connected node can scale as N^θ, with θ *arbitrarily close to 0* despite the fractal property of the network. We will see how this can be linked with the linear volume dependence of the MFPT previously stated [66]. We will then present an analytical approach which yields to

(i) rigorous bounds on the N dependence of the GMFPT, and

(ii) a simple criterion under which this bound is reached, which in particular provides a condition for a sub-linear scaling with N, which is independent of the scale-free, small-world, or fractal nature of the network.

Last, we show that a sub-linear scaling is never representative of the network, in the sense that the GMFPT averaged over the target site always scales faster than N.

Definition of the problem and notations.

We consider a set of graphs $\{\mathcal{G}_g\}_{g\in\mathbb{N}}$ where N_g denotes the number of sites of the graph $\mathcal{G}_g$ at generation g, such that $N_g \to \infty$ when $g \to \infty$.

We consider a discrete time random walker on $\mathcal{G}_g$. We assume that the transition probabilities ω_{ij} from site i to site j defining the walk are such that an equilibrium distribution P_{stat} satisfying detailed balance exists. We further assume that:

$$\sup_{X\in\mathcal{G}_g} P_{\text{stat}}(X) \underset{g\to\infty}{\to} 0. \tag{3.34}$$

We denote by $\text{FPT}_{S\to T}(\mathbf{T}=n)$ the probability that the walker reaches the target site T starting from site S for the first time after $\mathbf{T}=n$ steps, and write $\langle\mathbf{T}\rangle_{S\to T}$ for the MFPT from S to T. Note that this first average $\langle\cdot\rangle$ is taken over the realizations of the random walk. Taking the average of the MFPT over the starting point, we define the GMFPT according to:

$$\text{GMFPT}(T) = \overline{\langle\mathbf{T}_{S\to T}\rangle}_{S\neq T} = \frac{\displaystyle\sum_{S\neq T} P_{\text{stat}}(S)\langle\mathbf{T}\rangle_{S\to T}}{1-P_{\text{stat}}(T)}. \tag{3.35}$$

Note that this quantity depends on the target point T. Here the space average $\overline{\dots}$ is taken over the equilibrium distribution P_{stat}, and slightly differs from the definition used in [241, 236, 238, 240, 2] where the average is taken over the flat distribution. It can be checked numerically on networks recently studied in the literature that both definitions lead to the same scaling with N_g.

This GMFPT somehow measures the average time to reach the target if the target appears suddenly on the network, the random walker being distributed according to the stationary probability. This could be the case for an animal looking for food: the food appears at a given moment, and the animals are already scattered everywhere looking for food when this food (fruits, prey, ...) appears. We can thus assume that the animal has already reached the stationary distribution. A small GMFPT will be the sign of an efficient trapping. Since the search process is here a nearest neighbor random walk, only the network topology will be studied. To compare two different networks, one has to compare the GMFPT for a given size. If we focus on the large volume behavior, we will consider that the least this GMFPT depends on N, the more efficient the search process on this network is.

Efficient trapping on a fractal network.

We first exhibit a set of fractal networks which extends the so-called (u,u)-flowers [177] introduced in the first chapter, and whose GMFPT to the most connected node scales as N^θ, with θ arbitrarily close to 0. Note that a similar scaling has been reported in reference [119] for an example of non fractal network. The first generation of graph consists in two nodes connected by one link; then, at each iteration, every link is broken and replaced by k paths of $u \geq 2$ links. Figure 3.13 shows the example of the first three generations of this network for $k=3$ and $u=2$. As for the (u,v)–flowers, this network is fractal with a fractal dimension $d_f = \ln(ku)/\ln(u)$ since the diameter of the network at generation g is $L_g \sim u^g$ while the number of sites is $N_g \sim (ku)^g$ (the usual (u,u)–flowers in reference [177] correspond to the special case $k=2$).

Taking as target one of the initial nodes, it is easily seen that the GMFPT determination becomes a simple 1D problem since all the points $n(r)$ at the same distance r of the target are equivalent by symmetry, and thus lead all to the same $\langle\mathbf{T}\rangle(r)$. Noting next that $\forall r \in$

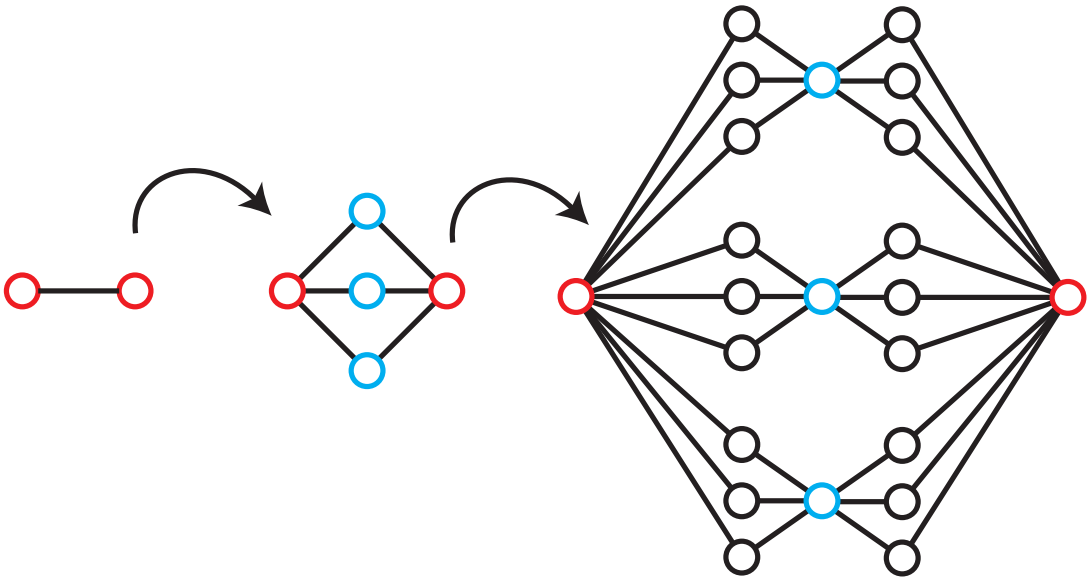

Figure 3.13: New fractal network: case of $k = 3$, $u = 2$ between generations $g = 1$ and $g = 3$ (left to right).

$[1, u^g - 1]$, $P_{\text{stat}}(r)n(r) = 2P_{\text{stat}}(T)$ and $P_{\text{stat}}(u^g)n(u^g) = P_{\text{stat}}(T)$, and using the classical 1D expression $\langle \mathbf{T} \rangle = r(2u^g - r)$ [168], we obtain the following exact expression:

$$\text{GMFPT}(T) = \frac{\sum_r P_{\text{stat}}(r)n(r)\langle \mathbf{T} \rangle(r)}{1 - P_{\text{stat}}(T)} = \frac{u^g\,(2u^g + 1)}{3} \propto N^{2\ln(u)/\ln(ku)}. \tag{3.36}$$

In other words, for k large enough, a random walk is *arbitrarily efficient* on this network to find a target place in one of the original node, despite its fractal property. Namely, we can obtain a network where the search time is independent of the network size if we place the target in the right position.

Lower bound of the GMFPT.

In order to gain understanding in the real parameters relevant to the scaling of the GMFPT with the network size N, we now derive a general lower bound for the GMFPT. This derivation follows from the generalization of the Kac formula [6, 64] which we briefly recall here.

We start from the discrete backward equation satisfied by the first passage density $\text{FPT}_{S\to T}$ for $n \geq 2$ (see [168]):

$$\text{FPT}_{S\to T}(n) = \sum_{j \neq T} \omega_{Sj}\,\text{FPT}_{j\to T}(n-1), \tag{3.37}$$

which is completed by $\text{FPT}_{S\to T}(n = 1) = \omega_{ST}$. Laplace transforming and averaging this equation over S (with a weight $P_{\text{stat}}(S)$ as in equation (3.35)) yields the generalized Kac formula

$$\frac{P_{\text{stat}}(T)}{1 - P_{\text{stat}}(T)} \left(\widehat{\text{FPT}}_{T\to T}(s) - e^{-s} \right) = (e^{-s} - 1)\langle \widehat{\text{FPT}}_{S\to T}(s) \rangle_{S \neq T}, \tag{3.38}$$

where

$$\widehat{\text{FPT}}_{S\to T}(s) \equiv \sum_{n=1}^{\infty} e^{-sn}\text{FPT}_{S\to T}(n). \tag{3.39}$$

This very general equation, derived in a similar form in [64], relates the distribution of the first return time to a site T to the distribution of the global first-passage time to T. Expanding equation (3.38) to first order in s yields the classical Kac formula $\langle \mathbf{T} \rangle_{T \to T} = 1/P_{\text{stat}}(T)$ [6, 64]. In turn, the second order in s gives :

$$\text{GMFPT}(T) = \frac{1}{2} \frac{P_{\text{stat}}(T) \langle \mathbf{T}^2 \rangle_{T \to T} - 1}{1 - P_{\text{stat}}(T)}. \tag{3.40}$$

Using next $\langle \mathbf{T}^2 \rangle_{T \to T} \geq \langle \mathbf{T} \rangle^2_{T \to T}$ and the classical Kac formula, the above exact expression gives a lower bound for the GMFPT:

$$\text{GMFPT}(T) \geq \frac{1}{2P_{\text{stat}}(T)}. \tag{3.41}$$

Note that this lower bound is in close analogy with the one obtained in [21] in the context of continuous space Pearson random walks in confinement.

We now discuss under which conditions this lower bound is reached. A strict equality requires the very restrictive condition that the variance of $\mathbf{T}_{T \to T}$ is zero. More generally we can discuss under which conditions the right and the left hand side of equation (3.41) share the same scaling in the large size limit. To do so, we consider a sequence of target sites $\{T_g \in \mathcal{G}_g\}_{g \in \mathbb{N}}$, which can be for instance hubs of the networks at each generation as in [2, 236, 238]. Using equation (3.41), and recalling that we have assumed $P_{\text{stat}}(T_g) \to 0$ for $g \to \infty$, we define the minimal scaling of the GMFPT for $g \to \infty$ by

$$\text{GMFPT}(T_g) = O(1/P_{\text{eq}}(T_g)). \tag{3.42}$$

Equation (3.40) then shows straightforwardly that this minimal scaling is reached as soon as the reduced variance of the first return time is finite in the large size limit, namely:

$$\frac{\langle \mathbf{T}^2 \rangle_{T \to T} - \langle \mathbf{T} \rangle^2_{T \to T}}{\langle \mathbf{T} \rangle^2_{T \to T}} = O(1). \tag{3.43}$$

We now show that this condition for a minimal scaling with the network size N_g is actually equivalent to the transience property of the random walk at the target site T_g in the large size limit. We first derive an alternative exact expression for the GMFPT. We will use the pseudo-Green functions, as introduced previously in chapter 1:

$$H_{S \to T} = \sum_{n=1}^{\infty} (P(T, n|S) - P_{\text{stat}}(T)), \tag{3.44}$$

where $P(T, n|S)$ is the propagator, namely the probability for the walker to be at T after n steps starting from S.

As shown in the first chapter:

$$\langle \mathbf{T} \rangle_{S \to T} = \frac{1}{P_{\text{stat}}(T)} (H_{T \to T} - H_{S \to T}). \tag{3.45}$$

Making use of the relation $P_{\text{stat}}(S) H_{S \to T} = P_{\text{stat}}(T) H_{T \to S}$, which follows from detailed balance (see also [148]), we obtain a second exact expression for the GMFPT:

$$\text{GMFPT}(T) = \frac{H_{T \to T}}{P_{\text{stat}}(T) (1 - P_{\text{stat}}(T))}. \tag{3.46}$$

This equation provides an alternative condition under which the minimal scaling is realized, given by $H_{T\to T} = O(1)$ in the large g limit. Using the definitions of recurrence and compacity, introduced in chapter 1, we can link this condition with the recurrence of the random walk. From the $H_{T\to T}$ definition of equation (3.44), the condition $H_{T\to T} = O(1)$ states that in the limit $g \to \infty$, a random walker returns on average only a *finite* number of times to T_g [46], which is equivalent to the fact that the random walk is transient at site T_g in this limit [168]. Conversely, equation (3.46) indicates that if the walk is recurrent at T_g for $g \to \infty$, that is if $H_{T_g\to T_g}$ diverges for $g \to \infty$, then GMFPT(T_g) grows faster than $1/P_{\text{stat}}(T_g)$.

The lower bound (3.41) and minimal scaling (3.42) for the GMFPT obtained above call for comments.

(i) First, our analysis puts forward a very general criterion to have a minimal scaling of the GMFPT with the network size, namely the type (transient or recurrent) [46] of the random walk at the target site. We stress that this criterion is independent of the scale-free, small world or fractal properties of the network. Note that for a generic set of graphs $\{\mathcal{G}_g\}_{g\in\mathbb{N}}$, the type of the random walk for $g \to \infty$ is a site dependent property [38, 46, 177].

(ii) Second, the minimal scaling (3.42) is fully determined by the equilibrium distribution at the target site, which is generally much easier to obtain than dynamical quantities, and which crucially depends on the connectivity of the target site. Let us take the classical example of a neares neighbor random walk, for which $\omega_{ij} = 1/k_i$ if i and j are neighbors and else 0, where k_i denotes the connectivity of site i. The minimal scaling of the GMFPT to a target T_g then reads $N_g\langle k\rangle/k_{T_g}$, where $\langle k\rangle$ is the connectivity averaged over all sites.

(iii) Note finally that in the case of a recurrent random walk at the target the minimal scaling is not realized, but the scaling of the GMFPT can however be sub-linear if the growth of the connectivity at the target is fast enough. In this case the scaling of the GMFPT depends on the scaling of $H_{T_g\to T_g}$, which generally depends both on the network and on the target T_g.

It is noteworthy that our analysis provides a comprehensive view of recent papers highlighting a sub-linear dependence of the GMFPT to a hub on different examples of networks.

(i) In the example of deterministic scale-free graph proposed in [2], the minimal scaling that we predict in equation (3.42) is indeed realized and the transience of the random walk at the target site (as defined above) is shown in the limit of large size (since the probability to come back at the hub in a finite time is null, as can be seen in equation (36) from [2] in the large size limit), in agreement with our approach.

(ii) The authors of [236, 238] have studied different examples of small world scale-free networks (Apollonian networks [236] and (u, v) flowers [238]) where the GMFPT to the main hub displays a sub-linear scaling. In these examples the scaling of GMFPT(T_g) is strictly faster than our predicted minimal scaling $1/P_{\text{stat}}(T_g)$ (and satisfies the upper bound given in (3.52)). Our criterion therefore implies that random walks on such structures are recurrent at the target site in the large size limit.

Bounds on the averaged GMFPT.

As demonstrated previously, the GMFPT highly depends on the target site, especially in the case of scale-free networks where the connectivity can be very heterogeneous. Therefore the GMFPT

to a specific target site cannot be taken as a general characteristic of the network. Actually, as we proceed to show the GMFPT averaged over target sites, defined by

$$\langle \text{GMFPT} \rangle = \sum_T P_{\text{stat}}(T) \text{GMFPT}(T), \tag{3.47}$$

has scaling properties with N_g which can widely differ from the case of a fixed target site studied above. The inequality (3.41) gives straightforwardly the following lower bound for $\langle \text{GMFPT} \rangle$ (see also [6]):

$$\langle \text{GMFPT} \rangle \geq \frac{N_g}{2}. \tag{3.48}$$

Hence, the averaged GMFPT always scales faster than N_g, and sub-linear scalings discussed above are point-wise properties which are *never* representative of the network. This general inequality sheds some light on the result obtained by Bollt and ben Avraham [38] in the case of a specific network ((1,2) flowers), where the GMFPT averaged over a fraction of nodes of the network scales sub-linearly with N_g, while the GMFPT averaged over all the nodes is linear with N_g.

Interestingly, we can also propose an upper bound for $\langle \text{GMFPT} \rangle$ following [6]. First we define (see also [12, 38]) the mean commute time as:

$$\tau_{ij} = \langle \mathbf{T} \rangle_{i \to j} + \langle \mathbf{T} \rangle_{j \to i}. \tag{3.49}$$

The quantity τ_{ij} can actually be bounded using the electrical analogy. Let us assign a unitary resistance to each link of the graph. Then it can be shown (see [57]) that the following general relation holds

$$\tau_{ij} = N_g \langle k \rangle R_{ij}, \tag{3.50}$$

where R_{ij} is the effective electrical resistance of the network between sites i and j. It is then straightforward to obtain that $R_{ij} \leq d_{ij}$ where d_{ij} is the distance between i and j. Indeed, d_{ij} is the resistance of a path of length d_{ij} between i and j, and any parallel paths can only lower the resistance. We therefore finally obtain:

$$\frac{N_g}{2} \leq \langle \text{GMFPT} \rangle \leq \frac{N_g \langle k \rangle \langle d \rangle}{2}, \tag{3.51}$$

where $\langle d \rangle$ is the weighted average over pairs of the point to point distance d_{ij}.

Importantly, this shows that the scaling of $\langle \text{GMFPT} \rangle$ is much more constrained than the scaling of the GMFPT for a fixed target. This is particularly striking in the case of small world networks for which $\langle d \rangle \propto \ln(N_g)$: hence in the case of small-world networks with finite $\langle k \rangle$, widespread in nature [13], this shows that $\langle \text{GMFPT} \rangle$ always scales linearly with N_g (up to log corrections). Note also that the bounds of equation (3.51) are compatible with the linear scaling of $\langle \text{GMFPT} \rangle$ with N_g reported in the case of Apollonian networks [110] and (1,2) flowers [38]. The conditions for which the scaling of each of the bounds in equation (3.51) is realized can also be discussed. As for the scaling of the lower bound, a sufficient condition for its realization is that for any sequence of targets $\{T_g \in \mathcal{G}_g\}_{g \in \mathbb{N}}$, the random walk is transient at T_g in the limit $g \to \infty$. Note however that this condition is not necessary, and the bound can be reached for networks having mixed type properties, as in the case of (1,2) flowers already mentioned [38]. As for the scaling of the upper bound, first notice that for any tree graph, R_{ij} is exactly the distance d_{ij} as discussed above using the electrical analogy. We conclude that for any tree the scaling of the upper bound is realized. In particular we find that $\langle \text{GMFPT} \rangle \sim N_g \ln(N_g)$ for *any* small world tree (see [237] for an example).

Additional comments are in order.

(i) First, equation (3.51) provides as a by-product an upper bound for the GMFPT itself, leading finally to:

$$\frac{1}{2P_{\text{stat}}(T)} \leq \text{GMFPT}(T) \leq \frac{N_g \langle k \rangle \langle d \rangle}{2P_{\text{stat}}(T)}. \tag{3.52}$$

(ii) Second, this upper bound for GMFPT inductively gives an upper bound of the trapping time in the case of a moving target using the Pascal principle [143].

(iii) Last, we underline that in the case of fractal networks, characterized by a fractal dimension d_f and a walk dimension d_w [204] an explicit scaling of $\langle \text{GMFPT} \rangle$ can be obtained (see [1, 38]). Indeed, using for instance the asymptotics of the MFPT between points separated by a distance d [66] and averaging over d, one gets the following scaling

$$\langle \text{GMFPT} \rangle \sim \begin{cases} N_g & \text{if } d_w < d_f \\ N_g \ln(N_g) & \text{if } d_w = d_f \\ N_g^{d_w/d_f} & \text{if } d_w > d_f \end{cases}, \tag{3.53}$$

which depends on the type of the random walk (transient if $d_f > d_w$, recurrent else).

Conclusion.

We have presented a general framework, applicable to a broad class of networks, which provides rigorous bounds on the size dependence of the GMFPT to a target site. We have shown that the GMFPT has the same scaling in the large size limit as this lower bound under the condition that the random walk is transient at the target site. This shows that the type of the random walk (transient or recurrent) is a crucial criterion to determine the scaling of the GMFPT, widely independent of its scale free, small world, or fractal properties. Those results reconciles recent works on GMFPT for random walks on various network examples. Additionally, we have demonstrated that the scaling of the GMFPT to a specific target is not a representative property of the network, since the target averaged GMFPT satisfies much more restrictive bounds, which in particular forbid any sub-linear scaling with the network size.

QUICK SUMMARY

We have presented a measure of the search efficiency, the GMFPT, and made explicit the target and the network dependence of this GMFPT, with analytical boundaries on the volume dependence. We have explicit expressions for fractal networks, and exact scalings for trees and for transient target nodes.

3.3.2 Target connectivity

A striking topological feature of many real-world complex networks is the wide distribution of the number k of links attached to a node – the connectivity – , as exemplified by the class of scale-free networks introduced in chapter 1, such as internet [86], biological networks [217], stock markets [117] or urban traffic [230]. For all those networks, the connectivity is distributed according to a power law. The impact of connectivity on transport properties has been put forward in [12, 119, 132, 209], where it was found in different examples that transport towards a target node can be favored by a high connectivity of the target.

We studied in the previous sub-section the impact of topological properties of a network on its transport properties, with a particular focus on complex networks[12, 38, 118, 148, 178]. We saw the target influence on the global mean first passage time (GMFPT) in equation (3.46). We will here look at the mean first-passage time (MFPT) [168] to a target node. This quantity is also an indicator of transport efficiency [1, 2, 93, 103, 110, 124, 238, 242]: after averaging over all couples at a given distance, the MFPT gives the average time to find a target located at a distance r. The GMFPT was global and target dependent, we will here go a step further in accuracy (and in complexity), by taking into account both the target connectivity and the source-target distance.

The dependence of the MFPT on geometric properties, such as the volume of the network and the source to target distance, is already known [29, 66, 68, 151], as presented in chapter 1: the starting position of the random walker plays a crucial role in the target search problem. We thus have two effects: a high connectivity and a small distance seem to minimize the MFPT. We will try to quantify the relative importance of distance and connectivity effects on transport properties on complex networks, which is equivalent to answer the following question: is it faster for a random walker to find either a close, or a highly connected target?

We will here propose a general framework, applicable to a broad class of networks, which deciphers the dependence of the MFPT both on the target connectivity and on the source to target distance, and provides a global understanding of recent results obtained on specific examples. Our approach highlights two strongly different behaviors depending on the so-called type – compact or non compact – of the random walk. In the case of non compact exploration, the MFPT is found to scale linearly with the inverse connectivity of the target, and to be widely independent of the starting point. On the contrary, in the compact case the MFPT is controlled by the source to target distance, and we find that unexpectedly the target connectivity is *irrelevant* for remote targets.

This analytical approach, validated numerically on various examples of networks, can be extended to other relevant first-passage observables such as splitting probabilities or occupations times [68].

Model and notations

We are interested in the MFPT denoted $\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)$ of a discrete Markovian random walker to a target $\mathbf{r}_T$, starting from a source point $\mathbf{r}_S$, and evolving in a network of N sites. We denote by $k(\mathbf{r})$ the connectivity (number of nearest neighbors) of site $\mathbf{r}$, and by $\langle k\rangle$ its average over all sites with a flat measure. The corresponding degree distribution is denoted by $p(k)$. We assume that at each time step n, the walker, at site $\mathbf{r}$, jumps to one of the neighboring site with probability $1/k(\mathbf{r})$. Let $P(\mathbf{r}, n|\mathbf{r}')$ be the propagator, as defined in chapter 1: the propagator is the probability that the walker is at $\mathbf{r}$ after n steps, starting from $\mathbf{r}'$. The stationary probability distribution is then given by $P_{\text{stat}}(\mathbf{r}) = k(\mathbf{r})/N\langle k\rangle$, and it can be shown that detailed balance yields the following symmetry relation :

$$P(\mathbf{r}, n|\mathbf{r}')P_{\text{stat}}(\mathbf{r}') = P(\mathbf{r}', n|\mathbf{r})P_{\text{stat}}(\mathbf{r}), \tag{3.54}$$

which will prove to be useful.

We consider networks with only short range degree correlations, namely such that $\langle k(\mathbf{r})k(\mathbf{r}')\rangle = \langle k\rangle^2$ for $|\mathbf{r}-\mathbf{r}'|$ larger than a cut-off distance R, where the average is taken over all pairs $\mathbf{r}, \mathbf{r}'$ with $\|\mathbf{r}-\mathbf{r}'\|$ fixed. $\|\mathbf{r}-\mathbf{r}'\|$ is here the chemical distance between $\mathbf{r}$ and $\mathbf{r}'$ (number of links in the shortest path linking $\mathbf{r}$ and $\mathbf{r}'$). This hypothesis is verified in particular by networks whose Pearson assortativity coefficient [147] is 0, such as Erdös-Rényi networks. It is however

less restrictive since local degree correlations can exist, and many networks actually comply with this assumption, as exemplified below. The hypothesis of short range degree correlations implies in particular that the degree distribution in a shell of radius $r > R$ is identical to the degree distribution $p(k)$ over the whole network, so that

$$\sum_{\mathbf{r}' \setminus |\mathbf{r}-\mathbf{r}'|=r} P_{\text{stat}}(\mathbf{r}') \simeq N_{\mathbf{r}}(r)/N \tag{3.55}$$

where $N_{\mathbf{r}}(r)$ is the number of sites $\mathbf{r}'$ such that $|\mathbf{r} - \mathbf{r}'| = r$. We then introduce the weighted average at distance r of a function f of two space variables defined by

$$\{f(\mathbf{r}, \mathbf{r}')\}_{\mathbf{r}'} = \frac{N}{N_{\mathbf{r}}(r)} \sum_{\mathbf{r}'/|\mathbf{r}-\mathbf{r}'|=r} f(\mathbf{r}, \mathbf{r}') P_{\text{stat}}(\mathbf{r}'), \tag{3.56}$$

and the standard flat average

$$\langle f(\mathbf{r}, \mathbf{r}') \rangle_{\mathbf{r}'} = \frac{1}{N_{\mathbf{r}}(r)} \sum_{\mathbf{r}'/|\mathbf{r}-\mathbf{r}'|=r} f(\mathbf{r}, \mathbf{r}'). \tag{3.57}$$

Scaling form of the propagator for scale invariant processes

We focus hereafter on transport processes having scale invariant properties, in order to use the approximation of chapter 1 for the infinite propagator. In this case, we will assume that the propagator in the infinite network size limit P_0, after averaging over points at a distance r from the starting point, satisfies the standard scaling for $|\mathbf{r} - \mathbf{r}'| > R$:

$$\langle P_\infty(\mathbf{r}, n|\mathbf{r}') \rangle_{\mathbf{r}} \propto n^{-d_f/d_w} \Pi\left(\frac{r}{n^{1/d_w}}\right), \tag{3.58}$$

where, as usual, the fractal dimension d_f characterizes the accessible volume $V_r \propto r^{d_f}$ within a sphere of radius r, and the walk dimension d_w characterizes the distance $r \propto n^{1/d_w}$ covered by a random walker in n steps. $\Pi(u)$ is here a scaling function: for instance $\Pi(u) = \exp(-u^{d_w})$ for a classical fractal medium [154].

A first central result of this section is to show numerically that the dependence of the propagator on the connectivity of the target site can be actually made explicit and reads

$$\langle P_\infty(\mathbf{r}, n|\mathbf{r}') \rangle_{\mathbf{r},k} \propto k n^{-d_f/d_w} \Pi\left(\frac{r}{n^{1/d_w}}\right), \tag{3.59}$$

where the average is taken over sites $\mathbf{r}$ at a distance r from $\mathbf{r}'$ with fixed connectivity k.

This scaling form is an assumption based on several observations:

i) the k dependence hypothesized in equation (3.59) satisfies the symmetry relation of equation (3.54);

ii) an average of equation (3.59) over k gives back the scaling of equation (3.58);

iii) an average over all possible starting points should gives a GMFPT scaling like $k \propto 1/P_{\text{stat}}(T)$. The k dependence of $H(\mathbf{r}_T|\mathbf{r}_T)$ is not trivial, but using equation (3.46), one could expect indeed a GMFPT growing like k.

We will see that this scaling leads in fact to a finer k dependence for the GMFPT that the one of equation (3.46).

Numerical simulations on various examples of scale invariant networks, such as percolation clusters and (u, v)–flowers (introduced in chapter 1) validate this assumption, as shown in Figures 3.14 and 3.15.

The upper plot of Figure 3.14 shows a rescaled propagator for a random walker diffusing on a supercritical 3D percolation networks ($p = 0.8$) of different sizes. We plot:

$$t^{d_f/d_w} \frac{\langle P(\mathbf{r}, t|\mathbf{r}')\rangle_{\mathbf{r},k}}{k} = f\left(\frac{r}{n^{1/d_w}}\right), \tag{3.60}$$

Following equation (3.59), we expect to obtain the same plot $\Pi(u)$ independently of the target connectivity. In Figure 3.14, we plot this rescaled propagator for different $k(\mathbf{r}_T)$. $\mathbf{r}_s$ is chosen in the center of the network, and t is small enough to avoid hits on the network's border. Black, red, green, blue, magenta and orange symbols stand respectively for $k = 1$, 2, 3, 4, 5 and 6. Circles, triangles, diamonds and squares stand respectively for networks of size 20^3, 25^3, 30^3 and 40^3. We see a perfect superposition of all plots, meaning that the scaling form of equation (3.59) is satisfied numerically.

The lower plot of Figure 3.14 shows the same rescaled propagator for a random walker diffusing on $(2, 2, 2)$–flowers (see chapter 1 for definition), for different target connectivity $k(\mathbf{r}_T)$. Black, red and green circles stand respectively for $k = 2$, 6 and 18. This network is not random as percolation clusters: we have a very marked order at short range, with in particular a very strong degree correlation at short range. If for small $r/t^{1/d_w}$, the scaling of equation (3.59) seems false, after some steps, we obtain a very decent merging of all curves.

The upper plot of Figure 3.15 shows the rescaled propagator for a random walker diffusing on a critical 3D percolation networks ($p = 0.2488$) of different sizes, and for different target connectivity $k(\mathbf{r}_T)$. $\mathbf{r}_s$ is chosen in the center of the network, and t is small enough to avoid hits on the network's border. Black, red, green, blue, magenta and orange symbols stand respectively for $k = 1$, 2, 3, 4, 5 and 6. Circles, and triangles stand respectively for networks of size 40^3 and 50^3. We see once again a very good superposition of all plots.

At last, the lower plot of Figure 3.14 shows the rescaled propagator for a random walker diffusing on $(3, 3)$-flowers (see chapter 1 for definition) for different $k(\mathbf{r}_T)$. Black, red, green, blue circles stand respectively for $k = 2$, 4, 8 and 16. This regular fractal have also a very strong short-range degree correlation, but for $r/t^{1/d_w}$ great enough, the scaling of equation (3.59) is verified.

We stress that the scaling form (3.59) is verified in the cases of both compact ($d_w > d_f$) and non compact ($d_w < d_f$) exploration. We believe that this result on its own can be important in the analysis of transport processes on networks. We show next that it enables to obtain the explicit dependence of first-passage properties on the connectivity of the target site.

Mean first-passage time

We now extend the theory developed in [66] to compute the MFPT of a discrete Markovian random walker to a target $\mathbf{r}_T$, and obtain explicitly its dependence on $k(\mathbf{r}_T)$. As shown in chapter 1, the MFPT satisfies the following exact expression[6, 64, 66, 148]:

$$\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S) = \frac{H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_S)}{P_{\text{stat}}(\mathbf{r}_T)}, \tag{3.61}$$

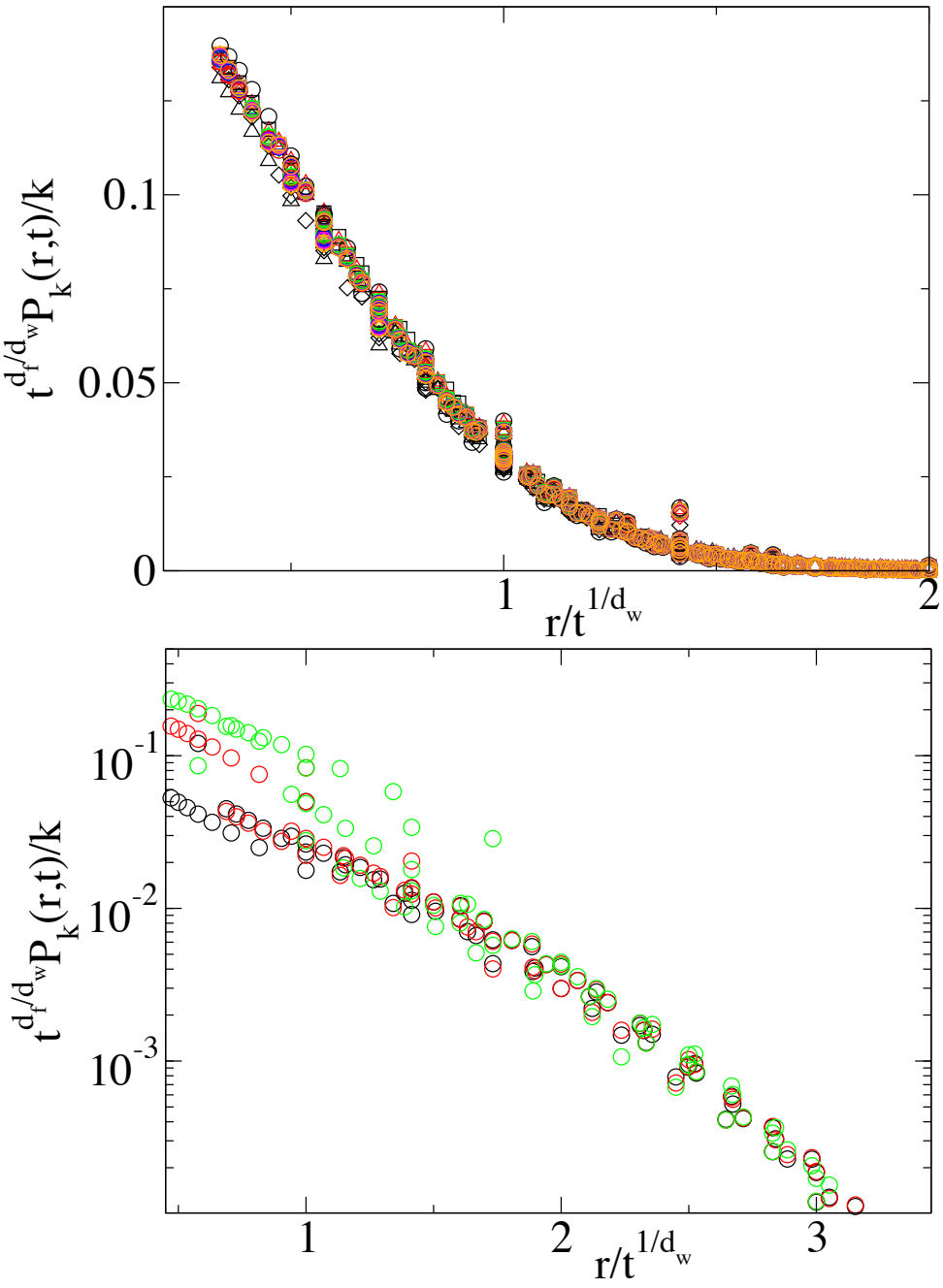

Figure 3.14: Plot of the propagator $P(\mathbf{r}_T, n|\mathbf{r}_S)$ for non compact explorations. (Upper plot) Supercritical 3D percolation networks ($p = 0.8$). Black, red, green, blue, magenta and orange symbols stand respectively for target connectivities $k(\mathbf{r}_T) = 1$, 2, 3, 4, 5 and 6. Circles, triangles, diamonds and squares stand respectively for networks of size 20^3, 25^3, 30^3 and 40^3. (Lower plot) $(2, 2, 2)$–flowers. Black, red and green circles stand respectively for target connectivities $k(\mathbf{r}_T) = 2$, 6 and 18.

where we have as usual the pseudo-Green function of the problem defined as[14]:

$$H(\mathbf{r}|\mathbf{r}') = \sum_{n=1}^{\infty} (P(\mathbf{r}, n|\mathbf{r}') - P_{\text{stat}}(\mathbf{r})) \tag{3.62}$$

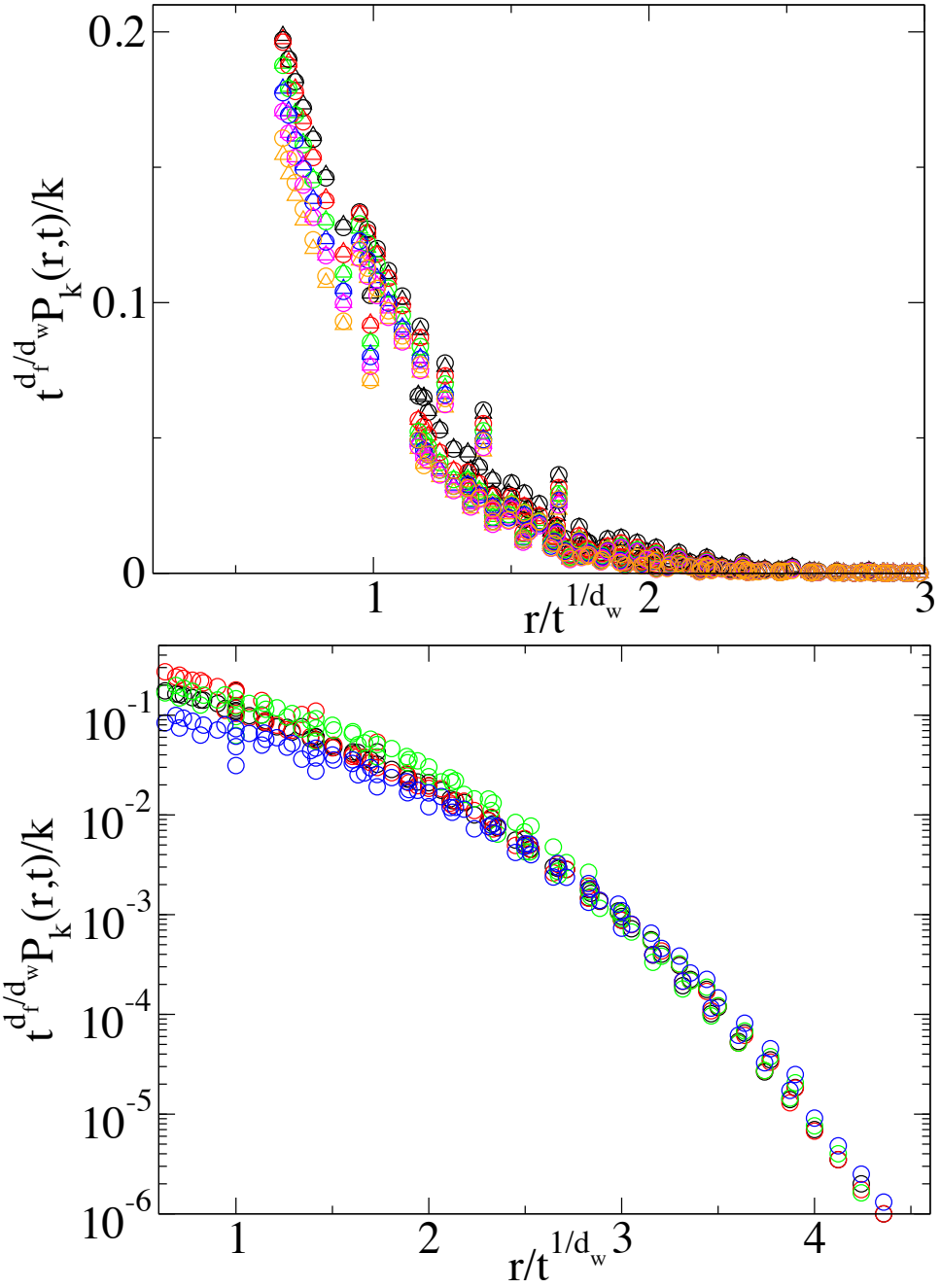

Figure 3.15: Plot of the propagator $P(\mathbf{r}_T, n|\mathbf{r}_S)$ for compact explorations. (Upper plot) critical 3D percolation networks ($p = 0.2488$). Black, red, green, blue, magenta and orange symbols stand respectively for target connectivities $k(\mathbf{r}_T) = 1$, 2, 3, 4, 5 and 6. Circles, and triangles stand respectively for networks of size 40^3 and 50^3. (Lower plot) $(3, 3)$–flowers. Black, red, green, blue circles stand respectively for target connectivities $k(\mathbf{r}_T) = 2$, 4, 8 and 16.

Note that averaging equation (3.61) for $\mathbf{r}_S$ covering the nearest neighbors of $\mathbf{r}_T$ gives the expression of the averaged MFPT $\langle\mathbf{T}\rangle_{\rm Kac}(\mathbf{r}_T)$ expected from Kac formula [6, 65]:

$$\langle\mathbf{T}\rangle_{\rm Kac}(\mathbf{r}_T) = \langle\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle_{\mathbf{r}_S}|_{\|\mathbf{r}_T-\mathbf{r}_S\|=1} = \frac{1}{P_{\rm stat}(\mathbf{r}_T)} - 1 = \frac{N\langle k\rangle}{k(\mathbf{r}_T)} - 1, \tag{3.63}$$

which we will use below.

Following [66], we consider the large N limit of equation (3.61). Making use of the symmetry relation of equation (3.54), we obtain

$$P_{\text{stat}}(\mathbf{r}_T)\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S) \sim G_0(\mathbf{r}_T|\mathbf{r}_T) - \frac{k(\mathbf{r}_T)}{k(\mathbf{r}_S)} G_0(\mathbf{r}_S|\mathbf{r}_T). \tag{3.64}$$

Here $\sim$ denotes equivalence for large N and G_0 is the usual infinite space Green function defined by

$$G_0(\mathbf{r}|\mathbf{r}') = \sum_{n=1}^{\infty} P_0(\mathbf{r}, n|\mathbf{r}'), \tag{3.65}$$

It is useful to notice that this leading term of the MFPT still satisfies the Kac formula (3.63). We next take the weighted average of equation (3.64) over the source points and obtain:

$$P_{\text{stat}}(\mathbf{r}_T)\mathbf{T}_{\mathbf{r}_T}(r) \sim G_0(\mathbf{r}_T|\mathbf{r}_T) - \frac{k(\mathbf{r}_T)}{\langle k \rangle} \langle G_0(\mathbf{r}_S|\mathbf{r}_T)\rangle_{\mathbf{r}_S}, \tag{3.66}$$

where we defined $\mathbf{T}_{\mathbf{r}_T}(r) \equiv \{\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\}_{\mathbf{r}_S}$. Substituting the scaling (3.58) in Eq. (3.66) then yields the large N equivalence of the MFPT to a target site $\mathbf{r}_T$ averaged over sources, which is valid for $r > R$:

$$\mathbf{T}_{\mathbf{r}_T}(r) \sim N\langle k \rangle \left(A_k + B r^{d_w - d_f} \right). \tag{3.67}$$

In this expression the constant A_k depends on the connectivity k of the target and B is a constant independent of k ad r, which depends on the scaling function Π. We now distinguish two regimes depending on the compact or non compact nature of the transport process, and focus on the large r regime.

Compact case $d_w \geq d_f$

In the compact case, $d_w \geq d_f$, which corresponds to recurrent random walks, we obtain that the MFPT scales in the large r limit as

$$\mathbf{T}_{\mathbf{r}_T}(r) \sim N\langle k \rangle B r^{d_w - d_f}. \tag{3.68}$$

This shows that unexpectedly the MFPT is asymptotically independent of the connectivity of the target, while the dependence on the distance r is crucial. Equation (3.67) is valid for r large enough (typically $r > R$). The dependence of A_k on k, which impacts on the MFPT for r small only, can be estimated by assuming that this expression still holds approximately for short distances. Following [28], we take $r = 1$ in equation (3.67) and use the Kac formula (3.63) to obtain:

$$1/k \simeq A_k + B, \tag{3.69}$$

which provides the k-dependence of A_k. We next aim at evaluating B. We introduce the weighted average of the MFPT over the target point, $\tau(r)$:

$$\tau(r) = \{\langle \mathbf{T}(\mathbf{r}_T, \mathbf{r}_S\rangle_{\mathbf{r}_S}\}_{\mathbf{r}_T} = \sum_{\mathbf{r}_T, \mathbf{r}_S / \ \|\mathbf{r}_T - \mathbf{r}_S\| = r} P_{\text{stat}}(\mathbf{r}_T)\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S). \tag{3.70}$$

Using equation (3.69), this quantity writes:

$$\tau(r) \sim N\left(1 + B\,\langle k \rangle\,(r^{d_w - d_f} - 1)\right). \tag{3.71}$$

In the case of compact exploration, the continuous space limit can be defined (see [28]) and imposes $\tau(r \to 0) = 0$. This extra equation, based on the existence of a continuous limit, enables to evaluate B as $B = 1/\langle k \rangle$. Note that for fractal trees ($d_w - d_f = 1$) we recover the exact result $\tau(r) = Nr$. Finally one has :

$$\mathbf{T}_{\mathbf{r}_T}(r) \sim N\langle k \rangle \left(\frac{1}{k} + \frac{1}{\langle k \rangle} \left(r^{d_w - d_f} - 1 \right) \right), \tag{3.72}$$

which fully elucidates the dependence of the MFPT on k and r. We recall here that this expression is originally derived for r large, and that the small r regime relies on the less controlled assumption that the scaling form of the propagator (3.59) holds for any distance r, and in particular that a continuous limit exists. It will however prove numerically to be accurate in various examples for all r values.

Figure 3.16 shows the MFPT for critical Erdös-Rényi networks, as a function of the source-target distance r, for a various target connectivity k. From up to down, black symbols correspond to $k = 1$, red to $k = 2$ and blue to $k = 3$. Circles stand for a network of size $N = 1000$, triangles stand for $N = 2000$. The straight lines stand for the zero-constant formula ($\langle k \rangle = 2$) of equation (3.72). Critical Erdös-Rényi are defined in chapter 1: they can be seen as percolation cluster ($p = 1/N$) on a complete graph. d_f has been estimated $1,9 - 2,0$[203] and we computed numerically $d_w \simeq 2,9$. We see that equation (3.72) is in very good agreement with numerical simulations.

Figure 3.17 shows the MFPT random $(2,2)$-flowers ($d_w = 2.5$ and $d_f = 1.9$), as a function of the source-target distance r, for a various target connectivity k. From up to down, black symbols stand for $k = 2$, red for $k = 3$, green for $k = 4$, blue for $k = 5$ and magenta for $k = 6$. Circles stand for simulations results, straight lines for $1/k + 1/\langle k \rangle (r^{d_w - d_f})$ ($\langle k \rangle = 3$) of equation (3.72). The prediction matches very well numerical simulations

Figure 3.18 shows the MFPT on compact Kozma networks[125] ($\alpha = 2.5$) of size $X = 50$, as a function of the source-target distance r, for a various target connectivity k. Kozma networks are defined in chapter 1, we briefly recall that in 1D, these networks are compact for $\alpha > 2$ and non compact for $\alpha < 2$. Symbols stand for simulation result: black is $k = 3$, red $k = 4$ and green $k = 5$. The expected scaling is in $r^{0.5}$: circles stand for simulation results, straight lines stand for $1/k + 1/\langle k \rangle (r^{0.5} - 1)$ ($\langle k \rangle = 2.5$) of equation (3.72).

Non compact case $d_w < d_f$

In the non compact (or transient) case, $d_w < d_f$, we obtain that the MFPT scales in the large r limit as

$$\mathbf{T}_{\mathbf{r}_T}(r) \sim N\langle k \rangle A_k. \tag{3.73}$$

This shows that the MFPT is independent of r for r large, as was already discussed in the literature [66]. The dependence on k is now fully contained in the constant A_k, which we now determine. Following [119], we assume that the FPT distribution is proportional to $\exp(-Akt/(N\langle k \rangle))$, with $A = \mathcal{O}(1)$, and widely independent of r in agreement with the result obtained in Eq.(3.73) for the first moment. This implies that the GMFPT, as defined in the beginning of this chapter and denoted by GMFPT, scales as

$$\text{GMFPT} = \{\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\}_{\mathbf{r}_S} \propto N\langle k \rangle / k, \tag{3.74}$$

where we used the weighted average of equation (3.56) on all possible $\mathbf{r}_S$. We recall the exact result derived in the first section of this chapter:

$$\text{GMFPT} = \frac{H(\mathbf{r}_T|\mathbf{r}_T)}{P_{\text{stat}}(\mathbf{r}_T)}, \tag{3.75}$$

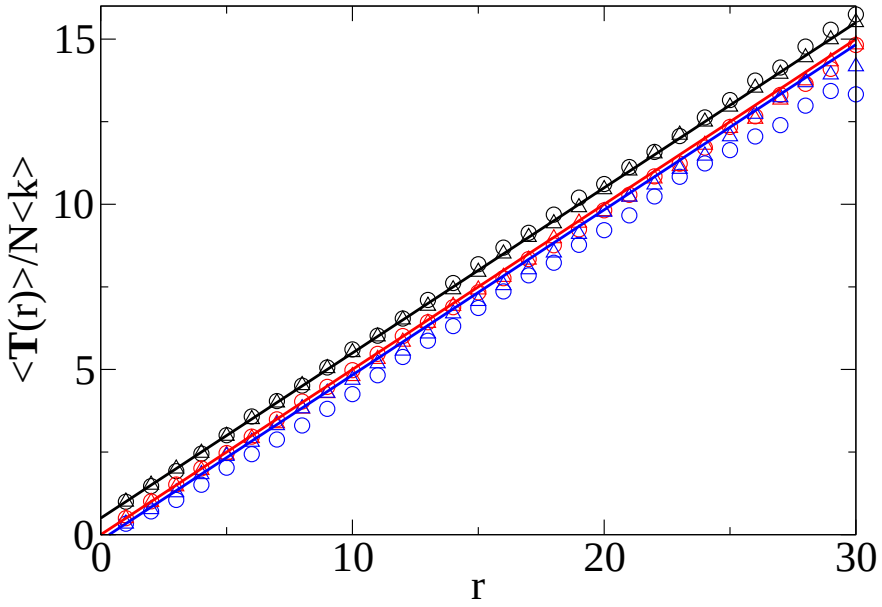

Figure 3.16: Rescalled MFPT ($\langle \mathbf{T}_{\mathbf{r}_T}(r) \rangle / N\langle k \rangle$) for critical Erdös-Rényi networks for various target connectivities k. From up to down, black symbols correspond to $k = 1$, red to $k = 2$ and blue to $k = 3$. Circles stand for a network of size $N = 1000$, triangles stand for $N = 2000$. The straight lines stand for the zero-constant formula ($\langle k \rangle = 2$) of equation (3.72).

where use the approximation $1 - P_{\text{stat}}(\mathbf{r}_T) \simeq 1$. We thus deduce from this equation that $H(\mathbf{r}_T|\mathbf{r}_T)$, and therefore asymptotically the infinite space Green function $G_0(\mathbf{r}_T|\mathbf{r}_T)$, is independent of k in the case of non compact exploration. As shown below, this has been checked numerically in Figure 3.22.

Identifying in equation (3.67) $A_k = G_0(\mathbf{r}_T|\mathbf{r}_T)/k$, which is finite in the case of non compact exploration, we finally obtain:

$$\mathbf{T}_{\mathbf{r}_T}(r) \sim N\langle k \rangle \left(\frac{A}{k} - B r^{d_w - d_f} \right). \tag{3.76}$$

Figure 3.19 shows the MFPT as a function of the source-target distance r, for a various target connectivity k, on supercritical 3D percolation network ($p = 0.8$). For this network, $d_w \simeq 2$ and $d_f = 3$, the exploration is non compact. From up to down, blue symbol stand for $k = 4$, magenta for $k = 5$ and orange for $k = 6$. Circles stand for simulations results, straight lines for a fit by $A/k + Br^{d_w - d_f}$, with $A \simeq 2.33$ and $B \simeq 0.8$. We observe that after a certain distance, Figure 3.19 shows a good agreement of numerical simulations with equation (3.76). As explained previously, our result does not apply very well in the small r area, since it relies on a scaling hypothesis on the propagator that is not fully satisfied for small r. But our general result (simple shift proportional to $1/k$ between the different curves) is still true as soon as r is great enough.

As in the compact case this expression is valid for r large, and becomes hypothetical for r small. It reveals that in the case of non compact exploration, the MFPT is independent of r for

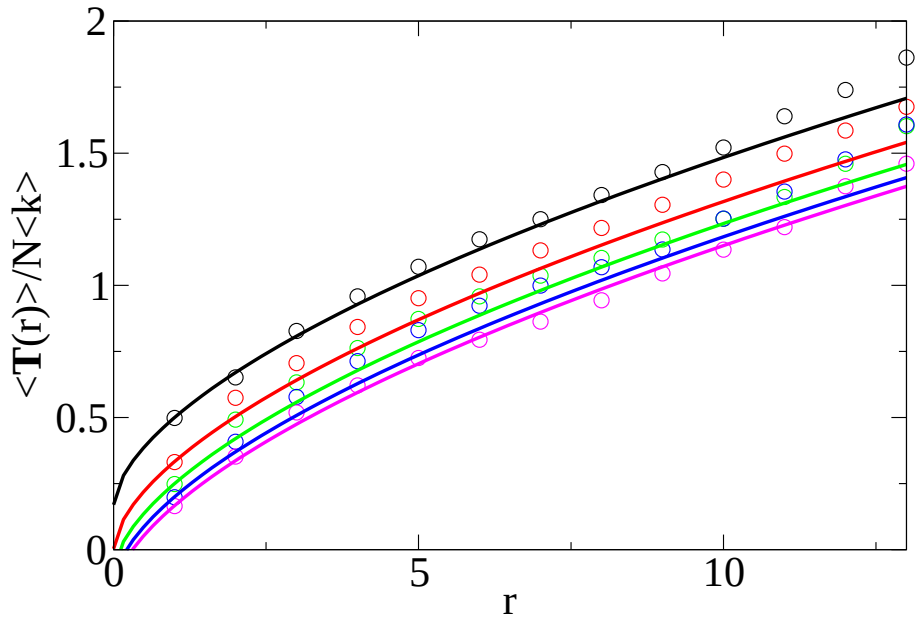

Figure 3.17: MFPT on random (u, v)-flowers for various target connectivities k. From up to down, black symbols stand for $k = 2$, red for $k = 3$, green for $k = 4$, blue for $k = 5$ and magenta for $k = 6$. Circles stand for simulations results, straight lines for $1/k + 1/\langle k\rangle(r^{d_w-d_f})$ ($\langle k\rangle = 3$) of equation (3.72).

r large, and scales as the inverse connectivity of the target. This behavior is in strong contrast with the case of compact exploration.

Discussion

Finally our central result can be summarized as follows, where the case of marginal exploration ($d_w = d_f$) has been obtained along the same line :

$$\frac{\mathbf{T}_{\mathbf{r}_T}(r)}{N\langle k\rangle} \sim \begin{cases} \dfrac{1}{k} + \dfrac{1}{\langle k\rangle}\left(r^{d_w-d_f} - 1\right) & \text{if } d_w > d_f \\ \dfrac{1}{k} + B\ln(r) & \text{if } d_w = d_f \\ \dfrac{A}{k} - Br^{d_w-d_f} & \text{if } d_w < d_f \end{cases} . \tag{3.77}$$

Figure 3.20 shows the MFPT regular $(2, 2)$-flowers, as a function of the source-target distance r, for a various target connectivity k. For thise networks, $d_w = d_f$. Black symbols stand for $k = 2$, red for $k = 4$ and blue for $k = 8$. Circles and crosses stand for simulation results, for two sizes of the network (respectively generations 4 and 5), straight lines stand for the formula $1/k + B\ln(r)$ of equation (3.77), with $B = 0.24$. The scaling works really well when $d_w = d_f$.

Expression (3.77) is very general and shows the respective impact of distance and connectivity on the MFPT. In particular the MFPT is *fully explicitly determined* in the compact case. The

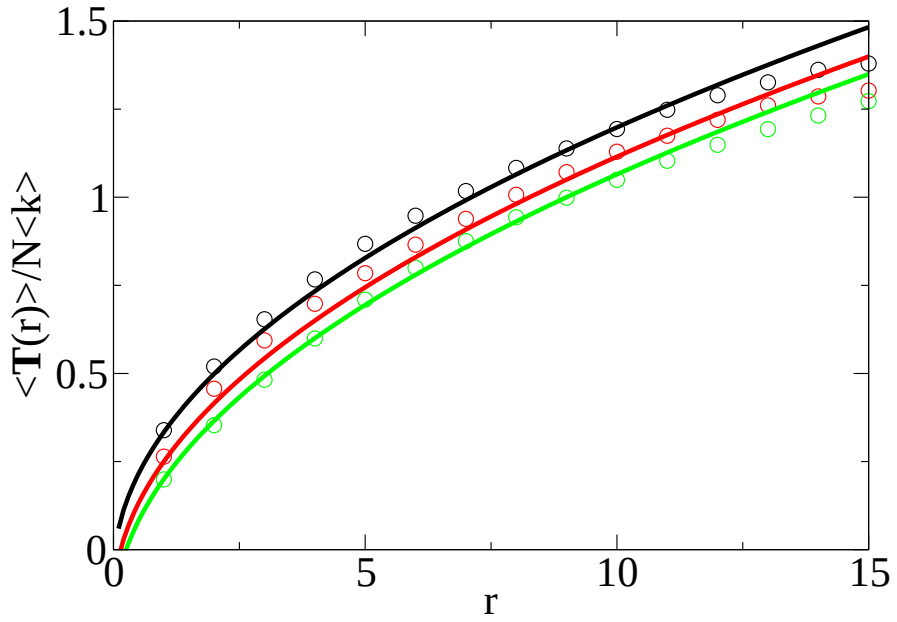

Figure 3.18: MFPT on compact Kozma networks, for various target connectivities k. Symbols stand for simulation result: black is $k = 3$, red $k = 4$ and green $k = 5$. The expected scaling is in $r^{0.5}$: circles stand for simulation results, straight lines stand for $1/k + 1/\langle k\rangle(r^{0.5} - 1)$ ($\langle k\rangle = 2.5$) of equation (3.72).

positive constants A and B depend on the network in the case of non compact exploration. We comment that in both cases the target connectivity k plays an important role at short distances r. However for large source-target distances r, the k-dependence is damped out in the compact case, while it remains important in the non compact case. The r-dependence is found to be important in the compact case and largely irrelevant in the non compact case in agreement with previous results [66]. The question raised in introduction can therefore be answered as follows : in the non compact case connected targets are found the fastest almost independently of their distance, while in the compact case close targets are found the fastest almost independently of their connectivity.

We can conclude that for self-similar networks with short range degree correlations, the main criterion that governs the behavior of $\mathbf{T}$ is the type (compact or non compact) of the random walk. In particular the existence of loops is irrelevant. Further comments are in order.

(i) As stressed above, equation (3.77) is derived in the large r regime. Its applicability to the small r regime relies on the assumption that the scaling form of the propagator (3.59) holds for all values of r, which is not always satisfied for real networks. In particular when degree correlations exist the relation $B = 1/\langle k\rangle$ obtained in the compact case gives only a rough estimate, and the result of equation (3.77) is valid only for r larger than the correlation length.

(ii) Our results can be extended to the case of non self-similar networks, still under the as-

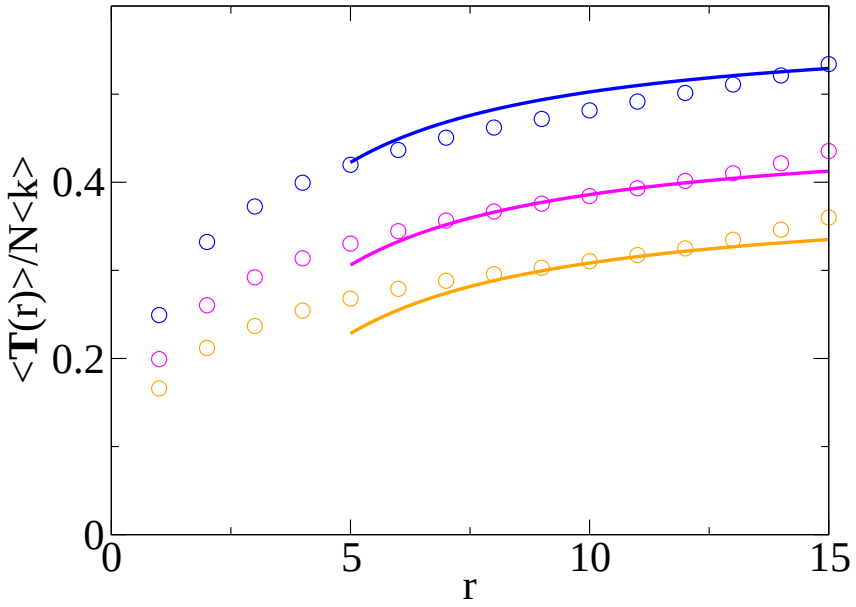

Figure 3.19: MFPT on supercritical 3D percolation networks ($p = 0.8$) for various target connectivities k. From up to down, blue symbol stand for $k = 4$, magenta for $k = 5$ and orange for $k = 6$. Circles stand for simulations results, straight lines for a fit by $A/k + Br^{d_w - d_f}$, with $A \simeq 2.33$ and $B \simeq 0.8$.

sumption that degree correlations are negligible. Following the method developed above, one can infer that

$$\mathbf{T}_{\mathbf{r}_T}(r) \sim N\langle k\rangle \left(A/k + g(r)\right) \tag{3.78}$$

where g does not depend on k and satisfies $g(r \to \infty) = C$ in the transient case, and $g(r \to \infty) = \infty$ in the recurrent case. The relative impact of connectivity and distance is therefore qualitatively the same as in the case of self-similar networks discussed above.

Figure 3.21 shows the MFPT on a non compact Kozma 1D network[125] ($\alpha = 1.0$) of size $X = 400$, as a function of the source-target distance r, for a various target connectivity k. Those networks are not strictly self-similar, we thus apply the result of equation (3.78). From up to down, green circles stand for $k = 6$, blue for $k = 7$ and magenta for $k = 8$. The insight shows a translation along the y axis of A/k with $A = 2.04$ according to equation (3.78). As predicted, this quantity does not depend on k.

(iii) Incidentally, our results straightforwardly yield the k dependence of the GMFPT. We find in the large N limit :

$$\text{GMFPT} \sim \begin{cases} CN^{d_w/d_f} & \text{if } d_w > d_f \\ CN\ln N & \text{if } d_w = d_f \\ CN/k & \text{if } d_w < d_f \end{cases} \tag{3.79}$$

which complements previous results obtained in the first section of this chapter. This expression, along with equation (3.75), yields as a by-product the large N asymptotics of

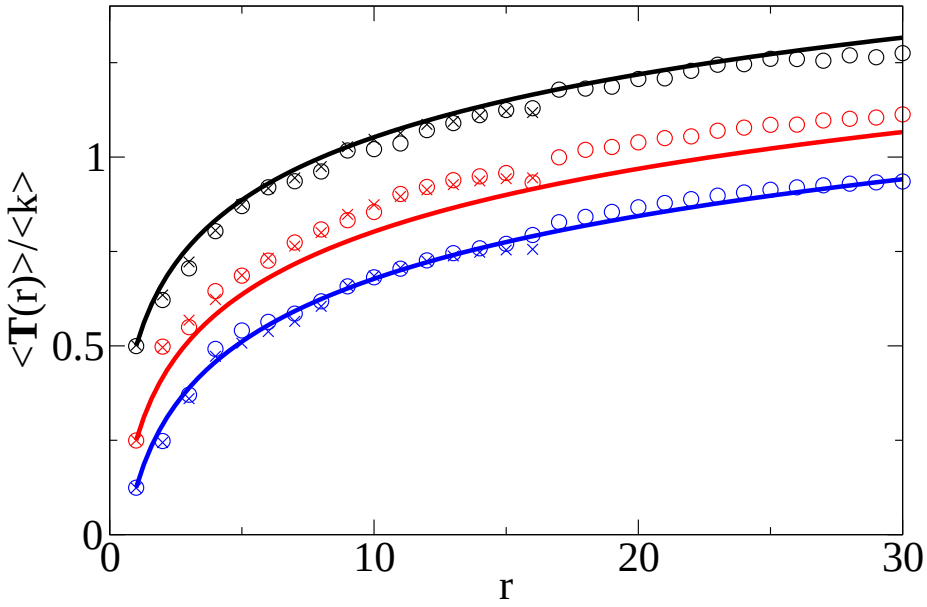

Figure 3.20: MFPT on (u, v)-flowers for various target connectivities k. Black symbols stand for $k = 2$, red for $k = 4$ and blue for $k = 8$. Circles and crosses stand for simulation results, for two sizes of the network (respectively generations 4 and 5), straight lines stand for the formula $1/k + B\ln(r)$ of equation (3.77), with $B = 0.24$.

$H(\mathbf{r}_T|\mathbf{r}_T)$:

$$H(\mathbf{r}_T|\mathbf{r}_T) \sim \begin{cases} CkN^{d_w/d_f-1} & \text{if } d_w > d_f \\ Ck\ln N & \text{if } d_w = d_f \\ C & \text{if } d_w < d_f \end{cases} . \tag{3.80}$$

This k-dependence of $H(\mathbf{r}_T|\mathbf{r}_T)$ is checked numerically in Figure 3.22 and directly validates the k–dependence of the GMFPT.

Figure 3.22 shows the pseudo-Green function $H(\mathbf{r}_T|\mathbf{r}_T)$ average over all target $\mathbf{r}_T$ of a given connectivity for several networks. The main plot shows the case of critical 3D percolation networks ($p = 0.2488$, compact case). The circle stand for numerical simulation, for a 10^3 network (black circles), 15^3 (red) and 20^3 (green). A fit in CkN^{d_w/d_f-1} (straight line) is performed, with the same fitting constant C for the three sizes. The inset shows the case of supercritical 3D percolation networks ($p = 0.8$, non-compact case). Three network sizes are plotted, 10^3 (black), 15^3 (red) and 20^3 (green). As predicted in equation (3.80), $H(\mathbf{r}_T|\mathbf{r}_T)$ does not depend either on k or on N. Figure 3.22 is in perfect agreement with equation (3.80) for both a compact and a non-compact examples.

Conclusion

To conclude, we have proposed in this section a general theoretical framework which elucidates the connectivity and source-target distance dependence of the MFPT for random walks on networks.

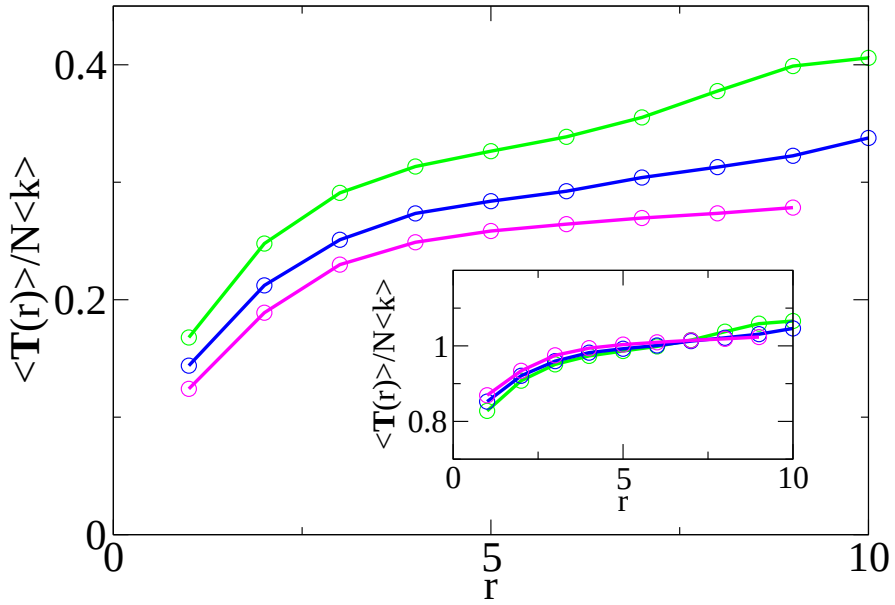

Figure 3.21: MFPT on non compact Kozma 1D network[125] ($\alpha = 1.0$) of size $X = 400$, for various target connectivities k. From up to down, green circles stand for $k = 6$, blue for $k = 7$ and magenta for $k = 8$. (Inset) translation along the y axis of A/k with $A = 2.04$ according to equation (3.78).

This approach leads to explicit solutions for self-similar networks and highlights two strongly different behaviors depending on the type – compact or non compact – of the random walk. In the case of non compact exploration, the MFPT is found to scale as the inverse connectivity of the target, and to be widely independent of the source-target distance. On the contrary, in the compact case the MFPT is controlled by the source-target distance, and we find that unexpectedly the target connectivity is *irrelevant* for remote targets. The question raised in introduction can therefore be answered as follows : in the non compact case connected targets are found the fastest almost independently of their position, while in the compact case close targets are found the fastest almost independently of their connectivity. Last, we stress that following [68], this explicit determination of MFPTs can be straightforwardly generalized to obtain other relevant first-passage observables, such as splitting probabilities or occupation times.

For our search strategy problem, the GMFPT can be used to quantify the search time. We have very different behaviors depending on the compact or non-compact character of the random walk: we expect the GMFPT to be independent of target connectivity for compact networks, and proportional to $1/k$ for non-compact one. We thus have a finer scaling with N and k than the bounds found in the previous sub-section.

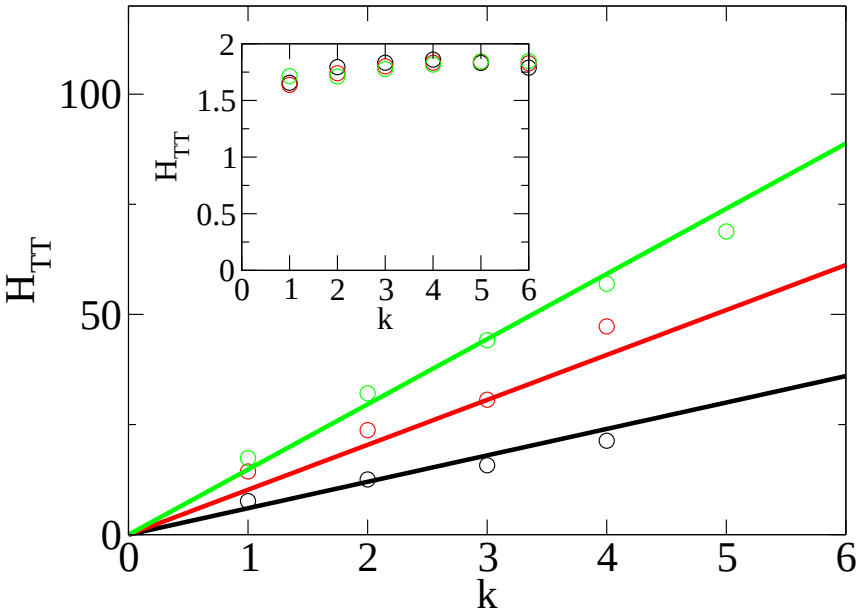

Figure 3.22: Numerical computation of $H(\mathbf{r}_T|\mathbf{r}_T)$ for critical 3D percolation networks ($p = 0.2488$, compact case). The circle stand for numerical simulation, for a 10^3 network (black circles), 15^3 (red) and 20^3 (green). A fit in CkN^{d_w/d_f-1} (straight line) is performed, with the same fitting constant C for the three sizes. (Inset) supercritical 3D percolation networks ($p = 0.8$, non-compact case). Three network sizes are plotted, 10^3 (black), 15^3 (red) and 20^3 (green).

3.3.3 Network perturbation

At last, we will investigate the impact on first passage properties of a topological change in the network.

This topological change can either be a protective measure to limit the propagation of a dangerous walker, or an aggressive attack to weaken a network. Recently, an increasing interest in modeling failures or attacks in networks has developed [144], for instance in the context of epidemic spreading [155], virus attack on networks [76], or electrical blackout [44, 53]. Most of these studies deal with static properties of networks after a given perturbation (typically the removal of nodes or links), with the notable exception of [195]. In particular it has been showed that scale free networks are very resilient to random perturbations, while the targeted removal of a hub can have dramatic consequences. Very recently, it has also been put forward that in the case of interdependent networks a broader degree distribution increases the vulnerability to random failure [44].

In this context, quantifying the impact of targeted perturbations of a network on the search efficiency and more generally on transport properties is an important and yet unexplored problem. We will here provide a general framework that quantifies the response of the MFPT to a target node to a local perturbation of the network, both in the context of attacks (damaged link) or strategies of transport enhancement (added link). This approach enables to determine explicitly

the dependence of this response on geometric parameters (such as the network size and the position of the perturbation) and on the intensity of the perturbation. In particular, it reveals that the relative variation of the MFPT is independent of the network size N in the large N limit, and remains significant even for very large networks ; additionally, in the case of non compact exploration of the network, a targeted perturbation keeps a substantial impact on transport properties for any localization of the damaged link.

To model those effects, we will consider a discrete time random walker on a network of N nodes, and assume that the transition probabilities ω_{uw} from u to w of the walker are such that a stationary distribution P_{stat} exists. In addition, we denote $P(x, n|y)$ the propagator, as previously defined as the probability to be at site x after n steps for a random walk starting from site y. In this sub-section, we will be interested in the influence of targeted perturbations of the network on the MFPT $\langle \mathbf{T}_{TS} \rangle$ from a starting site $\mathbf{r}_S$ to a target site $\mathbf{r}_T$.

Response of the MFPT to a damaged link

We first introduce the general method to study the influence of a single link perturbation on the MFPT, and take the example of the weakening or removal of a link. More precisely, we assume that the transition probability ω_{uw} from u to w is changed by $\delta\omega_{uw} < 0$. Without loss of generality, we here assume that this perturbation is compensated on the probability ω_{uu} to stay at u during the elementary step, *i.e.* $\delta\omega_{uu} = -\delta\omega_{uw}$, and that all other transition rates are unchanged. Note that the important particular case of a broken link is then given by taking $\delta\omega_{uw} = -\omega_{uw}$. This means that when the random walker wants to use the path $u \to w$, with a probability δ_{uw}, he will remain in u instead of reaching w. Figure 3.23 shows the effect of a damage link between u and w. Those two points do not play a symmetric role in the network perturbation. In the sequel, all quantities denoted with a prime correspond to the perturbed situation.

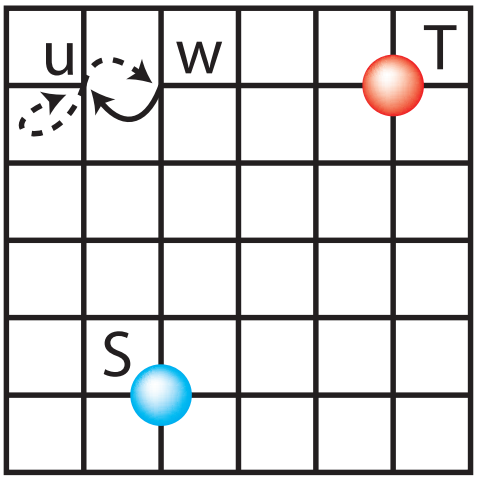

Figure 3.23: Scheme of the damaged link on a discrete network.

The MFPT in the perturbed situation can be calculated by first noting that the perturbation affects only the trajectories that pass through u before reaching T, so that, for any starting point x :

$$\langle \mathbf{T}'_{Tx} \rangle - \langle \mathbf{T}_{Tx} \rangle = P_{ux}^{T} \left(\langle \mathbf{T}'_{Tu} \rangle - \langle \mathbf{T}_{Tu} \rangle \right), \tag{3.81}$$

where P_{ux}^T is the splitting probability to reach u before T starting from x. Writing next the Kolmogorov backward equation for the MFPT [168]:

$$\langle \mathbf{T}'_{Tu} \rangle = 1 + \sum_v \omega'_{uv} \langle \mathbf{T}'_{Tv} \rangle \tag{3.82}$$

both in the perturbed and unperturbed situations, we obtain

$$\langle \mathbf{T}'_{Tu} \rangle - \langle \mathbf{T}_{Tu} \rangle = \sum_v \omega_{uv} \left(\langle \mathbf{T}'_{Tv} \rangle - \langle \mathbf{T}_{Tv} \rangle \right) + \delta\omega_{uw} (\langle \mathbf{T}'_{Tw} \rangle - \langle \mathbf{T}'_{Tu} \rangle). \tag{3.83}$$

Using equations (3.81) and (3.83), the variation of the MFPT starting from site u is found to be given by :

$$\langle \mathbf{T}'_{Tu} \rangle - \langle \mathbf{T}_{Tu} \rangle = \frac{\delta\omega_{uw} (\langle \mathbf{T}_{Tw} \rangle - \langle \mathbf{T}_{Tu} \rangle)}{1 - \bar{P}_{uu}^T + \delta\omega_{uw}(1 - P_{uw}^T)}, \tag{3.84}$$

where we have introduced

$$\bar{P}_{uu}^T \equiv \sum_v \omega_{uv} P_{uv}^T, \tag{3.85}$$

defined as the probability to come back to u before reaching T. Using equation (3.81), we finally obtain the relative MFPT variation δ_{TS}, defined as:

$$\delta_{TS} \equiv (\langle \mathbf{T}'_{TS} \rangle - \langle \mathbf{T}_{TS} \rangle) / \langle \mathbf{T}_{TS} \rangle \tag{3.86}$$

for any starting site S:

$$\delta_{TS} = \frac{P_{uS}^T}{\langle \mathbf{T}_{TS} \rangle} \frac{\delta\omega_{uw} (\langle \mathbf{T}_{Tw} \rangle - \langle \mathbf{T}_{Tu} \rangle)}{1 - \bar{P}_{uu}^T + \delta\omega_{uw}(1 - P_{uw}^T)}. \tag{3.87}$$

Using [28, 64, 65, 66], δ_{TS} can be expressed as a function of the perturbation $\delta\omega_{uw}$, the unperturbed stationary distribution P_{stat}, and the pseudo-Green function, H_{xy} of the unperturbed problem, defined as:

$$H_{xy} \equiv \sum_{n=0}^{\infty} (P(x, n|y) - P_{\text{stat}}(x)) \tag{3.88}$$

We obtain

$$\delta_{TS} = \frac{H_{Tu} - H_{Tw}}{H_{TT} - H_{TS}} \frac{\delta\omega_{uw} \left[P_{\text{stat}}(T)(H_{uS} - H_{uT}) + P_{\text{stat}}(u)(H_{TT} - H_{TS}) \right]}{P_{\text{stat}}(T) + \delta\omega_{uw} \left[P_{\text{stat}}(T)(H_{uu} - H_{uw}) + P_{\text{stat}}(u)(H_{Tw} - H_{Tu}) \right]}. \tag{3.89}$$

This central result has several important implications.

First, we stress that in the particular case of a regular d–dimensional hypercubic parallelepipedic network with constant probability transitions between nearest neighbors the pseudo Green functions are known exactly [14] (see chapter 1). Using next that the stationary probability is in this case uniform ($P_{\text{stat}}(x) = 1/N$ for any node x), equation (3.89) provides an exact and fully explicit result for the effect of an arbitrary modification of a given link on the MFPT.

Second, in the case of more general networks, but possessing scale-invariant properties, the dependence on the geometrical parameters can still be determined, by taking the large network size limit. This is for example the case of self-similar networks like (u, v)–flowers [202, 204] defined in chapter 1. More precisely, the large network size limit can be conveniently discussed when the random walk is a scale-invariant process, *i.e.* when the infinite volume propagator satisfies the scaling

$$P(x, n|y) \propto n^{-d_f/d_w} \Pi(|x - y|/n^{1/d_w}) \tag{3.90}$$

where d_w and d_f denote as usual the walk dimension and the fractal dimension of the network, and where $|x-y|$ denote the distance between nodes x and y. Indeed, in this case all the differences $H_{xy} - H_{xz}$ that enters equation (3.87) can be rewritten in terms of differences $H_{xx} - H_{xy}$ which satisfy in the large volume limit [28, 66] (see chapter 1 for details)

$$H_{xx} - H_{xy} \sim \begin{cases} A + B|x-y|^{d_w - d_f} & \text{if } d_w < d_f \\ A + B \ln |x-y| & \text{if } d_w = d_f \\ B|x-y|^{d_w - d_f} & \text{if } d_w > d_f \end{cases} \tag{3.91}$$

where A and B are numerical constants depending only on the infinite volume propagator, and more precisely on the scaling function Π.

Figure 3.24 shows the relative variation of the MFPT as a function of the distance $|S-T|$ between source and target (here the relative positions of T, u and w are fixed and averages over triplets are taken). Symbols stand for numerical simulation of 3D regular lattices (circles, $d_w - d_f = -1$), 2D critical percolation clusters (triangles, $d_w - d_f \simeq 0.98$) and $(2,2)$–flowers (crosses, $d_w = d_f$). For each network, several size are plotted in different colors. The plain lines are the result of equation (3.89), where the pseudo-Green functions are either exact (3D regular lattices) or approximated by the scaling of equation (3.91). The very good fitting shows that the approximation of equation (3.91) is good enough to reproduce accurately the MFPT variations using equation (3.89).

Equations (3.89) and (3.91) have two consequences.

(i) The independence of N of equation (3.91) readily gives that the relative variation δ_{TS} is *independent of* N in the large volume limit. This quite unexpected effect, illustrated by the data collapse for different volumes in Figure 3.24 on various networks, implies that the effect of a targeted perturbation is not diluted but remains finite even for extremely large networks. Actually this universal asymptotic independence on N of the relative variation δ_{TS} for fixed starting node S strongly differs from the relative variation δ_T of the GMFPT, defined as previously by

$$\text{GMFPT} \equiv \sum_S P_{\text{stat}}(S) \langle \mathbf{T}_{TS} \rangle. \tag{3.92}$$

More precisely, assuming next detailed balance, it can be shown using symmetry relations of H_{xy} that

$$\begin{aligned} \delta_T &\equiv \frac{\text{GMFPT}' - \text{GMFPT}}{\text{GMFPT}} \\ &= \frac{H_{Tu} - H_{Tw}}{H_{TT}} \frac{\delta\omega_{uw} P_{\text{stat}}(u)(H_{TT} - H_{Tu})}{P_{\text{stat}}(T) + \delta\omega_{uw}[P_{\text{stat}}(T)(H_{uu} - H_{uw}) + P_{\text{stat}}(u)(H_{Tw} - H_{Tu})]} \end{aligned} \tag{3.93}$$

In the large volume limit, this exact expression leads to

$$\delta_T \sim \begin{cases} C & \text{if } d_w < d_f \\ C/\ln N & \text{if } d_w = d_f \\ C/N^{d_w/d_f - 1} & \text{if } d_w > d_f \end{cases} \tag{3.94}$$

where we have derived the asymptotic expression of H_{TT} using [151, 209]. In other words, as for the relative variation δ_T, the N independence is recovered only in the case $d_f > d_w$ of so-called non compact exploration, while a strong dependence on N is found in the opposite case of compact exploration ($d_w \geq d_f$). Figure 3.25 shows the relative variation of the GMFPT, δ_T for different network sizes N for 1D (black crosses), 2D (red circles)

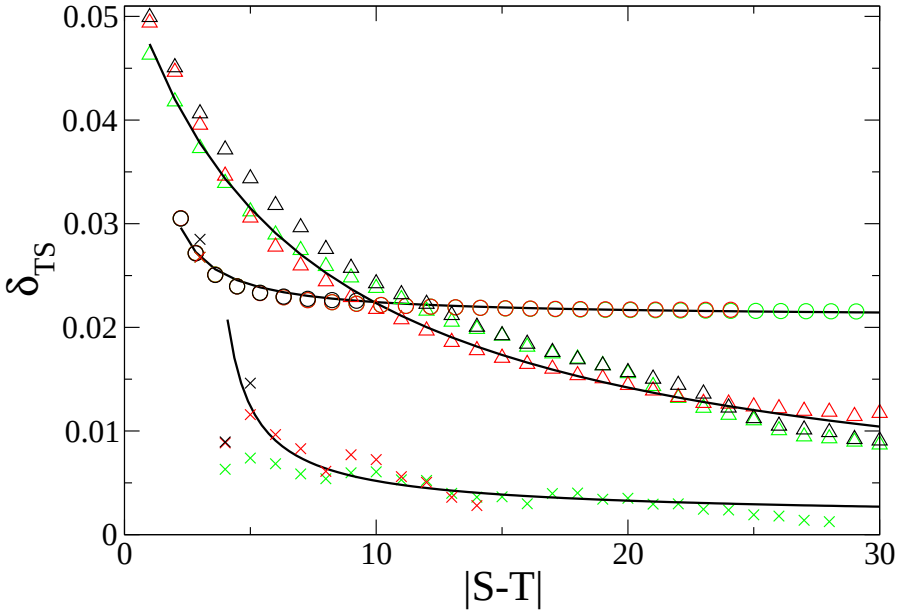

Figure 3.24: Relative variation of the MFPT δ_{TS} as a function of the distance $|S-T|$ for several networks. Symbols stand for numerical simulation of 3D regular lattices (circles, $d_w - d_f = -1$), 2D critical percolation clusters (triangles, $d_w - d_f \simeq 0.98$) and $(2,2)$–flowers (crosses, $d_w = d_f$). For each network, several size are plotted in different colors. The plain lines are the result of equation (3.89), where the pseudo-Green functions are either exact (3D regular lattices) or approximated by the scaling of equation (3.91).

and 3D (blue triangles) cubic lattices and a 2D critical percolation cluster (green circle - inset). The relative positions of T, u and w are fixed for all networks of a given kind. The circles stand for the simulated δ_T, the black lines for the theoretical prediction of equation (3.94), where C is a fitting parameter, different for all networks. The N dependence of the GMFPT strongly changes between compact and non-compact networks.

(ii) The asymptotic form (3.91) used in equation (3.89) also provides the explicit dependence of δ_{TS} on the relative distances between the nodes S, T, u, w. Such dependence has been checked numerically (see Figure 3.24) for various networks such as regular euclidian lattices ($d_w = 2$), 2D critical percolation clusters ($d_w = 2.88$ and $d_f = 91/48$), and (u,v)–flowers (see chapter 1 for definition). Figure 3.24 reveals a very different behavior in the case of compact exploration (illustrated by critical percolation clusters) for which δ_{TS} vanishes at larges distances, and in the case of non-compact exploration (illustrated by 3D regular lattices), for which δ_{TS} remains finite even for very large distances. Noteworthily this shows that in the non compact case a targeted perturbation keeps a substantial impact on transport properties for *any localization* of the damaged link.

Third, in the general case of a random walk on an arbitrary network, where the pseudo-

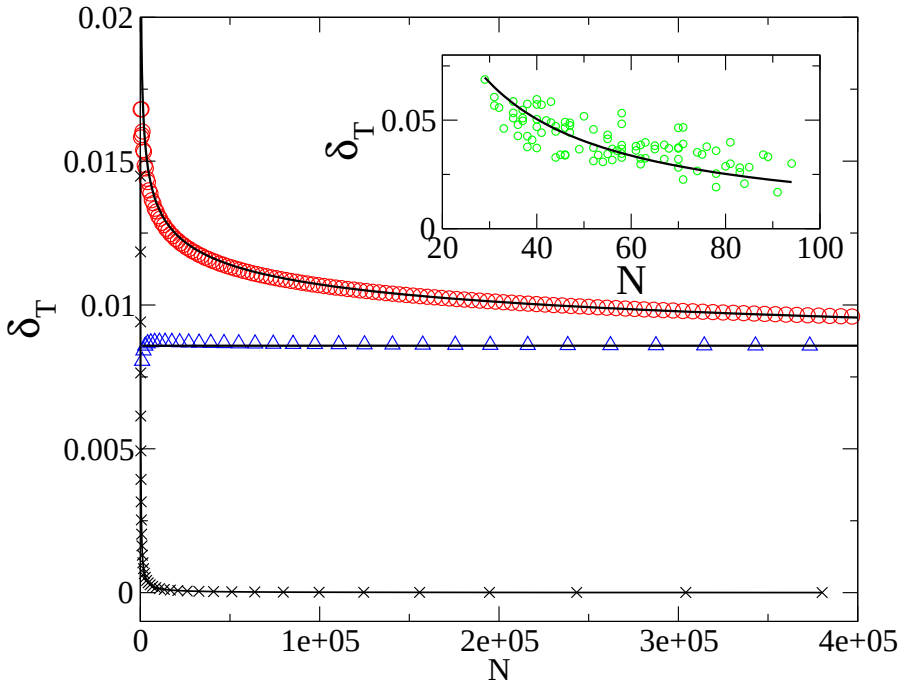

Figure 3.25: Relative variation of the GMFPT as a function of the network size N, for 1D (black crosses), 2D (red circles) and 3D (blue triangles) Euclidian lattices and a 2D critical percolation cluster (green circle - inset). The relative positions of T, u and w are fixed for all networks of a given kind. The circles stand for the simulated δ_T, the black lines for the theoretical prediction of equation (3.94), where C is a fitting parameter, different for all networks.

Green functions can be difficult to evaluate, equation (3.89) still gives explicitly the functional dependence of δ_{TS} on the perturbation $\delta\omega_{uw}$, which takes the form

$$\delta_{TS} = \frac{D\delta\omega_{uw}}{E + \delta\omega_{uw}} \tag{3.95}$$

where D and E do not depend on $\delta\omega_{uw}$ (note that E does not depend on S either). This general form has been validated by numerical simulations in Figure 3.26. This figure shows the relative variation of the MFPT as a function of the $u \to w$ link perturbation, for given T, u and w sites, and two different S for each network (one color for each S): 2D critical percolation clusters (circles) and $(3,3)$-flower of generation 3 (triangles). For the flower networks, the perturbed link leads to T. Numerical simulations are fitted with equation (3.95).

Additionally, provided that the differences $H_{xx} - H_{xy}$ involved in equation (3.89) have a finite limit in the large volume regime, D and E turn out to be independent of N. This shows that the independence of δ_{TS} on N still holds in this case, and makes this property very robust.

Finally, it should be noted that the relative variation δ_{TS} remains rather weak for an arbitrary perturbed link. We however stress that such local attack of a network is not affected by dilution effects and remains finite even in the large volume limit; additionally, in the non compact case, it is also widely independent of the localization of the perturbation and non vanishing even for

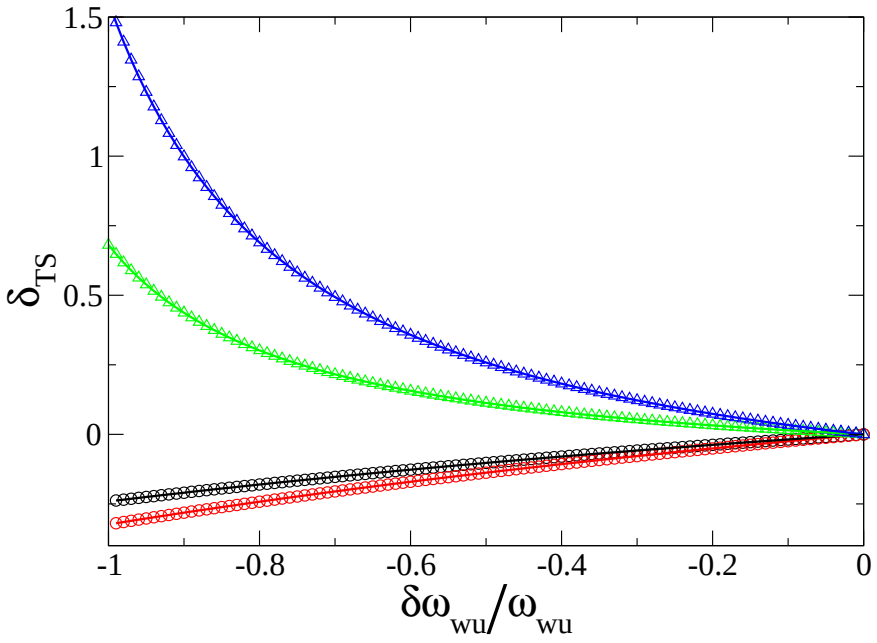

Figure 3.26: Relative variation of the MFPT δ_{TS} as a function of the $u \to w$ link perturbation, for given T, u and w sites. Two different starting sites S are chosen for each network (one color for each S): circles stand for 2D critical percolation clusters (red and black) and triangles for $(3,3)$–flowers of generation 3 (green and blue).

a very remote damaged link. Furthermore, the effect of a local perturbation can become much stronger if targeted to a link directly leading to the target, as can be expected intuitively (see Figure 3.26).

Response of the MFPT to an added link

Importantly, the above formalism can be extended to tackle the reciprocal problem of enhancing instead of damaging the transport abilities of a network. As a first step in this direction, we quantify the effect of adding a new link between two nodes. The definition of the perturbation has to be slightly modified in the case of an added link, since the case studied above is ill defined for $\delta\omega_{uw} > 0$ and $R_{uu} = 0$. We assume now that there is initially no link between u and w ($\omega_{uw} = 0$), and consider a perturbation $\delta\omega_{uw} > 0$. In turn, we set for all neighbors v of u that $\delta\omega_{uv} = -\delta\omega_{uw}/k(u)$, where $k(u)$ is the initial connectivity of u (note that $\delta\omega_{uw}$ is assumed to be small enough so that all transition probabilities remain positive). In this case, the equivalent of equation (3.87) can be obtained along the same line and reads:

$$\delta_{TS} = \frac{P^T_{uS}}{\langle \mathbf{T}_{TS} \rangle} \frac{\delta\omega_{uw} \left(\langle \mathbf{T}_{Tw} \rangle - \langle \mathbf{T}_{Tu} \rangle + 1 \right)}{1 - \bar{P}^T_{uu} + \delta\omega_{uw} \left(\bar{P}^T_{uu} - P^T_{uw} \right)}, \tag{3.96}$$

which, as previously, can be expressed only in terms of $\delta\omega_{uw}$ and pseudo-Green functions.

Figure 3.27 shows the relative variation of the MFPT in response to the addition of a new link, in the directed (red squares and black circles) and bidirectional case (green triangles) for

3D Euclidian lattices. Simulations (symbols) are plotted against the theoretical prediction of equation (3.96) (plain lines). Black circles: S, T, u and w are distinct, and the new link starts from a target neighbor. Red squares: the new link is between S and T. Green triangles : addition of a link between two 3D regular lattices (with $\delta\omega_{uw} = \delta\omega_{wu}$), S being in the first lattice, and T in the second. Here we define $\delta_{\mathrm{ST}} = (\mathrm{MFPT} - \min(\mathrm{MFPT}))/\min(\mathrm{MFPT})$.

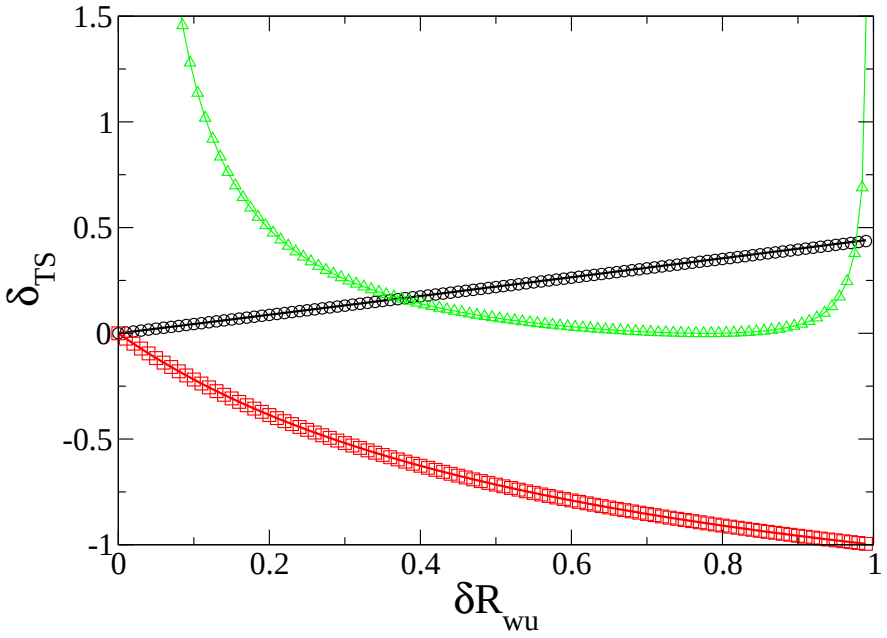

Figure 3.27: Relative variation of the MFPT in response to the addition of a new link in the directed (red squares and black circles) and bidirectional case (green triangles) for 3D Euclidian lattices. Simulations (symbols) are plotted against the theoretical prediction of equation (3.96) (plain lines). Black circles: S, T, u and w are distinct, and the new link starts from a target neighbor. Red squares: the new link is between S and T. Green triangles : addition of a link between two 3D regular lattices (with $\delta\omega_{uw} = \delta\omega_{wu}$), S being in the first lattice, and T in the second.

Note that the MFPT can be decreased very significantly if the new link points to the target (red line of Figure 3.27). Last, we stress that the case of a single added bidirectional link connecting two nodes u and w of initially distinct networks $\mathcal{N}_1$, $\mathcal{N}_2$ can be obtained using the same method, and yields an explicit results for the MFPT which is displayed in Figure 3.27 (green line). This constitutes a first step in designing interdependent networks as introduced in [44].

Conclusion

To conclude, we have presented a general framework that quantifies the response of the MFPT to a target node to a local perturbation of the network, both in the context of attacks (damaged link) or strategies of transport enhancement (added link). This approach enables to determine explicitly the dependence of this response on geometric parameters (such as the network size and

the position of the perturbation) and on the intensity of the perturbation. It reveals that the relative variation of the MFPT is independent of the network size, and remains significant even in the large size limit. Additionally, in the non compact case a targeted perturbation keeps a substantial impact on transport properties for any localization of the damaged link.

Conclusion

This third chapter aimed to understand how a random search process could be optimized in a confined media, namely how the search time could be minimized. If the target localization is unknown, the adequate observable is the GMFPT. If we have more insight on the relative position of the starting and the target sites, the MFPT becomes the relevant observable.

During the first section, we showed that for a walker evolving on a periodic Euclidian lattice, persistent random walks could be optimized. We focused on the GMFPT (MFPT have been studied in chapter 2), and found that the optimized persistent random walk performed better than any Lévy walk in the sense that the optimized GMFPT was lower.

The second section investigated the case of a moving target. We wanted to assess the impact of a moving target on first-passage observables. Since the problem was cumbersome, we focused on 1D subdomain, and found analytical and approximate expressions of the first encounter time, either its mean and its full density. The results obtained are non trivial, and in particular the first encounter density does not behaves as expected from [151]. The usual approximation on pseudo-Green functions for the first-passage time does not seem to apply well on extended targets (in the 2D problem).

The last section assessed the impact of several network topological features on the first-passage time. General boundaries of the GMFPT have been summarized in the first sub-section, showing that the compacity of the exploration was a key parameter to estimate the GMFPT. The tree structures also have a particular behavior. The second sub-section extended the result previously obtained on MFPT to take into account the target connectivity. The scaling form proposed for the propagator seems to apply very well on various scale invariant networks. At last, the influence of a network perturbation have been studied in the last sub-section, showing once again the differences between compact and non compact random walks. For compact random walks, a distant perturbation does not affect the MFPT in the large volume limit, while it affects a non compact random walks independently of the network volume.

Conclusion

Random walks and their first-passage properties govern the microscopical world, and in particular biochemical reactions. Understanding how the first-passage properties are influenced by the surrounding environment or by the diffusion mechanism can either help to probe the microscopical environment with single particle trajectories, or to control the behavior of a random walker by designing correctly the environment.

Our first approach was to use the random walk framework to extract information from single particle trajectories. Modern tracking setups allow to have access directly to trajectories, and not only to macroscopical averaged data such as diffusion coefficients. Experiments have observed repeatedly anomalous diffusion, in particular in biology, but so far the information extracted from those data was only the anomalous coefficient α. We have shown here that the first-passage properties of single particle trajectories were a powerful tool to investigate the origin of anomalous diffusion. We took into account the experimental constraints to propose several tests that can help discriminate between the three prominent model of anomalous diffusion, namely diffusion on fractals, CTRW and fBm. Those tests could be widely applied to single particle trajectories analysis, and enhance the accuracy of the extracted anomalous exponent α and anomalous diffusion coefficient K_α with the same data set. They could also help to identify some key features of experimental behaviors, like space correlations or environment fractality.

After applying our method to experimental data, we analyzed three models that were somehow finite versions of the three anomalous diffusion models studied, namely a model with only two scales instead of an infinity (cage networks), a model with a short-range correlation (persistent random walks), and a model with finite waiting time (hard-core crowding). Persistent random walks, cage networks and hard-core crowding could all lead to sub-diffusion at short times, and a Brownian behavior at longer ones. For cage networks, we have pointed out some discrepancies between the simple CTRW approach and the simulations, and explained them. We also point out that we obtained a very good approximation of the MFPT for hard-core crowding on Euclidian network. This crowding effect is often cited as a potential source of anomalous diffusion in cells, and is seen as a CTRW with a cut-off time. We have shown that in addition to the (finite) CTRW behavior, one had to take into account memory effects (anti-persistence).

The second approach was to optimize, assuming a given diffusion mechanism, the search time, namely the time needed to a random walker to find a hidden target. We showed in the first section that optimized persistent random walks could outrun any Lévy walk in a periodic Euclidian lattice, when the target position is unknown. This result is really important for search processes, since this optimization is in many cases crucial for the random walker (animal looking for food for instance).

The case of the moving target has been investigated in the second section, and we showed that the resulting first encounter time density was not trivial. An extension of this result in higher dimension would be really interesting, since in chemistry for instance, bimolecular reactions take place in 2D or in 3D, and not only in 1D. An exact result will probably be hard to obtain, but if we use approximations of pseudo-Green functions, we may manage to obtain the mean first encounter time, at least in the large volume limit.

Search processes can also occur on discrete networks, like social networks, computer networks or electrical grids. On such networks, other parameters appear and can be adjusted in order to optimize the first-passage time. The identification of those key parameters was still an open question in the beginning of this work. We first summarized several distinct results to find general bounds on the GMFPT, valid for *any* Markovian process on a discrete network, no matter how complex the network is. These bounds show that the compacity of the random walk and the target connectivity are parameters that influence greatly the GMFPT. We also found out that tree networks has tighter bounds than networks containing loops.

Going deeper into details, we assessed the impact of target connectivity. Assuming a scaling form for the propagator that depends on the target connectivity, we showed that for scale invariant processes, the target connectivity influence is crucial for non-compact explorations, and much less important for compact explorations.

At last, we estimated the influence on search efficiency of a little variation of the network. We determined the impact on the MFPT of a damaged link, or in the reverse of the creation of a new link. The modification of the network could be a strategy to optimize lab-on-a-chip. Indeed, we could design a microscopical maze, where several reactive sites have to be visited by a reactant diffusing freely. If the maze is optimized, we can hope that almost surely, the reactant will visit all the reactive sites, in the good order, before exiting the maze. We could also design the maze to sort out several random walkers with different characteristics (size, interaction with the support...) in continuous time.

We have here extended somehow the available knowledge of first-passage properties, mainly on networks. We hope that those results will help to analyze more efficiently real trajectories of single random walkers, and to gain a better understanding of first-passage properties of random walks on discrete networks.

1 Moments computation and experimental setups

We here detail the regular moment and maximal excursion moments computation, for brownian and anomalous diffusion.

We start by the Brownian motion, and extend those results to anomalous diffusion afterward.

1.1 Brownian motion

Diffusion equation

In d dimensions and for spherical geometry, the diffusion equation is

$$\frac{\partial P(r,t)}{\partial t} = \frac{D}{2dr^{d-1}} \frac{\partial}{\partial r} \left(r^{d-1} \frac{\partial P(r,t)}{\partial r} \right). \tag{97}$$

where D is the diffusion coefficient.

Regular moments

To compute the regular moments, we impose the normalization and initial conditions

$$\int_0^\infty P(r,t) r^{d-1} dr = 1 \text{ and } P(r,0) = \frac{\delta_+(r)}{r^{d-1}}, \tag{98}$$

where $\delta_+(r)$ is the one-sided δ function. The boundary conditions are normal, i.e., chosen such that $P(r \to \infty, t) = 0$. After obtaining an expression for the propagator $P(r,t)$ we compute the nth moment

$$\langle r^n \rangle = \int_0^\infty r^n P(r,t) r^{d_f - 1} dr. \tag{99}$$

Alternatively, this can be obtained by integration of r^n times the diffusion equation.

Maximal excursion moments

To compute maximal excursion moments, we first compute the probability that at time t the distance from the origin traveled by the random walker is less than r_0: $r_{\max} \leq r_0$. To this end we consider an absorbing sphere at the radius $r = r_0$ and obtain the propagator in the domain $0 \leq r \leq r_0$ using the following boundary and initial conditions:

$$P(r_0, t) = 0 \text{ and } P(r, 0) = \frac{\delta_+(r)}{r^{d-1}}. \tag{100}$$

The role of the absorbing sphere is to remove the cumulative probability that the random walker actually crossed the distance r_0 before t. The sought maximal excursion probability then becomes

$$\Pr\left(r_{\max} \leq r_0, t\right) = \int_0^{r_0} P(r,t) r^{d-1} dr. \tag{101}$$

The nth maximal excursion moment is obtained by integration of $r_{\max}^n$ times the derivative of the cumulative distribution, i.e., the density,

$$\begin{aligned} \langle r_{\max}^n \rangle &= \int_0^\infty r_0^n \frac{\partial \left(\Pr\left(r_{\max} \leq r_0, t\right)\right)}{\partial r_0} dr_0 \\ &= n \int_0^\infty r_0^{n-1} \left(1 - \Pr\left(r_{\max} \leq r_0, t\right)\right) dr_0. \end{aligned} \tag{102}$$

One dimension

The diffusion equation is simply

$$\frac{\partial P(r,t)}{\partial t} = \frac{D}{2} \frac{\partial^2 P(r,t)}{\partial r^2}. \tag{103}$$

After normalization the solution is the well-known Gaussian propagator

$$P(r,t) = \frac{1}{\sqrt{\frac{D\pi t}{2}}} \exp\left(-\frac{r^2}{2Dt}\right). \tag{104}$$

We then obtain for the second and the fourth moment

$$\langle r^2 \rangle = Dt \text{ and } \langle r^4 \rangle = 3(Dt)^2. \tag{105}$$

For an absorbing sphere placed at $r = r_0$ the MME propagator after becomes

$$P(r,t) = \frac{2}{r_0} \sum_{n=0}^{\infty} \cos\left(\frac{(2n+1)\pi r}{2r_0}\right) \exp\left(-\frac{D(2n+1)^2 \pi^2 t}{8 r_0^2}\right). \tag{106}$$

Integrating from 0 to r_0 we find the cumulative distribution function

$$\Pr\left(r_{\max} \leq r_0, t\right) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\left(-\frac{D(2n+1)^2\pi^2 t}{8r_0^2}\right). \tag{107}$$

This expression simplifies in the Laplace domain,

$$\Pr\left(r_{\max} \leq r_0, s\right) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \frac{1}{s + \frac{D(2n+1)^2\pi^2}{8r_0^2}}. \tag{108}$$

This expression can be simplified recalling that

$$\operatorname{sech}(z) = \pi \sum_{n=0}^{\infty} \frac{(-1)^n (2n+1)}{\pi^2 (n+1/2)^2 + z^2}. \tag{109}$$

The cumulative distribution can therefore be rewritten in the form

$$\Pr\left(r_{\max} \leq r_0, s\right) = \frac{1}{s}\left[1 - \operatorname{sech}\left(r_0\sqrt{\frac{2s}{D}}\right)\right]. \tag{110}$$

From this expression the Laplace transforms of the second and fourth MME moments become

$$\begin{aligned}
\langle r_{\max}^2(s)\rangle &= \frac{2}{s}\int_0^{\infty} r_0 \operatorname{sech}\left(r_0\sqrt{\frac{2s}{D}}\right) dr_0 && (111)\\
&= \frac{D}{s^2}\int_0^{\infty} u\operatorname{sech}(u) du = \frac{2DC}{s^2} && (112)\\
\langle r_{\max}^4(s)\rangle &= \frac{4}{s}\int_0^{\infty} r_0^3 \operatorname{sech}\left(r_0\sqrt{\frac{2s}{D}}\right) dr_0 && (113)\\
&= \frac{D^2}{s^3}\int_0^{\infty} u^3 \operatorname{sech}(u) du && (114)\\
&= \frac{D^2}{s^3}\left(6\imath(\mathrm{Li}_4(-\imath) - \mathrm{Li}_4(\imath))\right). && (115)
\end{aligned}$$

Here $C \approx 0.916$ is Catalan's constant, and Li_4 the polylogarithm. Inverse Laplace transform leads to the following results:

$$\begin{aligned}
\langle r_{\max}^2(t)\rangle &= 2DCt \approx 1.83Dt && (116)\\
\langle r_{\max}^4(t)\rangle &= 3\imath(\mathrm{Li}_4(-\imath) - \mathrm{Li}_4(\imath))(Dt)^2 && (117)\\
&\approx 5.93(Dt)^2 && (118)
\end{aligned}$$

We compare the variance and the dispersion of regular and maximal excursion moments in table 1.

Two dimensions

We write the diffusion equation in Cartesian coordinates,

$$\frac{\partial P(x,y,t)}{\partial t} = \frac{D}{4}\left(\frac{\partial^2 P(x,y,t)}{\partial x^2} + \frac{\partial^2 P(x,y,t)}{\partial y^2}\right). \tag{119}$$

	$\sigma^2 = \langle X^4 \rangle - \langle X^2 \rangle^2$	$\gamma = \sqrt{\frac{\langle X^4 \rangle - \langle X^2 \rangle^2}{\langle X^2 \rangle^2}}$
RM	$2(Dt)^2$	1.41
ME	$2.58(Dt)^2$	0.876

Table 1: Regular and maximal excursion moments in 1D.

After normalization the solution is

$$P(r,t) = \frac{2}{Dt} \exp\left(-\frac{r^2}{Dt}\right). \tag{120}$$

The second and fourth moments follow,

$$\langle r^2 \rangle = Dt \text{ and } \langle r^4 \rangle = 2(Dt)^2. \tag{121}$$

To place an absorbing circle at $r = r_0$ we need to pass to spherical coordinates,

$$\frac{\partial P(r,t)}{\partial t} = \frac{D}{4r}\left(r\frac{\partial^2 P(r,t)}{\partial r^2} + \frac{\partial P(r,t)}{\partial r}\right). \tag{122}$$

Rescaling of time by D leads to the solution

$$P(r,t) = \sum_{n=0}^{\infty} A_n J_0\left(K_n r\right) \exp\left(-\frac{DK_n^2}{4}t\right), \tag{123}$$

where A_n and K_n are the free parameters. Here we chose the Bessel function J_0 to ensure strict positivity of the density in $r = 0$. The boundary condition $P(r_0, t) = 0$ then imposes $K_n = x_{0n}/r_0$ (x_{0n} is the nth zero of J_0). For the normalization we use the Fourier-Bessel expansion formalism

$$\begin{aligned} A_n &= \frac{\displaystyle\int_0^{r_0} \frac{\delta(r)}{r} J_0\left(K_n r\right) 2\pi r dr}{\displaystyle\int_0^{r_0} J_0\left(K_n r\right)^2 2\pi r dr} \\ &= \frac{1}{\pi r_0^2 J_1(x_{0n})^2}. \end{aligned} \tag{124}$$

Integrating $P(r,t)$ from 0 to r_0 we find

$$\Pr\left(r_{\max} \leq r_0, t\right) = \sum_{n=0}^{\infty} \frac{2}{x_{0n} J_1(x_{0n})} \exp\left(-\frac{Dx_{0n}^2 t}{4r_0^2}\right). \tag{125}$$

In the Laplace space,

$$\Pr\left(r_{\max} \leq r_0, s\right) = \sum_{n=0}^{\infty} \frac{2}{x_{0n} J_1(x_{0n})} \left(\frac{1}{s + \frac{Dx_{0n}^2}{4r_0^2}}\right). \tag{126}$$

By a non trivial contour integral (see reference [165] for a complete proof) we finally obtain

$$\Pr\left(r_{\max} \leq r_0, s\right) = \frac{1}{s}\left(1 - \frac{1}{I_0\left(r_0\sqrt{\frac{4s}{D}}\right)}\right). \tag{127}$$

	$\sigma^2 = \langle X^4\rangle - \langle X^2\rangle^2$	$\gamma = \sqrt{\frac{\langle X^4\rangle - \langle X^2\rangle^2}{\langle X^2\rangle^2}}$
MSD	$(Dt)^2$	1
MME	$1.13(Dt)^2$	0.694

Table 2: Regular and maximal excursion moments in 2D.

This expression cannot be simplified any further. We compute the second and the fourth maximal excursion moment numerically, in Laplace and real space,

$$\langle r^2_{\max}(s)\rangle = \frac{D}{2s^2}\int_0^\infty \frac{u}{I_0(u)}du \tag{128}$$

$$\Rightarrow \langle r^2_{\max}(t)\rangle = \frac{Dt}{2}\int_0^\infty \frac{u}{I_0(u)}du \tag{129}$$

$$\langle r^4_{\max}(s)\rangle = \frac{D^2}{4s^3}\int_0^\infty \frac{u^3}{I_0(u)}du \tag{130}$$

$$\Rightarrow \langle r^4_{\max}(t)\rangle = \frac{(Dt)^2}{8}\int_0^\infty \frac{u^3}{I_0(u)}du. \tag{131}$$

Table 2 summarizes our findings in two dimensions.

Three dimensions

The diffusion equation in cartesian coordinates reads

$$\frac{\partial P(x,y,z,t)}{\partial t} = \frac{D}{6}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right))P(x,y,z,t). \tag{132}$$

with the normalized solution

$$P(x,y,z,t) = \frac{1}{\left(\frac{2\pi Dt}{3}\right)^{3/2}}\exp\left(-\frac{r^2}{\frac{2Dt}{3}}\right), \tag{133}$$

with $r = \sqrt{x^2+y^2+z^2}$. We then obtain for the second and the fourth moments

$$\langle r^2\rangle = Dt \text{ and } \langle r^4\rangle = \frac{5}{3}(Dt)^2. \tag{134}$$

To calculate the solution with an absorbing sphere located at $r = r_0$ we turn to spherical coordinates,

$$\frac{\partial P(r,t)}{\partial t} = \frac{D}{6r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial P(r,t)}{\partial r^2}\right) = \frac{D}{r}\frac{\partial^2 rP(r,t)}{\partial r^2}. \tag{135}$$

The quantity $rP(r,t)$ satisfies the 1D diffusion equation. $rP(r,t)$ can be written in a Fourier expansion, but now the value at $r = 0$ is 0:

$$\frac{\partial P(r,t)}{\partial t} = \sum_{n=1}^{\infty}\frac{2n\pi}{r_0^2}\frac{\sin(n\pi r/r_0)}{4\pi r}\exp\left(-\frac{Dn^2\pi^2 t}{6r_0^2}\right). \tag{136}$$

	$\sigma^2 = \langle X^4 \rangle - \langle X^2 \rangle^2$	$\gamma = \sqrt{\frac{\langle X^4 \rangle - \langle X^2 \rangle^2}{\langle X^2 \rangle^2}}$
MSD	$2/3(Dt)^2$	0.816
MME	$0.712(Dt)^2$	0.602

Table 3: Regular and MME moment statistics in 3D.

Integration from 0 to r_0 and Laplace transformation leads to the expression

$$\Pr\left(r_{\max} \leq r_0, t\right) = \sum_{n=1}^{\infty} 2(-1)^{n+1} \exp\left(-\frac{Dn^2\pi^2 t}{6r_0^2}\right) \tag{137}$$

$$\Pr\left(r_{\max} \leq r_0, s\right) = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{s + \frac{Dn^2\pi^2}{6r_0^2}}. \tag{138}$$

This expression can be simplified by use of a contour integral, obtaining

$$\Pr\left(r_{\max} \leq r_0, s\right) = \frac{1}{s}\left(1 - \frac{\sqrt{\frac{6s}{D}} r_0}{\sinh\left(\sqrt{\frac{6s}{D}} r_0\right)}\right). \tag{139}$$

We deduce the Laplace transform of the moments, as well as their expression after inverse Laplace transform,

$$\langle r_{\max}^2(s) \rangle = \frac{D}{3s^2} \int_0^{\infty} \frac{u^2}{\sinh(u)} du \tag{140}$$

$$\Rightarrow \langle r_{\max}^2(t) \rangle = \frac{7\zeta(3)}{6} Dt \tag{141}$$

$$\langle r_{\max}^4(s) \rangle = \frac{D^2}{9s^3} \int_0^{\infty} \frac{u^4}{\sinh(u)} du \tag{142}$$

$$\Rightarrow \langle r_{\max}^4(t) \rangle = \frac{31\zeta(5)}{12} (Dt)^2. \tag{143}$$

Here ζ represents the Riemann zeta function. The results of the 3D case are summarized in table 3.

1.2 Experimental setups for data acquisition

We here detail how the experimental data set analyzed in chapter 2 have been obtained.

Qdots freely diffusing

The first set of data was obtained by Ralf Jungmann and Friedrich Simmel at the Technical University of Munich (Garching, Germany), from fluorescence video tracking of single quantum dots (Qdot 565 streptavidin conjugate, Invitrogen Corporation, USA) freely diffusing in glycerol (G7757, Sigma Aldrich, USA) at room temperature (25C). The quantum dot stock solution was diluted in glycerol to a final concentration of 1 nM prior to usage. Samples were imaged using microscopic chamber slides (μ-slide 8, ibidi GmbH, Germany), which were incubated with 10 mg/ml

BSA solution (A-7906, Sigma Aldrich, USA) for 2 h to prevent non-specific adsorption of quantum dots to the surface. Fluorescence imaging was carried out on an inverted fluorescence microscope (IX-71, Olympus Corporation, Japan) in wide field configuration with an oil-immersion objective (PlanApo, 100X, NA 1.45 Oil TIRFM, Olympus Corporation, Japan) corresponding to a pixel size of 160 nm. The illumination light (X-Cite Series 120, EXFO Life Sciences, Canada) was filtered using an excitation bandpass filter (HQ 470/40, AHF Analysetechnik, Germany) and passed through a beam splitter (z 488 RDC, AHF Analysetechnik, Germany). Fluorescence light was filtered with an emission filter (HQ 510 LP, AHF Analysetechnik, Germany) and imaged on an EMCCD camera (Andor iXon+ DU897E, Andor Technology, North Ireland) using a frame rate of 31 fps and an exposure time of 0.03029 s. Single quantum dot trajectories were calculated using the ParticleTracker Plugin [184] for ImageJ [164].

Lipid granule observation

Cell preparation

For the both data set (the small one and the extensive one), the cells were cultured on AA-Leu agar plates for 12-14 hours at 33C and after growth stored at 4C. Prior to the experiments, the cells were suspended into liquid AA medium. The fluid chamber for the experiments consisted of two glass slides glued together by a layer of double sticky tape as a spacer. The chamber was then filled with the cell suspension. The cells were allowed to relax on the surface for approximately 20minutes, so that they do not move during data acquisition. The experiments were performed at room temperature (25C).

Optical microscopy protocol

Bright-field microscopy imaging was carried out on a Leica DMIRE2 microscope, using an oil immersion objective (Leica PL Apo 100x), and the AVT Pike F100B camera (Allied Vision Technologies) and the SmartView software(Allied Vision Technologies) for recording. Particle tracking was carried out using the PolyParticleTracker Package in Matlab [176]. The error in the determination of the particle position of this method is 20 nm.

For the extensive data set, we only wanted cells in interphase, to avoid the variability that could arise if we compare cells in interphase or in mitosis. To confirm that the cells were in interphase, the GFP fluorescence of the membranes was observed. Particle tracking was carried out using the algorithm of [176]. This algorithm is well-established for intracellular granule tacking [175].

Optical tweezers protocol

Optical tweezers were created using a Nd:YVO_4 laser (Neodymium-doped yttrium orthovanadate, 5W Spectra Physics BL-106C, $\lambda = 1064$ nm, TEM_{00}), implemented in an inverted fluorescence microscope (Leica, DMI 6000) equipped with a cooled CCD camera (Andor, Ixon cooled EMCCD). We used an oil immersion objective from Leica (HCX PL Apo, 100/NA=1.4 oil Cs). The forward scattered light of the highly-refractive intracellular granules was recorded by a silicon quadrant photodiode (S5981, Hamamatsu) in the back focal plane, allowing nanometer precision detection of the granule position in the laser focus. The motion of a single granule in the laser focus was tracked for approximately 3 sec at a sampling frequency of 22 kHz. We used a laser power of approximately 0.05 W at the sample, rendering this technique nearly non-invasive. The voltage signal from the photodiode was digitalized using home-written Labview algorithms.

While the voltage signal is directly proportional to the particle position with respect to the trap center, due to lack of knowledge of the precise optical conditions in the cell we could not convert the voltage signal to absolute distance. The time averaged mean square displacement obtained from this method is therefore given in arbitrary units.

2 First-passage properties of persistent random walks

We here detail the computations of the first passage properties in the persistent random walker case.

We consider a cubic lattice in d dimensions, of size $(X_1, X_2, \ldots, X_d)$, with periodic boundary conditions. $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d)$ is an orthogonal base of this lattice. The random walker starts at $\mathbf{r}_0$ toward a given direction $\mathbf{e}_i$, and at each step, he has a probability p_1 to continue in the same direction, p_2 to go in the opposite one, and $p_3 = (1 - p_1 - p_2)/(2d - 2)$ to choose one of the orthogonal directions. Following [85], we note $p_1 = p_3 + \epsilon$ and $p_2 = p_3 - \delta$.

2.1 Mean first passage time

We start from the backward Kolomogorov equation:

$$\langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}, \mathbf{e}_i) = p_1 \langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}+\mathbf{e}_i, \mathbf{e}_i) + p_2 \langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}-\mathbf{e}_i, -\mathbf{e}_i) + p_3 \sum_{\pm \mathbf{e}_j \in \mathcal{B}, j \neq i} \langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}+\mathbf{e}_j, \mathbf{e}_j) + 1 \quad (144)$$

where $\langle \mathbf{T} \rangle(\mathbf{r}_T|\mathbf{r}, \mathbf{e}_i)$ is the ensemble-averaged mean first passage time to reach a target located in $\mathbf{r}_T$, starting from $\mathbf{r}$ and when the last step has been along $\mathbf{e}_i$ direction.

Using a discrete Fourier transform, this leads to:

$$\begin{aligned} \langle \widetilde{\mathbf{T}} \rangle(\mathbf{r}_T|\mathbf{q}, \mathbf{e}_i) + V e^{-2\imath\pi \mathbf{q}.\mathbf{r}_T} &= p_1 \langle \widetilde{\mathbf{T}} \rangle(\mathbf{r}_T|\mathbf{q}, \mathbf{e}_i) e^{2\imath\pi \mathbf{q}.\mathbf{e}_i} + p_2 \langle \widetilde{\mathbf{T}} \rangle(\mathbf{r}_T|\mathbf{q}, -\mathbf{e}_i) e^{-2\imath\pi \mathbf{q}.\mathbf{e}_i} \\ &\quad + p_3 \sum_{\pm \mathbf{e}_j \in \mathcal{B}, j \neq i} \langle \widetilde{\mathbf{T}} \rangle(\mathbf{r}_T|\mathbf{q}, \mathbf{e}_j) e^{2\imath\pi \mathbf{q}.\mathbf{e}_j} + V\delta(\mathbf{q}) \end{aligned} \quad (145)$$

The second term on the left hand is due to the fact that equation (144) is not valid for $\mathbf{r} = \mathbf{r}_T$. Indeed, equation (144) gives the first return time to $\mathbf{r}_T$, on the left hand, and the first passage time from $\mathbf{r}_T$ to $\mathbf{r}_T$ (equal to 0) on the right hand. We thus have to add the first return time, given by the Kac' formula $1/P_{\text{stat}}(\mathbf{r}_T) = V$, on the left hand.

Considering equation (144) for both $\mathbf{e}_i$ and $-\mathbf{e}_i$, we obtain

$$\begin{pmatrix} \langle \widetilde{\mathbf{T}} \rangle(\mathbf{r}_T|\mathbf{q}, \mathbf{e}_i) \\ \langle \widetilde{\mathbf{T}} \rangle(\mathbf{r}_T|\mathbf{q}, -\mathbf{e}_i) \end{pmatrix} = \begin{pmatrix} 1 - \epsilon e^{2\imath\pi \mathbf{q}.\mathbf{e}_i} & \delta e^{-2\imath\pi \mathbf{q}.\mathbf{e}_i} \\ \delta e^{2\imath\pi \mathbf{q}.\mathbf{e}_i} & 1 - \epsilon e^{-2\imath\pi \mathbf{q}.\mathbf{e}_i} \end{pmatrix}^{-1} \cdot \begin{pmatrix} V\left(\delta(\mathbf{q}) - e^{-2\imath\pi \mathbf{q}.\mathbf{r}_T}\right) + p_3 f(\mathbf{q}) \\ V\left(\delta(\mathbf{q}) - e^{-2\imath\pi \mathbf{q}.\mathbf{r}_T}\right) + p_3 f(\mathbf{q}) \end{pmatrix} \quad (146)$$

with

$$f(\mathbf{q}) = \sum_{\mathbf{e}_j \in \mathcal{B}} \left(\langle \widetilde{\mathbf{T}} \rangle(\mathbf{r}_T|\mathbf{q}, \mathbf{e}_j) e^{2\imath\pi \mathbf{q}.\mathbf{e}_j} + \langle \widetilde{\mathbf{T}} \rangle(\mathbf{r}_T|\mathbf{q}, -\mathbf{e}_j) e^{-2\imath\pi \mathbf{q}.\mathbf{e}_j} \right) \quad (147)$$

A summation over all $\mathbf{e}_i \in \mathcal{B}$ gives for $f(\mathbf{q})$

$$f(\mathbf{q}) = 2V\left(\delta(\mathbf{q}) - e^{-2\imath\pi \mathbf{q}.\mathbf{r}_T}\right) \frac{\displaystyle\sum_{\mathbf{e}_j \in \mathcal{B}} \frac{\cos(2\pi \mathbf{q}.\mathbf{e}_j) - (\epsilon + \delta)}{1 + \epsilon^2 - 2\epsilon \cos(2\pi \mathbf{q}.\mathbf{e}_j) - \delta^2}}{1 - 2p_3 \displaystyle\sum_{\mathbf{e}_j \in \mathcal{B}} \frac{\cos(2\pi \mathbf{q}.\mathbf{e}_j) - (\epsilon + \delta)}{1 + \epsilon^2 - 2\epsilon \cos(2\pi \mathbf{q}.\mathbf{e}_j) - \delta^2}} \quad (148)$$

We thus obtain

$$\langle\widetilde{\mathbf{T}}\rangle(\mathbf{r}_T|\mathbf{q},\mathbf{e}_i) = V\frac{1-(\epsilon+\delta)e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}}{1+\epsilon^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_i)-\delta^2}\frac{\delta(\mathbf{q})-e^{-2\imath\pi\mathbf{q}.\mathbf{r}_T}}{1-2p_3\sum_{\mathbf{e}_j\in\mathcal{B}}\frac{\cos(2\pi\mathbf{q}.\mathbf{e}_j)-(\epsilon+\delta)}{1+\epsilon^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_j)-\delta^2}} \tag{149}$$

After an inverse Fourier transform, and a shift of λ so that $\langle\mathbf{T}\rangle(\mathbf{r}_T|\mathbf{r}_T,\mathbf{e}_i)=0$ (equation (144) is defined up to a constant), we retrieve equation (2.68):

$$\langle\mathbf{T}\rangle(\mathbf{r}_T|\mathbf{r}_S,\mathbf{e}_S) = \sum_{\mathbf{q}\neq\mathbf{0}}\left(\frac{1-(\epsilon+\delta)e^{-2\imath\pi\mathbf{q}.\mathbf{e}_S}}{1+\epsilon^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_S)-\delta^2}\frac{1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}}{1-2p_3\sum_{\mathbf{e}_j\in\mathcal{B}}\frac{\cos(2\pi\mathbf{q}.\mathbf{e}_j)-(\epsilon+\delta)}{1+\epsilon^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_j)-\delta^2}}\right) \tag{150}$$

2.2 Approximations

$\epsilon \to 1$

We can approximate the MFPT for the case $\epsilon \to 1$, using a small reorganization:

$$\langle\mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle = \frac{V(\delta-\epsilon)}{1-\epsilon+\delta}+\frac{2\epsilon}{(1+\epsilon+\delta)(1-\epsilon+\delta)}\sum_{\mathbf{q}\neq\mathbf{0}}\frac{1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}}{\frac{1}{d}\sum_{\mathbf{e}_j\in\mathcal{B}}\frac{1}{1+\frac{(\epsilon-1)^2-\delta^2}{2\epsilon(1-\cos(2\pi\mathbf{q}.\mathbf{e}_j))}}} \tag{151}$$

We focus on an square ($d = 2$), but the same calculation can be made in any dimension:

$$\begin{aligned}
\langle \mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle &= \frac{X^2(\delta-\epsilon)}{1-\epsilon+\delta} + \frac{2\epsilon}{(1+\epsilon+\delta)(1-\epsilon+\delta)} \\
&\times \left(\sum_{k=1}^{X-1}\sum_{l=1}^{X-1} \frac{1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}}{\frac{1}{2}\sum_{\mathbf{e}_j\in\mathcal{B},\mathbf{q}.\mathbf{e}_j\neq\mathbf{0}} \frac{1}{1+\frac{(\epsilon-1)^2-\delta^2}{2\epsilon(1-\cos(2\pi\mathbf{q}.\mathbf{e}_j))}}} \right. \\
&\left. + \sum_{k=1|_{l=0}}^{X-1} \frac{2\left(1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}\right)}{1+\frac{(\epsilon-1)^2-\delta^2}{2\epsilon(1-\cos(2\pi\mathbf{q}.\mathbf{e}_x))}} + \sum_{l=1|_{k=0}}^{X-1} \frac{2\left(1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}\right)}{1+\frac{(\epsilon-1)^2-\delta^2}{2\epsilon(1-\cos(2\pi\mathbf{q}.\mathbf{e}_y))}} \right) \\
\underset{\epsilon\to1}{\simeq} & \frac{X^2(\delta-\epsilon)}{1-\epsilon+\delta} + \frac{2\epsilon}{(1+\epsilon+\delta)(1-\epsilon+\delta)} \left(\sum_{k=1}^{X-1}\sum_{l=1}^{X-1} \left(1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}\right) \right. \\
&\times \left(1+\frac{1}{2}\sum_{\mathbf{e}_j\in\mathcal{B},\mathbf{q}.\mathbf{e}_j\neq\mathbf{0}} \frac{(\epsilon-1)^2-\delta^2}{2\epsilon(1-\cos(2\pi\mathbf{q}.\mathbf{e}_j))}\right) + 2\sum_{k=1|_{l=0}}^{X-1}\left(1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}\right) \\
&\times \left(1+\frac{(\epsilon-1)^2-\delta^2}{2\epsilon(1-\cos(2\pi\mathbf{q}.\mathbf{e}_x))}\right) + 2\sum_{l=1|_{k=0}}^{X-1}\left(1-e^{2\imath\pi\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}\right) \\
&\left. \times \left(1+\frac{(\epsilon-1)^2-\delta^2}{2\epsilon(1-\cos(2\pi\mathbf{q}.\mathbf{e}_y))}\right)\right) \\
\underset{\epsilon\to1}{\simeq} & \frac{X^2(\delta-\epsilon)}{1-\epsilon+\delta} + \frac{2\epsilon}{(1+\epsilon+\delta)(1-\epsilon+\delta)} \\
&\times\Big((X-1)(X+3) - \left(X\delta_{(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_x}+1\right)\left(X\delta_{(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_y}+1\right)+4\Big) \\
&+\frac{1-\epsilon-\delta}{1+\epsilon+\delta} \\
&\times\left((X-1)\frac{X^2-1}{6} + (\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_x\left(X-(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_x\right) + (\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_y\left(X-(\mathbf{r}_S-\mathbf{r}_T).\mathbf{e}_y\right)\right)
\end{aligned}$$

$r \to \infty$

We can approximate the discrete sum of equation (2.68) with an integral when all $X_i \to \infty$:

$$\begin{aligned}
\langle \mathbf{T}(\mathbf{r}_T|\mathbf{r}_S)\rangle &\simeq \frac{V(\delta-\epsilon)}{1-\epsilon+\delta}(1-\delta(\mathbf{r}_S,\mathbf{r}_T)) + \frac{V(1+\epsilon^2-\delta^2)}{(2\pi)^d(1+\epsilon+\delta)(1-\epsilon+\delta)} \\
&\times \int_{[0,2\pi]^d} \frac{\left(1-e^{\imath\mathbf{q}.(\mathbf{r}_S-\mathbf{r}_T)}\right)d\mathbf{q}}{1-\frac{(\epsilon-1)^2-\delta^2}{d}\sum_{\mathbf{e}_j\in\mathcal{B}}\frac{\cos(\mathbf{q}.\mathbf{e}_j)}{1+\epsilon^2-2\epsilon\cos(\mathbf{q}.\mathbf{e}_j)-\delta^2}}
\end{aligned}$$

At large distances, $\mathbf{r} = \mathbf{r}_T - \mathbf{r}_S$ is great enough, namely that $r = \|\mathbf{r}\| = \sqrt{\sum_i (\mathbf{r}.\mathbf{e}_i)^2} \gg 1$. We develop the integral for small $\mathbf{q}$, and we average this result for a given $r = \|\mathbf{r}\|$:

$$\begin{aligned}
\langle \mathbf{T}(r) \rangle &\simeq V \left(\frac{\delta - \epsilon}{1 - \epsilon + \delta} + \Delta \right) (1 - \delta(r)) + \frac{V}{(2\pi)^d} \frac{1 - \epsilon - \delta}{1 + \epsilon + \delta} \\
&\qquad \times \int_{[-\pi,\pi]^d} \frac{d\mathbf{q}}{\|\mathbf{q}\|^2} \frac{\displaystyle\int_0^\pi \left(1 - e^{\imath \|\mathbf{q}\| r \cos(\theta)}\right) \sin(\theta)^{d-2} d\theta}{\displaystyle\int_0^\pi \sin(\theta)^{d-2} d\theta} \\
&\simeq V \left(\frac{\delta - \epsilon}{1 - \epsilon + \delta} + \Delta \right) (1 - \delta(r)) + \frac{V}{(2\pi)^d} \frac{1 - \epsilon - \delta}{1 + \epsilon + \delta} \\
&\qquad \times \int_{[-\pi,\pi]^d} \frac{d\mathbf{q}}{\|\mathbf{q}\|^2} \left(1 - \frac{J_{d/2-1}(\|\mathbf{q}\| r) \Gamma\left(\frac{d}{2}\right)}{\left(\frac{\|\mathbf{q}\| r}{2}\right)^{d/2-1}} \right) \\
&\simeq V \left(\left(\frac{\delta - \epsilon}{1 - \epsilon + \delta} + \Delta \right) (1 - \delta(r)) + \frac{1 - \epsilon - \delta}{1 + \epsilon + \delta} \left(G_{0,d}(0) - G_{0,d}(r) \right) \right), \qquad (152)
\end{aligned}$$

where $J_\nu(x)$ are the Bessel functions of the first kind, and $G_{0,d}$ the infinite Green function in d dimensions. Δ is a constant correction due to the small $\mathbf{q}$ expansion, that we can estimate:

$$\begin{aligned}
\Delta \simeq \frac{(1 + \epsilon^2 - \delta^2)}{(1 + \epsilon + \delta)(1 - \epsilon + \delta)} \int_{[0,2\pi]^d} \frac{d\mathbf{q}}{(2\pi)^d} & \frac{1}{1 - \dfrac{(\epsilon - 1)^2 - \delta^2}{d} \displaystyle\sum_{\mathbf{e}_j \in \mathcal{B}} \frac{\cos(\mathbf{q}.\mathbf{e}_j)}{1 + \epsilon^2 - 2\epsilon \cos(\mathbf{q}.\mathbf{e}_j) - \delta^2}} \\
& - \frac{(\epsilon - 1)^2 - \delta^2}{1 + \epsilon^2 - \delta^2} \frac{1}{1 - \dfrac{1}{d} \displaystyle\sum_{\mathbf{e}_j \in \mathcal{B}} \cos(\mathbf{q}.\mathbf{e}_j)}
\end{aligned}$$

2.3 Persistent pseudo-Green functions

We note $P_{\mathbf{e}_i}(\mathbf{r}, t | \mathbf{r}_S, \mathbf{e}_S)$ the conditional propagator, namely the probability to be at the position $\mathbf{r}$ at time t, when the last step have been following the vector $\mathbf{e}_i$, starting from $\mathbf{r}_S$ with orientation $\mathbf{e}_S$. For simplicity sake, we will skip $\mathbf{r}_S$ and $\mathbf{e}_S$ in the notation. The conditional propagator satisfies the master Chapman-Kolmogorov equation:

$$P_{\mathbf{e}_i}(\mathbf{r}, t+1) = p_1 P_{\mathbf{e}_i}(\mathbf{r} - \mathbf{e}_i, t) + p_2 P_{-\mathbf{e}_i}(\mathbf{r} - \mathbf{e}_i, t) + p_3 \sum_{\mathbf{e}_j \in \mathcal{B}, j \neq i} \left(P_{\mathbf{e}_j}(\mathbf{r} - \mathbf{e}_i, t) + P_{-\mathbf{e}_j}(\mathbf{r} - \mathbf{e}_i, t) \right) \quad (153)$$

We can transform this equation successively using a discrete Fourier transform (assuming periodic boundary conditions) and a Z transform:

$$\widehat{P}_{\mathbf{e}_i}(\mathbf{q}, z) = \sum_{t=0}^{\infty} z^t \widetilde{P}_{\mathbf{e}_i}(\mathbf{q}, t) = \sum_{t=0}^{\infty} z^t \sum_{\mathbf{r}=(0,\ldots,0)}^{(X_1-1,\ldots,X_d-1)} P_{\mathbf{e}_i}(\mathbf{r}, t) e^{-2\imath\pi \mathbf{q}.\mathbf{r}}, \qquad (154)$$

where $\mathbf{q} = (k_1/X_1, k_2/X_2, \ldots, k_d/X_d)$. Equation (2.83) then become:

$$z^{-1}\left(\widehat{P}_{\mathbf{e}_i}(\mathbf{q},z) - \widetilde{P}_{\mathbf{e}_i}(\mathbf{q},t=0)\right) = \left(p_1\widehat{P}_{\mathbf{e}_i}(\mathbf{q},z) + p_2\widehat{P}_{-\mathbf{e}_i}(\mathbf{q},z) + p_3\sum_{\pm\mathbf{e}_j\in\mathcal{B},j\neq i}\widehat{P}_{\mathbf{e}_j}(\mathbf{q},z)\right)e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i} \quad (155)$$

$$= \epsilon\widehat{P}_{\mathbf{e}_i}(\mathbf{q},z)e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i} - \delta\widehat{P}_{-\mathbf{e}_i}(\mathbf{q},z)e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i} + p_3\widehat{P}(\mathbf{q},z)e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i} \quad (156)$$

Considering this equation for both $\mathbf{e}_i$ and $-\mathbf{e}_i$, we obtain:

$$\begin{pmatrix}\widehat{P}_{\mathbf{e}_i}(\mathbf{q},z)\\ \widehat{P}_{-\mathbf{e}_i}(\mathbf{q},z)\end{pmatrix} = \begin{pmatrix}1-\epsilon z e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i} & \delta z e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}\\ \delta z e^{2\imath\pi\mathbf{q}.\mathbf{e}_i} & 1-\epsilon z e^{2\imath\pi\mathbf{q}.\mathbf{e}_i}\end{pmatrix}^{-1}\cdot\begin{pmatrix}\widetilde{P}_{\mathbf{e}_i}(\mathbf{q},0) + p_3 z\widehat{P}(\mathbf{q},z)e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}\\ \widetilde{P}_{-\mathbf{e}_i}(\mathbf{q},0) + p_3 z\widehat{P}(\mathbf{q},z)e^{2\imath\pi\mathbf{q}.\mathbf{e}_i}\end{pmatrix} \quad (157)$$

In particular, we have:

$$\widehat{P}_{\mathbf{e}_i}(\mathbf{q},z) = \frac{(1-\epsilon z e^{2\imath\pi\mathbf{q}.\mathbf{e}_i})\widetilde{P}_{\mathbf{e}_i}(\mathbf{q},0) - \delta z e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}\widetilde{P}_{-\mathbf{e}_i}(\mathbf{q},0) + p_3 z(e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i} - (\epsilon+\delta)z)\widehat{P}(\mathbf{q},z)}{1+\epsilon^2z^2-\delta^2z^2-2\epsilon z\cos(2\pi\mathbf{q}.\mathbf{e}_i)} \quad (158)$$

We then can determine $\widehat{P}(\mathbf{q},z)$ by summation over all $\mathbf{e}_i$:

$$\widehat{P}(\mathbf{q},z) = \frac{\displaystyle\sum_{\mathbf{e}_i\in\mathcal{B}}\frac{\left(1-(\epsilon+\delta)ze^{2\imath\pi\mathbf{q}.\mathbf{e}_i}\right)\widetilde{P}_{\mathbf{e}_i}(\mathbf{q},0) + \left(1-(\epsilon+\delta)ze^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}\right)\widetilde{P}_{-\mathbf{e}_i}(\mathbf{q},0)}{1+\epsilon^2z^2-\delta^2z^2-2\epsilon z\cos(2\pi\mathbf{q}.\mathbf{e}_i)}}{1-2p_3z\displaystyle\sum_{\mathbf{e}_i\in\mathcal{B}}\frac{\cos(2\pi\mathbf{q}.\mathbf{e}_i)-(\epsilon+\delta)z}{1+\epsilon^2z^2-\delta^2z^2-2\epsilon z\cos(2\pi\mathbf{q}.\mathbf{e}_i)}} \quad (159)$$

We then have an explicit expression for all $\widehat{P}_{\mathbf{e}_i}(\mathbf{q},z)$. We can then get an explicit expression for pseudo-Green functions $H(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S)$:

$$H(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S) = \sum_{t=0}^{\infty}\left(P_{\mathbf{e}_i}(\mathbf{r},t|\mathbf{r}_S,\mathbf{e}_S) - P_{\text{stat}}(\mathbf{r},\mathbf{e}_i)\right), \quad (160)$$

where $\mathbf{r}_S$ is the initial position, $\mathbf{e}_S$ the first direction chosen, and $P_{\text{stat}}(\mathbf{r},\mathbf{e}_i) = \prod_i X_i^{-1}/(2d)$ is the stationary probability. The initial condition gives $\widetilde{P}_{\mathbf{e}_i}(\mathbf{q},0) = \delta(\mathbf{e}_i,\mathbf{e}_S)\exp(-2\imath\pi\mathbf{q}.\mathbf{r}_S)$, where $\delta(\mathbf{x},\mathbf{y}) = 1$ if $\mathbf{x}=\mathbf{y}$, else 0. We finally obtain:

$$H(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S) = \frac{1}{\displaystyle\prod_i X_i}\sum_{\mathbf{q}\neq\mathbf{0}}e^{2\imath\pi\mathbf{r}.\mathbf{q}}\left(\frac{\left((1-\epsilon e^{2\imath\pi\mathbf{q}.\mathbf{e}_i})\delta(\mathbf{e}_i,\mathbf{e}_S) - \delta e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}\delta(-\mathbf{e}_i,\mathbf{e}_S)\right)e^{-2\imath\pi\mathbf{q}.\mathbf{r}_S}}{1+\epsilon^2-\delta^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_i)}\right.$$
$$\left. + \frac{p_3(e^{-2\imath\pi\mathbf{q}.\mathbf{e}_i}-(\epsilon+\delta))\widehat{P}(\mathbf{q},z=1)}{1+\epsilon^2-\delta^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_i)}\right) \quad (161)$$

and

$$\widehat{P}(\mathbf{q},z=1) = \frac{\dfrac{\left(1-(\epsilon+\delta)e^{2\imath\pi\mathbf{q}.\mathbf{e}_S}\right)}{1+\epsilon^2-\delta^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_S)}e^{-2\imath\pi\mathbf{q}.\mathbf{r}_S}}{1-2p_3\displaystyle\sum_{\mathbf{e}_i\in\mathcal{B}}\frac{\cos(2\pi\mathbf{q}.\mathbf{e}_i)-(\epsilon+\delta)}{1+\epsilon^2-\delta^2-2\epsilon\cos(2\pi\mathbf{q}.\mathbf{e}_i)}} \quad (162)$$

We remove the $\mathbf{q}=\mathbf{0}$ case in the summation, because it correspond to the P_{stat} term. We can check that we retrieve the brownian motion if we choose $\epsilon=\delta=0$:

$$\sum_{\pm\mathbf{e}_i\in\mathcal{B}} H(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S)=\frac{1}{\prod_i X_i}\sum_{\mathbf{q}\neq\mathbf{0}}\frac{e^{2\imath\pi\mathbf{q}.(\mathbf{r}-\mathbf{r}_S)}}{1-\frac{1}{d}\sum_{\mathbf{e}_i\in\mathcal{B}}\cos(2\pi\mathbf{q}.\mathbf{e}_i)} \tag{163}$$

2.4 Retrieving $\langle\mathbf{T}\rangle$ using pseudo-Green functions

The first passage time from a given site $\mathbf{r}_S$ with velocity $\mathbf{e}_S$ to a target site $\mathbf{r}_T$ is in fact a first passage problem with $2d$ sites: we consider that all sites $\mathbf{r}$ are divided in $2d$ distinct sites, each with a given velocity $\mathbf{e}_i$. Following [65], we have, for the mean first passage time $\langle\mathbf{T}\rangle(\mathbf{r}_T|\mathbf{r}_S,\mathbf{e}_S)$:

$$\frac{\langle\mathbf{T}\rangle(\mathbf{r}_T|\mathbf{r}_S,\mathbf{e}_S)}{2dV}=-H(\mathbf{r}_T,\mathbf{e}_S|\mathbf{r}_S,\mathbf{e}_S)+\sum_{\mathbf{e}_i\in\mathcal{B}}P_i^+H(\mathbf{r}_T,\mathbf{e}_i|\mathbf{r}_T,\mathbf{e}_S)+P_i^-H(\mathbf{r}_T,-\mathbf{e}_i|\mathbf{r}_T,\mathbf{e}_S) \tag{164}$$

$V=\prod_i X_i$ is the volume of the subdomain, and $P_i^{\pm}$ the conditional splitting probabilities to reach $\mathbf{r}_T$ with a velocity $\pm\mathbf{e}_i$ starting from $\mathbf{r}_S$ with a velocity $\mathbf{e}_S$. Those probabilities are given by:

$$\begin{pmatrix} P_1^+\\ P_2^+\\ \vdots\\ P_d^- \end{pmatrix} = \begin{pmatrix} H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_T\\ \mathbf{e}_1&\mathbf{e}_1\end{array}\right)-H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_T\\ \mathbf{e}_2&\mathbf{e}_1\end{array}\right) & \dots & H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_T\\ \mathbf{e}_1&-\mathbf{e}_d\end{array}\right)-H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_T\\ \mathbf{e}_2&-\mathbf{e}_d\end{array}\right)\\ \vdots & \ddots & \vdots\\ H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_T\\ \mathbf{e}_1&\mathbf{e}_1\end{array}\right)-H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_T\\ -\mathbf{e}_d&\mathbf{e}_1\end{array}\right) & \dots & H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_T\\ \mathbf{e}_1&-\mathbf{e}_d\end{array}\right)-H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_T\\ -\mathbf{e}_d&-\mathbf{e}_d\end{array}\right)\\ 1 & \dots & 1 \end{pmatrix}^{-1} \times \begin{pmatrix} H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_S\\ \mathbf{e}_1&\mathbf{e}_S\end{array}\right)-H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_S\\ \mathbf{e}_2&\mathbf{e}_S\end{array}\right)\\ \vdots\\ H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_S\\ \mathbf{e}_1&\mathbf{e}_S\end{array}\right)-H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_S\\ -\mathbf{e}_d&\mathbf{e}_S\end{array}\right)\\ 1 \end{pmatrix}, \tag{165}$$

where

$$H\left(\begin{array}{c|c}\mathbf{r}_T&\mathbf{r}_S\\ \mathbf{e}_i&\mathbf{e}_S\end{array}\right)=H\left(\mathbf{r}_T,\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S\right) \tag{166}$$

Knowing all the pseudo-Green functions, we retrieve an analytic expression for the first passage time. I did not manage to link analytically those two expressions, but the numerical computation of both expression (2.68) and (164) lead to the same result, as expected.

2.5 Reflective boundary extension

We can extend our formalism beyond periodic boundary conditions. We will here show how to transform our previous results to treat reflective boundary conditions.

We can consider that a system of size X with reflective boundary conditions is equivalent to a $2X$ large system with periodic boundary conditions. The slight difficulty is to define correctly what happens when a random walker hit the border in the original network: we consider here that the random walker stay at the same position, but the velocity is reflected by the wall. For

the $2X$ system, the random walker perform a classical persistent random walk. In $1D$, this can be written as follow:

$$H^{\square}(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S) = H^{\circlearrowleft}(\mathbf{r}+\mathbf{1}/2,\mathbf{e}_i|\mathbf{r}_S+\mathbf{1}/2,\mathbf{e}_S) + H^{\circlearrowleft}(s(\mathbf{r}+\mathbf{1}/2),s(\mathbf{e}_i)|\mathbf{r}_S+\mathbf{1}/2,\mathbf{e}_S), \quad (167)$$

where, $H^{\square}$ is the pseudo-Green function for a X-large system with reflective boundaries, $H^{\circlearrowleft}$ for a $2X$-large system with periodic boundaries, $\mathbf{1}=\{1\}$ a vector, and $s(.)$ the symmetry operator by the $x=X$ point. The vector $\mathbf{1}/2$ is added so that near the border, the random walker can stay on the same position with a velocity reflected by the wall. We can write this expression as :

$$H^{\square}(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S) = (\mathrm{Id}+s)\circ H^{\circlearrowleft}(\mathbf{r}+\mathbf{1}/2,\mathbf{e}_i|\mathbf{r}_S+\mathbf{1}/2,\mathbf{e}_S), \quad (168)$$

The operator $\mathrm{Id}+s$ apply only on $\mathbf{r}$ and on $\mathbf{e}_i$.

If we then go to the general case with d dimension, we can obtain a similar result:

$$H^{\square}(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S) = \prod_i(\mathrm{Id}+s_i)\circ H^{\circlearrowleft}(\mathbf{r}+\mathbf{1}/2,\mathbf{e}_i|\mathbf{r}_S+\mathbf{1}/2,\mathbf{e}_S), \quad (169)$$

with $H^{\square}$ is the pseudo-Green function for a $(X_1,\ldots,X_d)$-large system with reflective boundaries, $H^{\circlearrowleft}$ for a $(2X_1,\ldots,2X_d)$-large system with periodic boundaries, $\mathbf{1}=(1,\ldots,1)$ a vector, and $s_i(.)$ the symmetry operator by the $x_i=X_i$ hyperplane. We thus have a sum of $2d$ terms. In particular, $d=2$ gives:

$$\begin{aligned} H^{\square}(\mathbf{r},\mathbf{e}_i|\mathbf{r}_S,\mathbf{e}_S) \;=\; & H^{\circlearrowleft}(\mathbf{r}+\mathbf{1}/2,\mathbf{e}_i|\mathbf{r}_S+\mathbf{1}/2,\mathbf{e}_S) + H^{\circlearrowleft}(s_x(\mathbf{r}+\mathbf{1}/2),s_x(\mathbf{e}_i)|\mathbf{r}_S+\mathbf{1}/2,\mathbf{e}_S) \\ & + H^{\circlearrowleft}(s_y(\mathbf{r}+\mathbf{1}/2),s_y(\mathbf{e}_i)|\mathbf{r}_S+\mathbf{1}/2,\mathbf{e}_S) \\ & + H^{\circlearrowleft}(s_x\circ s_y(\mathbf{r}+\mathbf{1}/2),s_x\circ s_y(\mathbf{e}_i)|\mathbf{r}_S+\mathbf{1}/2,\mathbf{e}_S) \end{aligned} \quad (170)$$

This result has been checked numerically, and as shown in Figure 28, it fits perfectly the simulations.

3 First-passage properties for a moving target

We here detail the computations of the first passage properties in the moving target case.

We consider a 1D segment of size L, and searcher and a target diffusing with a coefficient D, and boundaries successively absorbing (first section) and reflective (second section). We transform this problem with two moving particle to a 2D problem, with a single moving particle.

3.1 Encounter probability

We start from the standard 2D diffusion equation:

$$\frac{\partial P}{\partial t}(x,y,t|x_0,y_0) = D\left(\frac{\partial^2 P}{\partial x^2}(x,y,t|x_0,y_0) + \frac{\partial^2 P}{\partial y^2}(x,y,t|x_0,y_0)\right) \quad (171)$$

where $P(x,y,t|x_0,y_0)$ is the propagator.

The solution for a square of size L with absorbing boundaries is

$$P_{\square}(x,y,t|x_1,x_2) = \frac{4}{L^2}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\sin\left(\frac{kx\pi}{L}\right)\sin\left(\frac{kx_1\pi}{L}\right)\sin\left(\frac{ly\pi}{L}\right)\sin\left(\frac{lx_2\pi}{L}\right)e^{-\frac{(k^2+l^2)D\pi^2 t}{L^2}}. \quad (172)$$

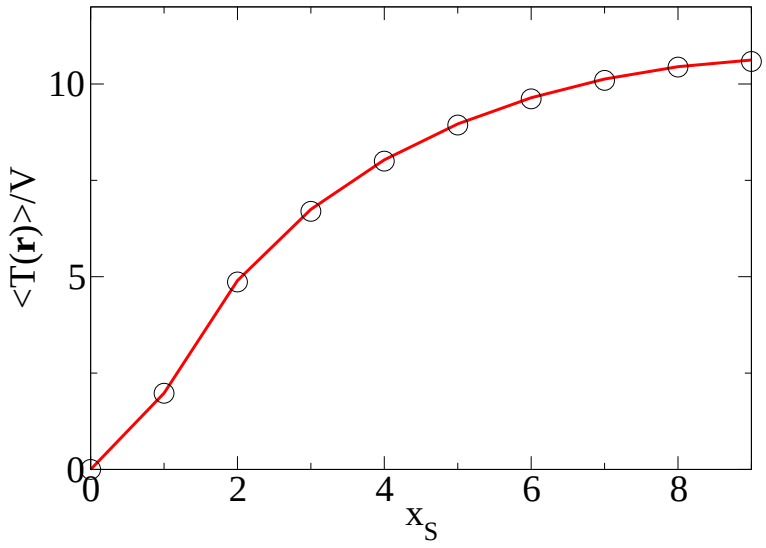

Figure 28: Mean first-passage time for a 10×10-large network with reflective boundaries, where $\mathbf{r}_s = (x_S, 0)$, $\mathbf{r}_T = (0, 0)$, $\mathbf{e}_S = (1, 0)$, with $x_S \in [0, 9]$. Here $\epsilon = -0.04$ and $\delta = 0.32$. The black circles stand for numerical simulations, the red line for theoretical result of equation (164).

The probability that a random walker diffusing in a 2D square of size L with absorbing boundary hit a given wall is:

$$\begin{aligned}
\text{Prob}_{\text{wall},\square}(x = L|x_1, x_2) &= \int_0^\infty \int_{y=0}^{L} -D \left.\frac{\partial P(x, y, t|x_1, x_2)}{\partial x}\right|_{x=L} dydt \\
&= -\frac{4}{\pi^2} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} (-1)^k \frac{k}{l} \left(1 - (-1)^l\right) \sin\left(\frac{kx_1\pi}{L}\right) \sin\left(\frac{lx_2\pi}{L}\right) \frac{1}{k^2 + l^2} \\
&= -\frac{8}{\pi^2} \sum_{l=0}^{\infty} \frac{\sin\left(\frac{(2l+1)x_2\pi}{L}\right)}{2l+1} \sum_{k=1}^{\infty} \frac{(-1)^k k}{k^2 + (2l+1)^2} \sin\left(\frac{kx_1\pi}{L}\right). \quad (173)
\end{aligned}$$

We then exploit the Fourier expansion of the sinh function (more precisely, the periodic function equal to sinh between $-\pi$ and π),

$$\frac{\sinh(ax)}{\sinh(a\pi)} = -\frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n n}{a^2 + n^2} \sin(nx), \quad (174)$$

to rewrite the probability $\text{Prob}_{\text{wall},\square}$ to hit the wall at $x = L$,

$$\text{Prob}_{\text{wall},\square}(x = L|x_1, x_2) = \frac{4}{\pi} \sum_{l=0}^{\infty} \frac{\sin\left(\frac{(2l+1)x_2\pi}{L}\right)}{2l+1} \frac{\sinh\left(\frac{(2l+1)x_1\pi}{L}\right)}{\sinh\left((2l+1)\pi\right)}. \quad (175)$$

This expression simplifies when we introduce the approximation $\sinh((2l+1)\pi) \approx \exp((2l+1)\pi)/2$, leading to

$$
\begin{aligned}
\mathrm{Prob}_{\mathrm{wall},\square}(x = L|x_1, x_2) \quad &\approx \frac{8}{\pi} \sum_{l=0}^{\infty} \frac{\sin\left(\frac{(2l+1)x_2\pi}{L}\right)}{2l+1} \sin \\
&= \frac{2}{\imath\pi} \sum_{l=0}^{\infty} \left(\frac{\exp\left(\frac{\pi}{L}(\imath x_2 + x_1 - L)\right)^{2l+1}}{2l+1} + \frac{\exp\left(\frac{\pi}{L}(-\imath x_2 - x_1 - L)\right)^{2l+1}}{2l+1} \right. \\
&\qquad \left. - \frac{\exp\left(\frac{\pi}{L}(\imath x_2 - x_1 - L)\right)^{2l+1}}{2l+1} - \frac{\exp\left(\frac{\pi}{L}(-\imath x_2 + x_1 - L)\right)^{2l+1}}{2l+1} \right) \\
&= \frac{1}{\imath\pi} \left(\ln\left(\frac{1+\exp\left(\frac{\pi}{L}(\imath x_2 + x_1 - L)\right)}{1-\exp\left(\frac{\pi}{L}(\imath x_2 + x_1 - L)\right)} \frac{1-\exp\left(\frac{\pi}{L}(-\imath x_2 + x_1 - L)\right)}{1+\exp\left(\frac{\pi}{L}(-\imath x_2 + x_1 - L)\right)} \right) \right. \\
&\qquad \left. - \ln\left(\frac{1+\exp\left(\frac{\pi}{L}(\imath x_2 - x_1 - L)\right)}{1-\exp\left(\frac{\pi}{L}(\imath x_2 - x_1 - L)\right)} \frac{1-\exp\left(\frac{\pi}{L}(-\imath x_2 - x_1 - L)\right)}{1+\exp\left(\frac{\pi}{L}(-\imath x_2 - x_1 - L)\right)} \right) \right) \\
&= \frac{1}{\imath\pi} \left(\ln\left(\frac{1+\imath\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(L-x_1)}{L}\right)}}{1-\imath\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(L-x_1)}{L}\right)}} \right) - \ln\left(\frac{1+\imath\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(L+x_1)}{L}\right)}}{1-\imath\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(L+x_1)}{L}\right)}} \right) \right) \\
&= \frac{2}{\pi} \left(\arctan\left(\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(L-x_1)}{L}\right)} \right) - \arctan\left(\frac{\sin\left(\frac{\pi x_2}{L}\right)}{\sinh\left(\frac{\pi(L+x_1)}{L}\right)} \right) \right)
\end{aligned}
$$

We can use the exact expression (3.17) to obtain an exact expression of the encounter proba-

bility:

$$
\begin{aligned}
P_M(x_1,x_2) &= = 1 - \mathrm{Prob}_{\mathrm{wall}}(x = L|x_1,x_2) - \mathrm{Prob}_{\mathrm{wall}}(y = 0|x_1,x_2) && (182)\\
&= 1 - (\mathrm{Pr}_{\mathrm{wall}}(x = L|x_1,x_2) - \mathrm{Pr}_{\mathrm{wall}}(x = L|x_2,x_1) \\
&\qquad + \mathrm{Pr}_{\mathrm{wall}}(y = 0|x_1,x_2) - \mathrm{Pr}_{\mathrm{wall}}(y = 0|x_2,x_1)) && (183)\\
&= 1 - (\mathrm{Pr}_{\mathrm{wall}}(x = L|x_1,x_2) - \mathrm{Pr}_{\mathrm{wall}}(x = 0|x_1,x_2) \\
&\qquad + \mathrm{Pr}_{\mathrm{wall}}(y = 0|x_1,x_2) - \mathrm{Pr}_{\mathrm{wall}}(y = L|x_1,x_2)) && (184)\\
&= 2\,(\mathrm{Pr}_{\mathrm{wall}}(x = 0|x_1,x_2) + \mathrm{Pr}_{\mathrm{wall}}(y = L|x_1,x_2)) && (185)\\
&= \frac{8}{\pi}\left(\sum_{l=0}^{\infty}\frac{\sin\left(\frac{(2l+1)x_2\pi}{L}\right)}{2l+1}\frac{\sinh\left(\frac{(2l+1)(L-x_1)\pi}{L}\right)}{\sinh\left((2l+1)\pi\right)}\right. \\
&\qquad \left. + \sum_{l=0}^{\infty}\frac{\sin\left(\frac{(2l+1)x_1\pi}{L}\right)}{2l+1}\frac{\sinh\left(\frac{(2l+1)x_2\pi}{L}\right)}{\sinh\left((2l+1)\pi\right)}\right) && (186)
\end{aligned}
$$

3.2 Mean first encounter time

We start with the propagator on the $\sqrt{2}L \times \sqrt{2}L$ square, with absorbing boundaries. As previously, starting from (x_0, y_0), we have

$$
P(x,y,t|x_0,y_0) = \frac{2}{L^2}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\sin\left(\frac{kx\pi}{\sqrt{2}L}\right)\sin\left(\frac{kx_0\pi}{\sqrt{2}L}\right)\sin\left(\frac{ly\pi}{\sqrt{2}L}\right)\sin\left(\frac{ly_0\pi}{\sqrt{2}L}\right)e^{-\frac{(k^2+l^2)D\pi^2 t}{2L^2}}.
$$

We then compute the survival probability, $\mathscr{S}(t|x_0,y_0)$, namely the probability to still be in the square at time t, starting from (x_0, y_0)

$$
\begin{aligned}
\mathscr{S}(t|x_0,y_0) &= \int_0^{\sqrt{2}L}\int_0^{\sqrt{2}L} P(x,y,t|x_0,y_0)dxdy \\
&= \frac{16}{\pi^2}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{2k+1}\frac{\sin\left(\frac{(2l+1)y_0\pi}{\sqrt{2}L}\right)}{2l+1}e^{-\frac{((2k+1)^2+(2l+1)^2)D\pi^2 t}{2L^2}}.
\end{aligned}
$$

$$
\begin{aligned}
\langle \mathbf{T} \rangle (x_0, y_0) &= \int_0^\infty \mathscr{S}(t|x_0, y_0) dt \\
&= \frac{32L^2}{D\pi^4} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{2k+1} \frac{\sin\left(\frac{(2l+1)y_0\pi}{\sqrt{2}L}\right)}{2l+1} \frac{1}{(2k+1)^2 + (2l+1)^2} \\
&= \frac{32L^2}{D\pi^4} \sum_{k=0}^{\infty} \frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{(2k+1)^3} \sum_{l=0}^{\infty} \left(\frac{\sin\left(\frac{(2l+1)y_0\pi}{\sqrt{2}L}\right)}{2l+1} - \frac{(2l+1)\sin\left(\frac{(2l+1)y_0\pi}{\sqrt{2}L}\right)}{(2k+1)^2 + (2l+1)^2} \right) \\
&= \frac{8L^2}{D\pi^3} \sum_{k=0}^{\infty} \frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{(2k+1)^3} \left(1 - \frac{\sinh\left(\frac{(2k+1)y_0\pi}{\sqrt{2}L}\right) + \sinh\left(\frac{(2k+1)(\sqrt{2}L-y_0)\pi}{\sqrt{2}L}\right)}{\sinh\left((2k+1)\pi\right)} \right) \\
&= \frac{x_0}{2D}\left(\sqrt{2}L - x_0\right) \\
&\qquad - \frac{8L^2}{D\pi^3} \sum_{k=0}^{\infty} \frac{\sin\left(\frac{(2k+1)x_0\pi}{\sqrt{2}L}\right)}{(2k+1)^3} \frac{\sinh\left(\frac{(2k+1)y_0\pi}{\sqrt{2}L}\right) + \sinh\left(\frac{(2k+1)(\sqrt{2}L-y_0)\pi}{\sqrt{2}L}\right)}{\sinh\left((2k+1)\pi\right)}
\end{aligned} \tag{187}
$$

List of publications

This research work has led to several publications, that are listed here. Publications 1 and 2 were related to my Master thesis, but are extensively used here. Publication 11 is not directly related to this thesis, even if it uses the same global formalism for an application in chemistry.

1. S. Condamin, O. Bénichou, V. Tejedor, R. Voituriez & J. Klafter, *First-passage times in complex scale-invariant media*", Nature, **450** (2007), 77–80.

2. S. Condamin,V. Tejedor & O. Bénichou, *Occupation times of random walks in confined geometries: From random trap model to diffusion-limited reactions*, Physical Review E, **76** (2007), 050102, R.

3. S. Condamin, V. Tejedor, R. Voituriez, O. Bénichou & J. Klafter, *Probing microscopic origins of confined subdiffusion by first-passage observables*, PNAS, **105** (2008), 5675-5680.

4. O. Bénichou, B. Meyer, V. Tejedor & R. Voituriez, *Zero Constant Formula for First-Passage Observables in Bounded Domains*, Physical Review Letters, **101** (2008), 130601.

5. R. Metzler, V. Tejedor, J.-H. Jeon, Y. He, W.H. Deng, S. Burov & E. Barkai, *Analysis of Single Particle Trajectories: From Normal to Anomalous Diffusion*, Acta Physica Polonica B, **40** (2009), 1315-1331.

6. V. Tejedor, O. Bénichou & R. Voituriez,*Global mean first-passage times of random walks on complex networks*, Physical Review E, **80** (2009), 065104(R).

7. V. Tejedor & R. Metzler, *Anomalous diffusion in correlated continuous time random walks*, Journal of Physics A-Mathematical and Theoretical, **43** (2010), 082002.

8. V. Tejedor, O. Bénichou, R. Voituriez, R. Jungmann, F. Simmel, C. Selhuber-Unkel, L.B. Oddershede & R. Metzler,*Quantitative analysis of single particle trajectories: mean maximal excursion method*, Biophysical Journal, **98** (2010), 1364-1372.

9. V. Tejedor, O. Bénichou, R. Voituriez & M. Moreau, *Response to targeted perturbations for random walks on networks*, Physical Review E, **82** (2010), 056106.

10. J-H. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unkel, K. Berg-Sørensen, L. Oddershede & R. Metzler, *In Vivo Anomalous Diffusion and Weak Ergodicity Breaking of Lipid Granules*, Physical Review Letters, **106** (2011), 048103.

11. V. Tejedor, *Réaction prépondérante : justification mathématique*, Bulletin de l'Union des Physiciens, **105** (2011), 157-174.

12. V. Tejedor, O. Bénichou, R. Metzler & R. Voituriez, *Residual mean first-passage time for jump processes: theory and applications to Levy flights and fractional Brownian motion*, Journal of Physics A-Mathematical and Theoretical, **44** (2011), 255003.

13. V. Tejedor, O. Bénichou, R. Voituriez, *Close or connected: Distance and connectivity effects on transport in networks*, Physical Review E, **83** (2011), 066102.

14. V. Tejedor, M. Schad, O. Bénichou, R. Voituriez & R. Metzler, *Encounter distribution of two random walkers on a finite one-dimensional interval*, Journal of Physics A-Mathematical and Theoretical, **44** (2011), 395005.

15. V. Tejedor, R. Voituriez, O. Benichou, *Optimizing persistent random searches*, Physical Review Letters, **108** (2012), 088103.

Bibliography

[1] Elena Agliari, *Exact mean first-passage time on the t-graph*, Physical Review E **77** (2008), no. 1, 011128–6.

[2] Elena Agliari and Raffaella Burioni, *Random walks on deterministic scale-free networks: Exact results*, Physical Review E **80** (2009), 031125.

[3] Reka Albert and Albert-Laszlo Barabasi, *Statistical mechanics of complex networks*, Reviews of Modern Physics **74** (2002), no. 1, 47 – 97.

[4] Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi, *Internet: Diameter of the world-wide web*, Nature **401** (1999), no. 6749, 130–131.

[5] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter, *Molecular biology of the cell*, Garland New York, 2002.

[6] David Aldous and James A. Fill, *Reversible markov chains and random walks on graphs*, Monograph in preparation, 1999.

[7] S. Alexander and Raymond L. Orbach, *Density of states on fractals: "fractons"*, Journal de Physique Lettres **43** (1982), no. 17, 625–631.

[8] F. Amblard, A. C. Maggs, B. Yurke, A. N. Pargellis, and S. Leibler, *Subdiffusion and anomalous local viscoelasticity in actin networks*, Phys. Rev. Lett. **77** (1996), 4470–4473.

[9] Daniel S. Banks and Cécile Fradin, *Anomalous diffusion of proteins due to molecular crowding*, Biophys. J. **89** (2005), no. 5, 2960–2971.

[10] Arie Bar-Haim and Joseph Klafter, *On mean residence and first passage times in finite one-dimensional systems*, The Journal of Chemical Physics **109** (1998), no. 13, 5187–5193.

[11] Eli Barkai and Yuan-Chung Cheng, *Aging continuous time random walks*, The Journal of Chemical Physics **118** (2003), no. 14, 6167–6178.

[12] Andrea Baronchelli, Michele Catanzaro, and Romualdo Pastor-Satorras, *Random walks on complex trees*, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **78** (2008), no. 1, 011114–9.

[13] A. Barrat, M. Barthélémy, and A. Vespignani, *Dynamical processes in complex networks*, Cambridge University Press, New-York, 2008.

[14] G. Barton, *Elements of green's functions and propagation*, Oxford Science Publications, 1989.

[15] F Bartumeus, M G E Da Luz, G M Viswanathan, and J Catalan, *Animal search strategies: A quantitative. random-walk analysis*, Ecology **86** (2005), no. 11, 3078–3087.

[16] G. Bel and E. Barkai, *Weak ergodicity breaking in the continuous-time random walk*, Physical Review Letters **94** (2005), no. 24, 240602–4.

[17] J. W. Bell, *Searching behaviour, the behavioural ecology of finding resources, animal behaviour series*, Chapman and Hall, London, 1991.

[18] S Benhamou, *How many animals really do the Levy walk?*, Ecology **88** (2007), 1962–1969.

[19] O. Benichou, M. Coppey, J. Klafter, M. Moreau, and G. Oshanin, *On the joint residence time of n independent two-dimensional brownian motions*, Journal of Physics A: Mathematical and General **36** (2003), no. 26, 7225–7231.

[20] O Benichou, M Coppey, M Moreau, and G Oshanin, *Kinetics of diffusion-limited catalytically activated reactions: an extension of the wilemski-fixman approach.*, J Chem Phys **123** (2005), no. 19, 194506 (eng).

[21] O. Bénichou, M. Coppey, M. Moreau, P. H. Suet, and R. Voituriez, *Averaged residence times of stochastic motions in bounded domains*, EPL (Europhysics Letters) **70** (2005), no. 1, 42–48.

[22] O Benichou, M Coppey, M Moreau, P-H Suet, and R Voituriez, *Optimal search strategies for hidden targets.*, Physical Review Letters **94** (2005), no. 19, 198101 (eng).

[23] O. Bénichou, M. Coppey, M. Moreau, and R. Voituriez, *Intermittent search strategies: When losing time becomes efficient*, EPL (Europhysics Letters) **75** (2006), no. 2, 349–354.

[24] O Benichou, C Loverdo, M Moreau, and R Voituriez, *Two-dimensional intermittent search processes: An alternative to levy flight strategies.*, Phys Rev E Stat Nonlin Soft Matter Phys **74** (2006), no. 2 Pt 1, 020102 (eng).

[25] O. Benichou, C. Loverdo, M. Moreau, and R. Voituriez, *Optimizing intermittent reaction paths*, Phys. Chem. Chem. Phys. **10** (2008), –.

[26] O Bénichou, C Loverdo, M Moreau, and R Voituriez, *Optimizing intermittent reaction paths*, Physical Chemistry Chemical Physics **10** (2008), 7059–7072.

[27] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, *Intermittent search strategies*, Rev. Mod. Phys. **83** (2011), 81–129.

[28] O. Benichou, B. Meyer, V. Tejedor, and R. Voituriez, *Zero constant formula for first-passage observables in bounded domains*, Physical Review Letters **101** (2008), no. 13, 130601–4.

[29] O. Benichou and R. Voituriez, *Narrow-escape time problem: Time needed for a particle to exit a confining domain through a small window*, Physical Review Letters **100** (2008), no. 16, 168105–4.

[30] Bénichou, O. , Loverdo, C. , and Voituriez, R. , *How gene colocalization can be optimized by tuning the diffusion constant of transcription factors*, EPL **84** (2008), no. 3, 38003.

[31] A. M. Berezhkovskii and A. V. Barzykin, *Extended narrow escape problem: Boundary homogenization-based analysis*, Phys. Rev. E **82** (2010), 011114.

[32] H C Berg, *E.coli in motion*, Springer, New York, 2004.

[33] O G Berg, R B Winter, and P H Von Hippel, *Diffusion-driven mechanisms of protein translocation on nucleid acids. 1. models an theory*, Biochemistry **20** (1981), 6929–6948.

[34] Roger Bidaux, Jérôme Chave, and Radim Vocka, *Finite time and asymptotic behaviour of the maximal excursion of a random walk*, Journal of Physics A: Mathematical and General **32** (1999), no. 27, 5009.

[35] S. Blanco and R. Fournier, *An invariance property of diffusive random walks*, EPL (Europhysics Letters) **61** (2003), no. 2, 168–173.

[36] Stephane Blanco and Richard Fournier, *Short-path statistics and the diffusion approximation*, Physical Review Letters **97** (2006), no. 23, 230604–4.

[37] A. Blumen, G. Zumofen, and J. Klafter, *Target annihilation by random walkers*, Phys. Rev. B **30** (1984), 5379.

[38] Erik M. Bollt and Daniel ben Avraham, *What is special about diffusion on scale-free nets?*, New Journal of Physics **7** (2005), 26–26.

[39] I Bonnet, A Biebricher, P-L Porté, C Loverdo, O Bénichou, R Voituriez, C Escudé, W Wende, A Pingoud, and P Desbiolles, *Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA*, Nucleid Acids Research **36** (2008), 4118–27.

[40] J. P. Bouchaud, *Weak ergodicity breaking and aging in disordered systems*, Journal de Physique I France **2** (1992), 1705 – 1713.

[41] Jean-Philippe Bouchaud and Antoine Georges, *Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications*, Physics Reports **195** (1990), no. 4-5, 127–293.

[42] Robert Brown, *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.*, Phil. Mag. **4** (1828), 161–173.

[43] M. J. A. M. Brummelhuis, *Applications of random walks : Tracer diffusion, lattice covering, and damage spreading*, Ph.D. thesis, Leiden University, 1991.

[44] Sergey V. Buldyrev, Roni Parshani, Gerald Paul, H. Eugene Stanley, and Shlomo Havlin, *Catastrophic cascade of failures in interdependent networks*, Nature **464** (2010), no. 7291, 1025–1028.

[45] A. Bunde and S. Havlin, *Fractals and disordered systems*, Springer, Berlin, 1991.

[46] R. Burioni and D. Cassi, *Random walks on graphs: ideas, techniques and results*, Journal of Physics A: Mathematical and General **38** (2005), no. 8, R45–R78.

[47] S. Burov and E. Barkai, *Occupation time statistics in the quenched trap model*, Physical Review Letters **98** (2007), no. 25, 250601–4.

[48] ———, *Critical exponent of the fractional langevin equation*, Phys. Rev. Lett. **100** (2008), 070601.

[49] ———, *Fractional langevin equation: Overdamped, underdamped, and critical behaviors*, Phys. Rev. E **78** (2008), 031112.

[50] S. Burov, R. Metzler, and E. Barkai, *Aging and nonergodicity beyond the khinchin theorem*, Proceedings of the National Academy of Sciences **107** (2010), no. 30, 13228–13233.

[51] Stas Burov, Jae-Hyung Jeon, Ralf Metzler, and Eli Barkai, *Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking*, Phys. Chem. Chem. Phys. **13** (2011), –.

[52] I Calvo and R Sánchez, *The path integral formulation of fractional brownian motion for the general hurst exponent*, Journal of Physics A: Mathematical and Theoretical **41** (2008), no. 28, 282002.

[53] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, *Critical points and transitions in an electric power transmission model for cascading failure blackouts*, Chaos **12** (2002), no. 4, 985–994.

[54] Avi Caspi, Rony Granek, and Michael Elbaum, *Enhanced diffusion in active intracellular transport*, Phys. Rev. Lett. **85** (2000), 5655–5658.

[55] ———, *Diffusion and directed motion in cellular transport*, Phys. Rev. E **66** (2002), 011916.

[56] L Champagne, R G Carl, and R Hill, *Search theory, agent-based simulation, and u-boats in the bay of biscay*, Proceedings of The 2003 Winter Simulation Conference **1-2** (2003), 991–998.

[57] A. K. Chandra, P. Raghavan, W. L. Ruzzo, and R. Smolensky, *The electrical resistance of a graph captures its commute and cover times*, STOC '89: Proceedings of the twenty-first annual ACM symposium on Theory of computing, 1989.

[58] Thipphaya Cherdhirankorn, Markus Retsch, Ulrich Jonas, Hans-Juergen Butt, and Kaloian Koynov, *Tracer diffusion in silica inverse opals*, Langmuir **26** (2010), no. 12, 10141–10146, PMID: 20232884.

[59] C Chevalier, O Bénichou, B Meyer, and R Voituriez, *First-passage quantities of brownian motion in a bounded domain with multiple targets: a unified approach*, Journal of Physics A: Mathematical and Theoretical **44** (2011), no. 2, 025002.

[60] Chia-Fu Chou, Olgica Bakajin, Stephen W. P. Turner, Thomas A. J. Duke, Shirley S. Chan, Edward C. Cox, Harold G. Craighead, and Robert H. Austin, *Sorting by diffusion: An asymmetric obstacle course for continuous molecular separation*, Proceedings of the National Academy of Sciences **96** (1999), no. 24, 13762–13765.

[61] S Condamin and O Benichou, *Exact expressions of mean first-passage times and splitting probabilities for random walks in bounded rectangular domains.*, J Chem Phys **124** (2006), no. 20, 206103 (eng).

[62] S Condamin, O Benichou, and J Klafter, *First-passage time distributions for subdiffusion in confined geometry.*, Phys Rev Lett **98** (2007), no. 25, 250602 (eng).

[63] S Condamin, O Benichou, and M Moreau, *First-exit times and residence times for discrete random walks on finite lattices.*, Phys Rev E Stat Nonlin Soft Matter Phys **72** (2005), no. 1 Pt 2, 016127 (eng).

[64] ———, *First-passage times for random walks in bounded domains.*, Phys Rev Lett **95** (2005), no. 26, 260601 (eng).

[65] ———, *Random walks and brownian motion: a method of computation for first-passage times and related quantities in confined geometries.*, Phys Rev E Stat Nonlin Soft Matter Phys **75** (2007), no. 2 Pt 1, 021111 (eng).

[66] S Condamin, O Benichou, V Tejedor, R Voituriez, and J Klafter, *First-passage times in complex scale-invariant media.*, Nature **450** (2007), no. 7166, 77–80 (eng).

[67] S Condamin, V Tejedor, and O Benichou, *Occupation times of random walks in confined geometries: from random trap model to diffusion-limited reactions.*, Phys Rev E Stat Nonlin Soft Matter Phys **76** (2007), no. 5 Pt 1, 050102 (eng).

[68] S. Condamin, V. Tejedor, R. Voituriez, O. Benichou, and J. Klafter, *Probing microscopic origins of confined subdiffusion by first-passage observables*, Proceedings of the National Academy of Sciences (2008), 0712158105–.

[69] M. Coppey, O. Benichou, R. Voituriez, and M. Moreau, *Kinetics of target site localization of a protein on dna: A stochastic approach*, Biophys. J. **87** (2004), no. 3, 1640–1649.

[70] Francis Crick, *Diffusion in embryogenesis*, Nature **225** (1970), no. 5231, 420–422.

[71] D.Ben-Avraham and S.Havlin, *Diffusion and reactions in fractals and disordered systems*, Cambridge University Press, 2000.

[72] P. G. de Gennes, *Kinetics of diffusion-controlled processes in dense polymer systems. i. nonentangled regimes*, The Journal of Chemical Physics **76** (1982), no. 6, 3316–3321.

[73] Pierre Gilles de Gennes, *La percolation : un concept unificateur*, La Recherche **7** (1976), 919.

[74] Weihua Deng and Eli Barkai, *Ergodic properties of fractional brownian-langevin motion*, Phys. Rev. E **79** (2009), 011112.

[75] Nicolas Destainville, Aude Saulière, and Laurence Salomé, *Comment to the article by michael j. saxton: A biological interpretation of transient anomalous subdiffusion. i. qualitative model*, Biophysical Journal **97** (2008), no. 7, 3117 – 3119.

[76] Z Dezso and AL Barabasi, *Halting viruses in scale-free networks*, Physical Review E **65** (2002), no. 5, 055103 (English).

[77] Da-Qiao Ding, Yuki Tomita, Ayumu Yamamoto, Yuji Chikashige, Tokuko Haraguchi, and Yasushi Hiraoka, *Large-scale screening of intracellular protein localization in living fission yeast cells by the use of a gfp-fusion genomic dna library*, Genes to Cells **5** (2000), no. 3, 169–190.

[78] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Critical phenomena in complex networks*, Reviews of Modern Physics **80** (2008), no. 4, 1275–61.

[79] Peter G. Doyle and J. Laurie Snell, *Random walks and electric networks*, Carus Mathematical Monographs, vol. 22, Mathematical Association of America, Washington DC, USA, 1984.

[80] Olga K. Dudko, Alexander M. Berezhkovskii, and George H. Weiss, *Time-dependent diffusion coefficients in periodic porous materials*, The Journal of Physical Chemistry B **109** (2005), no. 45, 21296–21299, PMID: 16853761.

[81] Andrew M. Edwards, Richard A. Phillips, Nicholas W. Watkins, Mervyn P. Freeman, Eugene J. Murphy, Vsevolod Afanasyev, Sergey V. Buldyrev, M. G. E. da Luz, E. P. Raposo, H. Eugene Stanley, and Gandhimohan M. Viswanathan, *Revisiting levy flight search patterns of wandering albatrosses, bumblebees and deer*, Nature **449** (2007), no. 7165, 1044–1048.

[82] A. Einstein, *Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen*, Annalen der Physik **322** (1905), no. 8, 549–560.

[83] Johan Elf, Gene-Wei Li, and X. Sunney Xie, *Probing transcription factor dynamics at the single-molecule level in a living cell*, Science **316** (2007), no. 5828, 1191–1194.

[84] Iddo Eliazar, Tal Koren, and Joseph Klafter, *Searching circular dna strands*, Journal of Physics: Condensed Matter **19** (2007), no. 6, 065140.

[85] Matthieu H. Ernst, *Random walks with short memory*, Journal of Statistical Physics **53** (1988), 191–201, 10.1007/BF01011552.

[86] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos, *On power-law relationships of the internet topology*, SIGCOMM '99: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication (New York, NY, USA), ACM, 1999, pp. 251–262.

[87] M. Ferraro and L. Zaninetti, *Statistics of visits to sites in random walks*, Physica A **338** (2004), no. 3 – 4, 307 – 318.

[88] Adolf Fick, *Ueber diffusion*, Annalen der Physik **170** (1855), no. 1, 59–86.

[89] Joseph Fourier, *Théorie analytique de la chaleur*, Firmin Didot, père et fils, Paris, 1822.

[90] Benjamin M. Friedrich, *Search along persistent random walks*, Physical Biology **5** (2008), no. 2, 026007.

[91] Agata Fronczak and Piotr Fronczak, *Biased random walks in complex networks: The role of local navigation rules*, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **80** (2009), no. 1, 016107–6.

[92] J R Frost and Lawrence D Stone, *Review of search theory: Advances and applications to search and rescue decision support*, Tech. report, United States Coast Guard, 2100 Second Street, SW, Washington, DC 20593 USA, 2001.

[93] Lazaros K. Gallos, Chaoming Song, Shlomo Havlin, and Hernan A. Makse, *Scaling theory of transport in complex biological networks*, Proceedings of the National Academy of Sciences **104** (2007), no. 19, 7746–7751.

[94] R.N. Ghosh and W.W. Webb, *Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules*, Biophysical Journal **66** (1994), no. 5, 1301 – 1318.

[95] Thomas Gilbert, Huu Chuong Nguyen, and David P Sanders, *Diffusive properties of persistent walks on cubic lattices with application to periodic lorentz gases*, Journal of Physics A: Mathematical and Theoretical **44** (2011), no. 6, 065001.

[96] J. Gillis, *Correlated random walk*, Mathematical Proceedings of the Cambridge Philosophical Society **51** (1955), no. 04, 639–651.

[97] C. Godrèche and J.M. Luck, *Statistics of the occupation time of renewal processes*, J. Stat. Phys. **104** (2001), 489.

[98] L. Le Goff, F. Amblard, and E.M. Furst, *Motor-driven dynamics in actin-myosin networks*, Phys. Rev. Lett. **88** (2002), no. 1, 018101.

[99] E. Golding and E.C. Cox, *Physical nature of bacterial cytoplasm*, Phys. Rev. Lett. **96** (2006), 098102.

[100] J Gorman and E C Greene, *Visualizing one-dimensional diffusion of proteins along DNA*, Nature structural and molecular biology **15** (2008), 768–774.

[101] Jing-Dong J. Han, Nicolas Bertin, Tong Hao, Debra S. Goldberg, Gabriel F. Berriz, Lan V. Zhang, Denis Dupuy, Albertha J. M. Walhout, Michael E. Cusick, Frederick P. Roth, and Marc Vidal, *Evidence for dynamically organized modularity in the yeast protein-protein interaction network*, Nature **430** (2004), no. 6995, 88–93.

[102] S. Havlin and D. ben Avraham, *Diffusion in disordered media*, Adv.in Phys. **36** (1987), no. 6, 695.

[103] C. P. Haynes and A. P. Roberts, *Global first-passage times of fractal lattices*, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **78** (2008), no. 4, 041111–9.

[104] Y. He, S. Burov, R. Metzler, and E. Barkai, *Random time-scale invariant diffusion and transport coefficients*, Physical Review Letters **101** (2008), no. 5, 058101–4.

[105] J. Hohlbein, M. Steinhart, C. Schiene-Fischer, A. Benda, M. Hof, and ChristianG. Hübner, *Confined diffusion in ordered nanoporous alumina membranes*, Small **3** (2007), no. 3, 380–385.

[106] D Holcman and I Kupka, *The probability of an encounter of two brownian particles before escape*, Journal of Physics A: Mathematical and Theoretical **42** (2009), no. 31, 315210.

[107] D. Holcman, A. Marchewka, and Z. Schuss, *Survival probability of diffusion with trapping in cellular neurobiology*, Phys.Rev.E **72** (2005), 31910.

[108] D. Holcman and Z. Schuss, *Escape through a small opening: Receptor trafficking in a synaptic membrane*, Journal of Statistical Physics **117** (2004), no. 5, 975–1014.

[109] J. R. M. Hosking, *Modeling persistence in hydrological time series using fractional differencing*, Water Resour. Res. **20** (1984), no. 12, 1898–1908.

[110] Z. G. Huang, X. J. Xu, Z. X. Wu, and Y. H. Wang, *Walks on apollonian networks*, Eur. Phys. J. B **51** (2006), no. 4, 549–553.

[111] B.D. Hughes, *Random walks and random environments*, Oxford University Press, New-York, 1995.

[112] Alex James, Michael J. Plank, and Andrew M. Edwards, *Assessing lévy walks as models of animal foraging*, Journal of The Royal Society Interface **8** (2011), no. 62, 1233–1247.

[113] Jae-Hyung Jeon, Vincent Tejedor, Stas Burov, Eli Barkai, Christine Selhuber-Unkel, Kirstine Berg-Sorensen, Lene Oddershede, and Ralf Metzler, *In vivo anomalous diffusion and weak ergodicity breaking of lipid granules*, Physical Review Letters **106** (2011), 048103.

[114] Aldo Jesorka, Natalia Stepanyants, Haijiang Zhang, Bahanur Ortmen, Bodil Hakonen, and Owe Orwar, *Generation of phospholipid vesicle-nanotube networks and transport of molecules therein*, Nature Protocols **6** (2011), no. 6, 7910–805.

[115] J.Klafter, A.Blumen, and M.F.Shlesinger, *Stochastic pathway to anomalous diffusion*, Phys.Rev.A **35** (1987), 3081–3085.

[116] Yacov Kantor and Mehran Kardar, *Anomalous diffusion with absorbing boundary*, Phys. Rev. E **76** (2007), 061121.

[117] HJ Kim, IM Kim, Y Lee, and B Kahng, *Scale-free network in stock markets*, Journal of the Korean Physical Society **40** (2002), no. 6, 1105–1108 (English), 7th International Workshop on Similarity in Diversity, TOKYO, JAPAN, SEP 10-12, 2001.

[118] Maksim Kitsak, Lazaros K. Gallos, Shlomo Havlin, Fredrik Liljeros, Lev Muchnik, H. Eugene Stanley, and Hernan A. Makse, *Identification of influential spreaders in complex networks*, Nat Phys **6** (2010), 888—893.

[119] A. Kittas, S. Carmi, S. Havlin, and P. Argyrakis, *Trapping in complex networks*, EPL **84** (2008), no. 4, 40008.

[120] J. Klafter, M. Shlesinger, and G. Zumofen, *Beyond brownian motion*, Phys. Today **49** (1996), no. 2, 33.

[121] Grigory Kolesov, Zeba Wunderlich, Olga N. Laikova, Mikhail S. Gelfand, and Leonid A. Mirny, *How gene order is influenced by the biophysics of transcription regulation*, Proceedings of the National Academy of Sciences **104** (2007), no. 35, 13948–13953.

[122] R. Kopelman, P. W. Klymko, J. S. Newhouse, and L. W. Anacker, *Reaction kinetics on fractals: Random-walker simulations and excition experiments*, Phys. Rev. B **29** (1984), 3747–3748.

[123] Tal Koren, Michael A. Lomholt, Aleksei V. Chechkin, Joseph Klafter, and Ralf Metzler, *Leapover lengths and first passage time statistics for lévy flights*, Phys. Rev. Lett. **99** (2007), no. 16, 160602.

[124] John J. Kozak and V. Balakrishnan, *Analytic expression for the mean time to absorption for a random walker on the sierpinski gasket*, Physical Review E **65** (2002), no. 2, 021105.

[125] Balazs Kozma, Matthew B. Hastings, and G. Korniss, *Diffusion processes on power-law small-world networks*, Physical Review Letters **95** (2005), no. 1, 018701–4.

[126] D L Kramer and R L McLaughlin, *The behavioral ecology of intermittent locomotion*, American Zoologist **41** (2001), 137–153.

[127] A. Kusumi, Y. Sako, and M. Yamamoto, *Confined lateral diffusion of membrane-receptors as studied by single-particle tracking (nanovid microscopy) - effects of calcium-induced differentiation in cultured epithelial-cells*, Biophysical Journal **65** (1993), no. 5, 2021–2040.

[128] L. Lizana and T. Ambjörnsson, *Single-file diffusion in a box*, Phys. Rev. Lett. **100** (2008), 200601.

[129] A. L. Lloyd and R. M. May, *Epidemiology - how viruses spread among computers and people*, Science **292** (2001), no. 5520, 1316–1317.

[130] Michael A. Lomholt, Koren Tal, Ralf Metzler, and Klafter Joseph, *Lévy strategies in intermittent search processes are advantageous*, Proceedings of the National Academy of Sciences **105** (2008), no. 32, 11055–11059.

[131] Michael A. Lomholt, Irwin M. Zaid, and Ralf Metzler, *Subdiffusion and weak ergodicity breaking in the presence of a reactive boundary*, Physical Review Letters **98** (2007), no. 20, 200603–4.

[132] Eduardo López, Sergey V. Buldyrev, Shlomo Havlin, and H. Eugene Stanley, *Anomalous transport in scale-free networks*, Phys. Rev. Lett. **94** (2005), 248701.

[133] C. Loverdo, O. Benichou, M. Moreau, and R. Voituriez, *Enhanced reaction kinetics in biological cells*, Nature Physics **4** (2008), no. 2, 134–137.

[134] C. Loverdo, O. Benichou, R. Voituriez, A. Biebricher, I. Bonnet, and P. Desbiolles, *Quantifying hopping and jumping in facilitated diffusion of dna-binding proteins*, Physical Review Letters **102** (2009), no. 18, 188101–4.

[135] Ariel Lubelski, Igor M. Sokolov, and Joseph Klafter, *Nonergodicity mimics inhomogeneity in single particle tracking*, Physical Review Letters **100** (2008), no. 25, 250602–4.

[136] S.M. Majumdar and A. Comtet, *Local and occupation time of a particle diffusing in a random medium*, Phys. Rev. Lett. **89** (2002), 60601.

[137] Benoit B. Mandelbrot and John W. Van Ness, *Fractional brownian motions, fractional noises and applications*, SIAM Review **10** (1968), no. 4, 422–437.

[138] James Clerk Maxwell, *A treatise on electricity and magnetism*, vol. I, Clarendon Press, Oxford, 1873 (English).

[139] A. Mazzolo, *Properties of diffusive random walks in bounded domains*, EPL (Europhysics Letters) **68** (2004), no. 3, 350–355.

[140] Y Meroz, I Eliazar, and J Klafter, *Facilitated diffusion in a crowded environment : from kinetics to stochastics*, Journal of Physics A - Mathematical and Theoretical **42** (2009), 434012.

[141] Minas M. Mezedur, Massoud Kaviany, and Wayne Moore, *Effect of pore structure, randomness and size on effective mass diffusivity*, AIChE Journal **48** (2002), no. 1, 15–24.

[142] Elliott W. Montroll, *Random walks on lattices. iii. calculation of first-passage times with application to exciton trapping on photosynthetic units*, Journal of Mathematical Physics **10** (1969), no. 4, 753–765.

[143] M Moreau, G Oshanin, O Benichou, and M Coppey, *Pascal principle for diffusion-controlled trapping reactions.*, Phys Rev E Stat Nonlin Soft Matter Phys **67** (2003), no. 4 Pt 2, 045104 (eng).

[144] André A. Moreira, José S. Andrade, Hans J. Herrmann, and Joseph O. Indekeu, *How to make a fragile network robust and vice versa*, Physical Review Letters **102** (2009), no. 1, 018701.

[145] Deqiang Mu, Zhong-Sheng Liu, Cheng Huang, and Ned Djilali, *Determination of the effective diffusion coefficient in porous media including knudsen effects*, Microfluidics and Nanofluidics **4** (2008), 257–260, 10.1007/s10404-007-0182-3.

[146] Thomas Neusius, Igor M. Sokolov, and Jeremy C. Smith, *Subdiffusion in time-averaged, confined random walks*, Phys. Rev. E **80** (2009), 011109.

[147] M. E. J. Newman, *Assortative mixing in networks*, Phys. Rev. Lett. **89** (2002), no. 20, 208701.

[148] Jae Dong Noh and Heiko Rieger, *Random walks on complex networks*, Physical Review Letters **92** (2004), no. 11, 118701–4.

[149] I. Nordlund, *Eine neue bestimmung der avogadroschen konstante aus der brownschen bewegung kleiner, in wasser suspendierten quecksilberkügelchen*, Zeitschrift für Physikalische Chemie **87** (1914), 40–62.

[150] Dmytro Nykypanchuk, Helmut H. Strey, and David A. Hoagland, *Brownian motion of dna confined within a two-dimensional array*, Science **297** (2002), no. 5583, 987–990.

[151] Benichou O., Chevalier C., Klafter J., Meyer B., and Voituriez R., *Geometry-controlled kinetics*, Nature Chemistry **2** (2010), no. 6, 472–477.

[152] W J O'Brien, H I Browman, and B I Evans, *Search strategies of foraging animals*, American Scientist **78** (1990), 152–160.

[153] G. Oshanin, H. S. Wio, K. Lindenberg, and S. F. Burlatsky, *Intermittent random walks for an optimal search strategy: one-dimensional case*, Journal of Physics: Condensed Matter **19** (2007), no. 6, 065142.

[154] Ben O'Shaughnessy and Itamar Procaccia, *Analytical solutions for diffusion on fractal objects*, Physical Review Letters **54** (1985), no. 5, 455 – 458.

[155] R Pastor-Satorras and A Vespignani, *Epidemic spreading in scale-free networks*, Physical Review Letters **86** (2001), no. 14, 3200–3203.

[156] K Pearson, *The problem of the random walk*, Nature **72** (1905), 294–294.

[157] Jean Perrin, *Les atomes*, Felix Alcan, 1913.

[158] S. Pillay, M. J. Ward, A. Peirce, and T. Kolokolnikov, *An asymptotic analysis of the mean first passage time for narrow escape problems: Part i: Two-dimensional domains*, Multiscale Modeling & Simulation **8** (2010), no. 3, 803–835.

[159] Inti Pineda, Marco-Vinicio Vazquez, Alexander M. Berezhkovskii, and Leonardo Dagdug, *Diffusion in periodic two-dimensional channels formed by overlapping circles: Comparison of analytical and numerical results*, The Journal of Chemical Physics **135** (2011), no. 22, 224101.

[160] Melpomeni Platani, Ilya Goldberg, Angus I. Lamond, and Jason R. Swedlow, *Cajal body dynamics and association with chromatin are atp-dependent*, Nat Cell Biol **4** (2002), no. 7, 502–508.

[161] Georg Pólya, *über eine aufgabe der wahrscheinlichkeitsrechnung betreffend die irrfahrt im straßennetz*, Mathematische Annalen **83** (1921), 149–160.

[162] Riccardo Raccis, Arash Nikoubashman, Markus Retsch, Ulrich Jonas, Kaloian Koynov, Hans-Jürgen Butt, Christos N. Likos, and George Fytas, *Confined diffusion in periodic porous nanostructures*, ACS Nano **5** (2011), no. 6, 4607–4616.

[163] Subramanian P. Ramanathan, Kara van Aelst, Alice Sears, Luke J. Peakman, Fiona M. Diffin, Mark D. Szczelkun, and Ralf Seidel, *Type iii restriction enzymes communicate in 1d without looping between their target sites*, Proceedings of the National Academy of Sciences **106** (2009), no. 6, 1748–1753.

[164] W S Rasband, *ImageJ*, U. S. National Institutes of Health, Bethesda, Maryland, USA (1997-2009), http://rsb.info.nih.gov/ij/.

[165] M.D. Rawn, *On the summation of fourier and bessel series*, Journal of Mathematical Analysis and Applications **193** (1995), no. 1, 282 – 295.

[166] Lord Rayleigh, *The problem of the random walk*, Nature **72** (1905), 318–318.

[167] A. Rebenshtok and E. Barkai, *Distribution of time-averaged observables for weak ergodicity breaking*, Physical Review Letters **99** (2007), no. 21, 210601–4.

[168] S. Redner, *A guide to first passage time processes*, Cambridge University Press, Cambridge, England, 2001.

[169] S. A. Rice, *Diffusion-limited reactions*, vol. 25, Elsevier, Amsterdam, 1985.

[170] A. C. Richardson, S. N. S. Reihani, and L. B. Oddershede, *Non-harmonic potential of a single beam optical trap*, Optics Express **16** (2008), 15709 – 15717.

[171] H R Richardson and L D Stone, *Operations analysis during the underwater search for Scorpion*, Naval Research Logistics Quarterly **18** (1971), 141 – 157.

[172] R.Metzler and J.Klafter, *The random walk's guide to anomalous diffusion: a fractionnal dynamics approach*, Physics Reports **339** (2000), no. 1, 1–77.

[173] ———, *The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics*, J.Phys.A **37** (2004), R161–R208.

[174] C.F. Robinow and J. S. Hyams, *Molecular biology of the fission yeast*, Academic Press, New York, 1989.

[175] Salman S. Rogers, Thomas A. Waigh, and Jian R. Lu, *Intracellular microrheology of motile amoeba proteus*, Biophysical Journal **94** (2008), 3313 – 3322.

[176] Salman S Rogers, Thomas A Waigh, Xiubo Zhao, and Jian R Lu, *Precise particle tracking against a complicated background: polynomial fitting with gaussian weight*, Physical Biology **4** (2007), no. 3, 220.

[177] Hernán D. Rozenfeld, Shlomo Havlin, and Daniel ben Avraham, *Fractal and transfractal recursive scale-free nets*, New Journal of Physics **9** (2007), no. 6, 175–175.

[178] A. N. Samukhin, S. N. Dorogovtsev, and J. F. F. Mendes, *Laplacian spectra of, and random walks on, complex networks: Are scale-free architectures really important?*, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **77** (2008), no. 3, 036115–19.

[179] L.P. Savtchenko and D.A. Rusakov, *The optimal height of the synaptic cleft*, PNAS **104** (2007), 1823.

[180] M. J. Saxton, *Anomalous diffusion due to obstacles: a monte carlo study*, Biophysical Journal **66** (1994), no. 2, 394 – 401.

[181] ———, *Anomalous diffusion due to binding: a monte carlo study*, Biophysical Journal **70** (1996), no. 1, 1250 – 1262.

[182] Michael J. Saxton, *A biological interpretation of transient anomalous subdiffusion. i. qualitative model*, Biophysical Journal **92** (2007), no. 4, 1178–1191.

[183] ———, *A biological interpretation of transient anomalous subdiffusion. ii. reaction kinetics*, Biophysical Journal **94** (2008), no. 3, 760–771.

[184] I.F. Sbalzarini and P. Koumoutsakos, *Feature point tracking and trajectory analysis for video imaging in cell biology*, Journal of Structural Biology **151** (2005), no. 2, 182 – 195.

[185] S.B.Yuste and K.Lindenberg, *Subdiffusion-limited reactions*, Chem.Phys. **284** (2002), no. 1-2, 169–180.

[186] H. Scher, G. Margolin, R. Metzler, J. Klafter, and B. Berkowitz, *The dynamical foundation of fractal stream chemistry: The origin of extremely long retention times*, Geophys.Res.Lett. **29** (2002), no. 1061.

[187] H. Scher and E.W. Montroll, *Anomalous transit-time dispersion in amorphous solids*, Physical Review B **12** (1975), 2455.

[188] Z. Schuss, A. Singer, and D. Holcman, *The narrow escape problem for diffusion in cellular microdomains*, Proceedings of the National Academy of Sciences **104** (2007), no. 41, 16098–16103.

[189] K L Sebastian, *Path integral representation for fractional brownian motion*, Journal of Physics A: Mathematical and General **28** (1995), no. 15, 4305.

[190] Georg Seisenberger, Martin U. Ried, Thomas Endreß, Hildegard Büning, Michael Hallek, and Christoph Bräuchle, *Real-time single-molecule imaging of the infection pathway of an adeno-associated virus*, Science **294** (2001), no. 5548, 1929–1932.

[191] Christine Selhuber-Unkel, Pernille Yde, Kirstine Berg-Sørensen, and Lene B Oddershede, *Variety in intracellular diffusion during the cell cycle*, Physical Biology **6** (2009), no. 2, 025015.

[192] Michael F Shlesinger, *Mathematical physics: first encounters.*, Nature **450** (2007), no. 7166, 40–41 (eng).

[193] Michael F. Shlesinger, *Random searching*, Journal of Physics A: Mathematical and Theoretical **42** (2009), 434001.

[194] Ludovico Silvestri, Leone Fronzoni, Angelo Gemignani, Paolo Grigolini, Danilo Menicucci, and Paolo Allegrini, *Event-driven power-law relaxation in weak turbulence: A liquid crystals mesoscopic experiment bridging quantum dots and the integration theory for the brain*, Journal of Physics: Conference Series **174** (2009), no. 1, 012070.

[195] Ingve Simonsen, Lubos Buzna, Karsten Peters, Stefan Bornholdt, and Dirk Helbing, *Transient dynamics increasing network vulnerability to cascading failures*, Physical Review Letters **100** (2008), no. 21, 218701–4.

[196] A. Singer, Z. Schuss, and D. Holcman, *Narrow escape, part ii: The circular disk*, Journal of Statistical Physics **122** (2006), no. 3, 465–489.

[197] A. Singer, Z. Schuss, D. Holcman, and R. Eisenberg, *Narrow escape, part i*, Journal of Statistical Physics **122** (2006), no. 3, 437–463.

[198] Michael Slutsky and Leonid A. Mirny, *Kinetics of protein-dna interaction: Facilitated target location in sequence-dependent potential*, Biophysical Journal **87** (2004), no. 6, 4021–4035.

[199] Patricia R. Smith, Ian E.G. Morrison, Keith M. Wilson, Nelson Fernández, and Richard J. Cherry, *Anomalous diffusion of major histocompatibility complex class i molecules on hela cells determined by single particle tracking*, Biophysical Journal **76** (1999), no. 6, 3331 – 3344.

[200] Igor M. Sokolov, Ralf Metzler, Kiran Pant, and Mark C. Williams, *Target search of n sliding proteins on a dna*, Biophysical Journal **89** (2005), no. 2, 895–902.

[201] I.M. Sokolov and J. Klafter, *Field-induced dispersion in subdiffusion*, Phys. Rev. Lett. **97** (2006), 140602.

[202] C. Song, S. Havlin, and H.A. Makse, *Origins of fractality in the growth of complex networks*, Nature Physics **2** (2006), 275–281.

[203] Chaoming Song, Lazaros K. Gallos, Shlomo Havlin, and Hernán A. Makse, *How to calculate the fractal dimension of a complex network: the box covering algorithm*, Journal of Statistical Mechanics: Theory and Experiment **2007** (2007), no. 03, P03006–P03006.

[204] Chaoming Song, Shlomo Havlin, and Hernan A. Makse, *Self-similarity of complex networks*, Nature **433** (2005), no. 7024, 392–395.

[205] _____, *Origins of fractality in the growth of complex networks*, Nat Phys **2** (2006), no. 4, 275–281.

[206] V. Sood, S. Redner, and D. ben Avraham, *First-passage properties of the erdös–rényi random graph*, Journal of Physics A: Mathematical and General **38** (2005), no. 1, 109–123.

[207] Jedrzej Szymanski and Matthias Weiss, *Elucidating the origin of anomalous diffusion in crowded fluids*, Physical Review Letters **103** (2009), no. 3, 038102–4.

[208] V Tejedor, O Bénichou, Ralf Metzler, and R Voituriez, *Residual mean first-passage time for jump processes: theory and applications to lévy flights and fractional brownian motion*, Journal of Physics A: Mathematical and Theoretical **44** (2011), no. 25, 255003.

[209] V. Tejedor, O. Bénichou, and R. Voituriez, *Global mean first-passage times of random walks on complex networks*, Physical Review E **80** (2009), no. 6, 065104(R).

[210] Vincent Tejedor, Olivier Bénichou, Raphael Voituriez, Ralf Jungmann, Friedrich Simmel, Christine Selhuber-Unkel, Lene B. Oddershede, and Ralf Metzler, *Quantitative analysis of single particle trajectories: Mean maximal excursion method*, Biophysical Journal **98** (2010), no. 7, 1364–1372.

[211] Liang Tian and Da-Ning Shi, *Scaling of disordered recursive scale-free networks*, EPL **84** (2008), no. 5, 58001.

[212] I M Tolic-Norrelykke, E L Munteanu, G Thon, L Oddershede, and K Berg-Sorensen, *Anomalous diffusion in living yeast cells*, Physical Review Letters **93** (2004), 078102.

[213] H.C. Tuckwell, *Introduction to theoretical neurobiology*, Cambridge University Press, 1988.

[214] Jeremie Unterberger, *Stochastic calculus for fractional brownian motion with hurst exponent $h > 1/4$: A rough path method by analytic extension*, Annals of Probability **37** (2009), no. 2, 565–614.

[215] M. v. Smoluchowski, *Drei vortrage über diffusion, brownsche molekularbewegung und koagulation von kolloidteilchen*, Physikalische Zeitschrift **17** (1916), 557 – 571.

[216] N. G. van Kampen, *Stochastic processes in physics and chemistry*, North-Holland, Amsterdam, 1992.

[217] V van Noort, B Snel, and MA Huynen, *The yeast coexpression network has a small-world, scale-free architecture and can be explained by a simple model*, EMBO REPORTS **5** (2004), no. 3, 280–284 (English).

[218] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley, *Levy flight search patterns of wandering albatrosses*, Nature **381** (1996), no. 6581, 413–415.

[219] G. M. Viswanathan, Sergey V. Buldyrev, Shlomo Havlin, M. G. E. da Luz, E. P. Raposo, and H. Eugene Stanley, *Optimizing the success of random searches*, Nature **401** (1999), no. 6756, 911–914.

[220] G. M. Viswanathan, E. P. Raposo, and M. G. E. da Luz, *Lèvy flights and superdiffusion in the context of biological encounters and random searches*, Physics of Life Reviews **5** (2008), no. 3, 133–150.

[221] P H Von Hippel, *From "simple" DNA-protein interactions to the macromolecular machines of gene expression*, Annual Review of Biophysics and Biomolecular Structure **36** (2007), 79–105.

[222] M. von Smoluchowski, *Experiments on a mathematical theory of kinetic coagulation of coloid solutions.*, Zeitschrift für Physikalische Chemie **92** (1917), 129–168.

[223] Malte Wachsmuth, Waldemar Waldeck, and Jrg Langowski, *Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy*, Journal of Molecular Biology **298** (2000), no. 4, 677–689.

[224] K.G. Wang and C.W. Lung, *Long-time correlation effects and fractal brownian motion*, Physics Letters A **151** (1990), no. 3–4, 119 – 121.

[225] Duncan J. Watts and Steven H. Strogatz, *Collective dynamics of /'small-world/' networks*, Nature **393** (1998), no. 6684, 440–442.

[226] Stephanie C. Weber, Andrew J. Spakowitz, and Julie A. Theriot, *Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm*, Phys. Rev. Lett. **104** (2010), 238102.

[227] Feller William, *An introduction to probability theory and its applications*, vol. 2, Wiley, January 1971.

[228] I. Y. Wong, M. L. Gardel, D. R. Reichman, Eric R. Weeks, M. T. Valentine, A. R. Bausch, and D. A. Weitz, *Anomalous diffusion probes microstructure dynamics of entangled f-actin networks*, Phys. Rev. Lett. **92** (2004), 178101.

[229] W.R.Schneider and W.Wyss, *Fractionnal diffusion and wave equations*, J.Math.Pys. **30** (1987), 134–144.

[230] JJ Wu, ZY Gao, HJ Sun, and HJ Huang, *Urban transit system as a scale-free network*, Modern Physics Letters B **18** (2004), no. 19-20, 1043–1049 (English).

[231] Zeba Wunderlich and Leonid A. Mirny, *Spatial effects on the speed and reliability of protein-dna search*, Nucleic Acids Research **36** (2008), no. 11, 3570–3578.

[232] Soichiro Yamada, Denis Wirtz, and Scot C. Kuo, *Mechanics of living cells measured by laser tracking microrheology*, Biophysical Journal **78** (2000), no. 4, 1736 – 1747.

[233] Haw Yang, Guobin Luo, Pallop Karnchanaphanurach, Tai-Man Louie, Ivan Rech, Sergio Cova, Luying Xun, and X. Sunney Xie, *Protein conformational dynamics probed by single-molecule electron transfer*, Science **302** (2003), no. 5643, 262–266.

[234] E A Yuzbasyan, H Lin, N C Darnton, J B Stock, P Silberzan, S Park, P M Wolanin, and R H Austin, *Influence of topology on bacterial social interaction*, Proceedings of the National Academy of Sciences **100** (2003), 13910–13915.

[235] Irwin M. Zaid, Michael A. Lomholt, and Ralf Metzler, *How subdiffusion changes the kinetics of binding to a surface*, Biophys. J. **97** (2009), no. 3, 710–721.

[236] Zhongzhi Zhang, Jihong Guan, Wenlei Xie, Yi Qi, and Shuigeng Zhou, *Random walks on the apollonian network with a single trap*, EPL (Europhysics Letters) **86** (2009), no. 1, 10006.

[237] Zhongzhi Zhang, Yi Qi, Shuigeng Zhou, Shuyang Gao, and Jihong Guan, *Determining mean first-passage time for random walks on deterministic uniform recursive trees using the eigenvalues of laplacian matrices*, 2009.

[238] Zhongzhi Zhang, Yi Qi, Shuigeng Zhou, Wenlei Xie, and Jihong Guan, *Exact solution for mean first-passage time on a pseudofractal scale-free web*, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **79** (2009), no. 2, 021127–6.

[239] Zhongzhi Zhang, Weilen Xie, Shuigeng Zhou, Mo Li, and Jihong Guan, *Distinct scalings for mean first-passage time of random walks on scale-free networks with the same degree sequence*, 2009.

[240] Zhongzhi Zhang, Wenlei Xie, Shuigeng Zhou, Shuyang Gao, and Jihong Guan, *Anomalous behavior of trapping on a fractal scale-free network*, 2009.

[241] Zhongzhi Zhang, Yichao Zhang, Shuigeng Zhou, Ming Yin, and Jihong Guan, *Influences of degree inhomogeneity on average path length and random walks in disassortative scale-free networks*, Journal of Mathematical Physics **50** (2009), no. 3, 033514–9.

[242] Zhongzhi Zhang, Shuigeng Zhou, Wenlei Xie, Lichao Chen, Yuan Lin, and Jihong Guan, *Standard random walks and trapping on the koch network with scale-free behavior and small-world effect*, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **79** (2009), no. 6, 061113–11.

[243] A. Zoia, E. Dumonteil, and A. Mazzolo, *Collision-number statistics for transport processes*, Phys. Rev. Lett. **106** (2011), 220602.

Résumé

Les propriétés de premier passage en général, et parmi elles le temps moyen de premier passage (MFPT), sont fréquemment utilisées dans les processus limités par la diffusion. Les processus réels de diffusion ne sont pas toujours Browniens : durant les dernires années, les comportements non-Browniens ont été observés dans un nombre toujours croissant de systmes. Les milieux biologiques sont un exemple frappant o ce genre ce comportement a été observé de faon répétée. Nous présentons dans ce manuscrit une méthode basée sur les propriétés de premier passage permettant d'obtenir des informations sur le processus réel de diffusion, ainsi que sur l'environnement o évolue le marcheur aléatoire. Cette méthode permet de distinguer trois causes possibles de sous-diffusion : les marches aléatoires en temps continu, la diffusion en milieu fractal et le mouvement brownien fractionnaire. Nous étudions également l'efficacité des processus de recherche sur des réseaux discrets. Nous montrons comment obtenir les propriétés de premier passage sur réseau afin d'optimiser ensuite le processus de recherche, et obtenons un encadrement général du temps moyen de premier passage global (GMFPT). Grâce à ces résultats, nous estimons l'impact sur l'efficacité de recherche de plusieurs paramtres, notamment la connectivité de la cible, la mobilité de la cible ou la topologie du réseau.

Mots-clés: Propriétés de premier passage, diffusion anormale, marches aléatoires, réseaux discrets.

Zusammenfassung

Eigenschaften der First Passage, dem erstmaligen Überschreiten eines festgelegten Grenzwertes, und die Mean First Passage Time (MFPT) im besonderen sind häufig benutzte Methoden zur Charakterisierung von stochastischen Prozessen. Dabei sind reale Prozesse nicht immer Brownscher Natur: in den letzten Jahren wurde nicht-Brownsches Verhalten in einer zunehmenden Anzahl von Systemen beobachtet. Insbesondere Single particle tracking Experimente in lebenden Zellen sind ein frappierende Beispiel, wo nicht-Brownsches Verhalten in der Form von Subdiffusion vielfach beobachtet wird. Wir schlagen hier eine Methode vor, die auf First Passage Eigenschaften beruht, um weitergehende physikalische Informationen über den zugrundeliegenden stochastischen Prozess sowie die physische Umgebung, in der der Diffusionsprozeß abläuft, zu gewinnen. Diese Methode erlaubt es uns, zwischen den drei prominenten Modelle der Subdiffusion zu unterscheiden: Continuous Time Random Walks, Diffusion auf Fraktalen und fractional Brownian motion. Wir untersuchen außerdem Random Walks auf diskreten Netzwerken. Wir zeigen, wie via First Passage Eigenschaften auf diesen Netzwerken die diffusive Suche nach einem bestimmten Ziel optimiert werden kann. Wir erhalten allgemeine Schranken für die globale MFPT. Mit diesen Ergebnissen schätzen wir die Effizienz des Suchprozesses ab, wenn verschiedene Parameter wie die Ziel-Konnektivität, Bewegung des Ziels oder Veränderungen der Netzwerktopologie stattfinden.

Stichworte: Eigenschaften der First Passage, Diffusionsprozess, nicht-Brownsches Verhalten, diskrete Netzwerke.

Abstract

First-passage properties in general, and the mean first-passage time (MFPT) in particular, are widely used in the context of diffusion-limited processes. Real processes are not always purely Brownian: in the last few years, non-Brownian behaviors have been observed in an increasing number of systems. Especially single particle experiments in living cells provide striking examples for systems in which non-Brownian behavior of subdiffusive kind has been repeatedly observed experimentally. Here we present a method based on first-passage properties to gain more detailed insight into the actual physical processes underlying the anomalous diffusion behavior, and to probe the environment in which this diffusion process evolves. This method allows us to discriminate between three prominent models of subdiffusion: continuous time random walks, diffusion on fractals, and fractional Brownian motion. We also investigate the search efficiency of random walks on discrete networks for a specific target. We show how to compute first-passage properties on those networks in order to optimize the search process, as well as general bounds on the global mean first-passage time (GMFPT). Using those results, we estimate the impact on the search efficiency of several parameters, namely the target connectivity, the target motion, or the network topology.

Keywords: First-passage properties, anomalous diffusion, random walks, discrete networks.