Universidade do Minho
Escola de Ciências

Tiago da Silva Alves de Nogueira Simões

Kirigami at the microscale

março de 2021

Tiago da Silva Alves de Nogueira Simões **Kirigami at the microscale**

Uminho | 2021

Universidade do Minho
Escola de Ciências

Tiago da Silva Alves de Nogueira Simões

Kirigami at the microscale

Dissertação de Mestrado
Mestrado em Física

Trabalho efetuado sob a orientação do
Professor Doutor Nuno Miguel Azevedo Machado de Araújo
e do
Professor Doutor Mikhail Igorevich Vasilevskiy

março de 2021

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

AGRADECIMENTOS

Ainda que seja de índole maioritariamente individual, a concretização deste trabalho não teria sido possível sem o apoio de um certo grupo de pessoas. Dedico, portanto, esta primeira página para lhes agradecer e reconhecer o seu contributo no âmbito desta dissertação.

Ao Professor Doutor Nuno Miguel Azevedo Machado de Araújo, o meu orientador, pelo tempo e vontade dispensadas para me guiar neste trabalho e etapa da vida, pelo rigor e excelência instruídas que prosperaram o meu método de trabalho, e por ter confiado nas minhas capacidades.

Ao Professor Doutor Mikhail Igorevich Vasilevskiy, o meu co-orientador, pela valiosa revisão do conteúdo deste trabalho e pela ajuda imprescindível ao longo do meu percurso académico, particularmente na orientação enriquecedora do meu projeto de licenciatura.

Ao Doutor Hygor Piaget Monteiro Melo, pela sua mentoria indispensável, camaradagem e disponibilidade para me ajudar a qualquer momento no decurso deste trabalho.

Agradeço ainda em geral a todos os membros do grupo de acolhimento do Centro de Física Teórica e Computacional da Faculdade de Ciências da Universidade de Lisboa por me terem recebido e pelo ambiente saudável e instrutivo que proporcionaram.

A todo o corpo docente da Universidade do Minho que me transmitiu ao longo do meu percurso de licenciatura e mestrado os conhecimentos e aptidões necessárias que possibilitaram a realização desta dissertação. Dirijo também os meus agradecimentos a todos os meus colegas nesta jornada académica que tornaram a experiência ainda mais engrandecedora.

Quero ainda agradecer a três tutores em particular, ao professor António José da Cunha Costa Dias, ao professor Domingos Manuel Machado Oliveira e à professora Maria do Carmo da Cunha Costa Dias, três pessoas importantes na minha vida escolar que certamente tiveram um papel fulcral na escolha e preparação da minha iniciação académica.

Finalmente, dirijo ainda os meus agradecimentos aos meus pais e avós, por definirem quem eu sou hoje, pelo seu incondicional apoio e constante interesse no meu trabalho.

A todos aqueles que, direta ou indiretamente, tornaram possível o galgar desta etapa, o meu sincero e profundo obrigado!

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

KIRIGAMI À MICRO-ESCALA

RESUMO

Kirigami, a antiga arte oriental da dobra e corte do papel, tem inspirado novas técnicas na fabricação de microestruturas tridimensionais que se dobram espontaneamente a partir de materiais planos, sem ser necessária intervenção manual. Estas estruturas oferecem uma larga gama de novas funcionalidades físicas comparativamente com as que são proporcionadas por materiais planos. Entender quais os mecanismos subjacentes e qual a escala temporal típica destes processos de dobra é então um desafio de relevância atual. Neste trabalho consideramos uma pirâmide regular com N faces laterais como estrutura alvo, obtida através da dobra devido a flutuações térmicas a partir de uma planificação em forma de estrela com N pontas. Assumiu-se que a base desta planificação estava fixa a uma superfície, suprimindo a possibilidade de faces laterais se aderirem em lados opostos da base. Para geometrias simples como esta, o movimento rotacional das faces laterais em torno das arestas que as ligam à base é consistente com um processo Browniano. O processo de dobra ocorre através de uma sequência de eventos de ligação entre faces laterais num ângulo específico de fecho. Propomos um modelo de rede para estudar a dinâmica deste processo, mapeando a posição angular das faces laterais em partículas difusivas unidimensionais. A evolução temporal do sistema é realizada através de um algoritmo de Monte Carlo cinético. Através de resultados numéricos independentes, mostramos que o tempo médio de dobra varia de uma forma não monótona com N, com um valor ótimo que minimiza este tempo. Para um N elevado e uma configuração inicial plana, a escala temporal típica é governada por uma lei de potência com o ângulo de fecho. Usando conceitos de processos de primeira passagem e teoria de ordem, estes resultados numéricos são suplementados com soluções analíticas. Esta escala temporal é também afetada pela maneira como as faces se aderem. Compararam-se também os tempos de dobra típicos para várias condições iniciais e concluiu-se que os tempos mais rápidos são obtidos em média quando as faces começam empilhadas sobre a base. Os tempos mais lentos estão associados a uma configuração inicial plana ou com faces posicionadas em lados opostos do espaço angular. Por último, estudou-se um sistema que se dobra em suspensão, onde defeitos são possíveis. Nestas condições, o número ótimo de faces para um processo rápido é mais elevado quando comparado com o sistema fixado. A eficiência do processo, definida pela fração de amostras que se dobram sem defeitos, decresce com N, pelo que um equilíbrio entre o tempo e eficiência tem de ser decidido com base no que é prioritário.

Palavras-chave: física de matéria mole; Monte Carlo cinético; passeios aleatórios; processos de primeira passagem; self-assembly.

KIRIGAMI AT THE MICROSCALE

ABSTRACT

Kirigami, the ancient eastern art of folding and cutting paper, has inspired novel techniques for the fabrication of 3D microstructures which self-fold spontaneously from flat templates, without the need of manual intervention. These folded structures offer a plethora of new physical functionalities compared to flat materials. Understanding how and how long it takes for the planar templates to fold into the desired 3D shape is then of growing interest and importance. Here we use a regular pyramid with N lateral faces as the target structure, which folds due to thermal fluctuations from a N-pointed star template. We assume that this template is pinned through the base at a substrate, preventing misfolding. For simple geometries, the rotational motion of the lateral faces is consistent with Brownian movement. Folding occurs through a sequence of binding events between lateral faces at a specific closing angle. We propose a lattice model to study the dynamics of self-folding, mapping the angular position of each lateral face into a unidimensional diffusing particle. The time evolution of the system is studied using a kinetic Monte Carlo approach. From independent numerical results we show that the average folding time varies non-monotonically with the number of lateral faces N, leading to the existence of an optimal value of N for fast folding. We also find that, if the template starts from a planar configuration, this time average is governed by a simple scaling law with the closing angle when N is large enough. Applying concepts from first passage processes and order theory, these numerical findings are supplemented with analytical results. The characteristic time scale of the folding process is also affected by the binding mechanism. We compare the average folding time of several initial configurations of the net and find that stacking the lateral faces on top of the base yields the fastest results. The slowest times are associated with either a planar conformation or when lateral faces are configured in opposite sides of the angular space. Lastly, we additionally study a system which folds in suspension, which allows for the possibility of defects in the target structure when lateral faces bind on both sides. Under these conditions, numerical independent results show that the optimal number of lateral faces for fast folding increases when compared to the case on a substrate. The folding efficiency, defined as the fraction of samples without defects, decreases with N and a specific balance between folding time and efficiency needs to be considered depending on the priorities.

Keywords: first passage processes; kinetic Monte Carlo; random walks; self-assembly; soft matter physics.

Contents

List of Abbreviations

BC - boundary condition

FPT - first passage time

IC - initial condition

kMC - kinetic Monte Carlo

MFPT - mean first passage time

PDF - probability density function

SC - standard conditions

List of Symbols

N - number of lateral faces

θ_i - outer angle of lateral face i

α_i - dihedral angle between the lateral face i and the base

t - time; total number of hops

ϕ - closing angle

L - number of lattice sites

Δt - time step

n - index identifying a lattice site

θ_u - uppermost possible value of the outer angles

θ_l - lowermost possible value of the outer angles

$\Delta\theta$ - lattice spacing

$P_n\left(t\right)$ - probability of finding the random walk at site n at time t

p - probability of hopping to the right

q - probability of hopping to the left

k_r - hopping rate to the right

k_l - hopping rate to the left

k - total hopping rate

$P\left(\theta, t\right)$ - probability density of finding an outer angle with value θ at time t

v - bias velocity

D - diffusion coefficient

$N_n(t)$ - number of particles at site n at time t

$\Delta\phi$ - length of the trapping region

t_i - i-th binding time

T_F - first binding time

T_L - last binding time

T - folding time

N_F - number of free particles

r_N - randomly generated integer used to choose a free particle

r_w - randomly generated number used to decide the hopping direction

$\delta_{x,y}$ - delta Kronecker function with arguments x and y

a_m - middle site of the lattice

r_L - randomly generated integer used to choose a lattice site

L_F - number of lattice sites where a particle may still be placed

S_n - list of available sites where a particle may still be placed

a_i - i-th randomly selected initial lattice site

$\overline{x}$ - mean of random variable x

N_S - number of samples

x_i,y_i - data points

$\langle x \rangle$ - best estimate of $\overline{x}$

δx - uncertainty in the best estimate of $\overline{x}$

P_{upper} - probability of the first binding event to occur at the upper side

B_i - number of successes of a Bernoulli trial specified by the subscript i

P_{lower} - probability of the first binding event to occur at the lower side

$P(r,t)$ - probability that a single random walk made r hops to the right out of a total of t hops

r - number of hops to the right in a single random walk

l - number of hops to the left in a single random walk

a - initial outer angle

t_D - diffusion time

$\delta(x)$ - delta Dirac function with argument x

$S(t)$ - survival probability

$f(t \mid a)$ - distribution of the diffusion time for a defined initial outer angle a

$F(t \mid a)$ - antiderivetive of $f(t \mid a)$

$f_0(a)$ - distribution of the initial outer angle

$f_R(t)$ - distribution of the diffusion time for a random initial outer angle

$\mathrm{Li}_s(z)$ - polylogarithm function of order s and argument z

$\zeta(s)$ - Riemann zeta function with argument s

$\Phi(z, s, a)$ - Lerch transcendent function with arguments z, s and a

ΔX - relative error

σ_x - standard deviation of random variable x

$t_F^{(i)}$ - set of first passage time associated with a 2D Brownian process

$f_F(t)$ - distribution from which the first passage time $t_F^{(i)}$ are drawn

$S_F(t)$ - probability that a particular measurement drawn from f_F yields a value larger than t

$F_F(t)$ - distribution of the first binding time T_F

N_P - number of data points

a_0,a_1 - fitting parameters

$t_L^{(i)}$ - set of first passage times associated with a 1D Brownian process

$f_L(t)$ - distribution from which the first passage times $t_L^{(i)}$ are drawn

$S_L(t)$ - probability that a particular measurement drawn from f_L yields a value larger than t

$P_{\vec{a}}(\vec{n}, t)$ - occupation probability of a 2D random walk at site $\vec{n}$ at time t, provided it started at $\vec{a}$

$F_{\vec{a}}(\vec{n}, t)$ - first passage probability for a 2D random walk to reach site $\vec{n}$ for the first time at time t, provided in started at $\vec{a}$

$\tilde{P}_{\vec{a}}(\vec{n}, z)$ - generating function of $P_{\vec{a}}(\vec{n}, t)$

$\tilde{F}_{\vec{a}}(\vec{n}, z)$ - generating function of $F_{\vec{a}}(\vec{n}, t)$

$\langle t_{\vec{a}\to\vec{n}} \rangle$ - mean first passage time for a 2D random walk to first reach site $\vec{n}$, provided it started at $\vec{a}$

$P_{\phi}(\theta, t)$ - probability density of the case $N = 3$ of finding the outer angle of the free lateral face with value θ at time t, as a function of ϕ

$H(x)$ - Heaviside step function with argument x

$P_F(\theta, \phi)$ - average conditional probability density for the case $N = 3$ of finding the outer angle of the free lateral face with value θ at the moment of the first binding event, as a function of ϕ

$f_{\phi}(t)$ - distribution of the last binding time T_L for the case $N = 3$, which takes into account the angular subdomain beyond ϕ

$\varepsilon_{\phi}(n)$ - average conditional occupation probability for the case $N = 3$ of finding the free particle at site n at the moment the other two get trapped at $n(\phi)$

j_R - index of a right adjacent lateral face

j_L - index of a left adjacent lateral face

$H[x]$ - discrete Heaviside step function with argument x

$P(\text{down} \mid \text{up})$ - conditional probability for the second binding event to occur at the lower side provided the first binding occurred at the upper side

$P(\text{up} \mid \text{down})$ - conditional probability for the second binding event to occur at the upper side provided the first binding occurred at the lower side

$P(\text{up} \mid \text{up})$ - conditional probability for both the first and second binding events to occur at the upper side

$P(\text{down} \mid \text{down})$ - conditional probability for both the first and second binding events to occur at the lower side

η - folding efficiency

μ_n - n-th distribution moment

χ^2 - chi-squared distribution

List of Figures

List of Tables

Introduction

Origami is the ancient art of folding a sheet of paper into a hollowed structure, with its most likely origin rooted in the regions of China, Korea and Japan [1]. Kirigami is a variant of this art which allows cutting the paper, expanding the range of accessible shapes. The words derive from the Japanese ori ('to fold'), kiri ('to cut') and kami (paper) [1, 2].

Let us consider, for example, a polyhedron. To obtain this 3D structure from folding a 2D template, one needs first to identify a 2D structure (net) that can fold into it. A net consists of simple polygons, which in the folded structure are the faces of the polyhedron, connected through flexible hinges. First uses of the concept of a net date back to the 15th century with the work of Albrecht Dürer who, for the first time, represented a polyhedron with a net [3]. It is assumed that all convex polyhedra have at least one net, since no one has found a convex polyhedron for which there is no net [1]. Nevertheless, a formal proof that there is always a net is still elusive [1].

Kirigami has inspired novel approaches for the design of new materials across a wide range of length scales, going from the macroscale to the mesoscopic scale ($\sim 10\mathrm{nm}$ to $\sim 100\mu\mathrm{m}$) [4, 5]. In material science, forming 3D nano- and microstructures is of growing interest due to the new characteristics and functionalities offered by its extra dimensionality, when compared to its planar counterparts [5]. As such, methods to obtain these 3D shapes are sought after, but traditional 3D nano- and microfabrication techniques impose limitations on both the efficiency of the process and the complexity of the accessible shapes [6]. The appeal of origami/kirigami inspired techniques stems mainly from the ability to obtain 3D structures 'solely' by folding and cutting planar structures, and being able to control the shape of the folded structures using only a few degrees of freedom [6]. These characteristics, along with the compatibility of this approach with current established fabrication methods such as compressive buckling or lithography [7–12], provide many possibilities to the development of new materials.

Applications are found across a wide range of length scales. At the macroscale, the folding pathway is deterministic [13], and the focus is on 3D functional structures whose actuation is driven by external applied

forces [4]. Examples include foldable solar trackers [14], shape-shifting smart materials and morphing components [15, 16] and power wearable technology [17]. Challenges at this length scale are mainly related to increasing the actuation response speeds and effectively turning the geometrical idealizations into practical functional platforms [4].

At the submillimeter-scale, stimuli-responsive materials are used to bypass the needs of manual intervention [4]. Planar templates of interconnected rigid panels can be produced through conventional lithographic tools, and the template self-folds into the desired structure, driven by thermal fluctuations, making the folding process stochastic [18]. This allows for large scale parallel productions and reduced cost [18, 19]. Applications for self-folding kirigami are varied, and include drug encapsulation and delivery [20–22], soft-robotics [23], micro-electromechanical systems (MEMSs) [24] and photonic materials [25]. Despite its advantages, self-folding micro kirigami suffers from being a generally slow method [6] and difficult to control, with the possibility of 2D structures not folding into the desired 3D shape [18].

To understand folding at the microscale and identify design strategies to optimize the process, several studies have been developed [18, 19, 26]. One of the challenges is that the same 3D structure can be obtained from several different nets. For example, in the case of polyhedrons with identical faces composed of convex regular polygons, the so called platonic solids, the tetrahedron has two nets, the cube and octahedron have 11, and the dodecahedron and icosahedron have 43,380 nets [18]. We see that, even for relatively simple shapes, the set of possible nets can be large, and each net exhibits different folding pathways and different propensity to fold into the desired 3D structure [27]. This raises the question: from all possible nets of a given polyhedron, which ones are optimal for self-folding, free of defects?

In an attempt to help answering this question, Pandey et al. [18] studied the surface-tension driven self-folding of several nets, by computing the yield, defined as the fraction of nets that, from the full sample set, self-folded without defects. They compared two criteria to optimise this fraction: geometrical compactness, measured by the radius of gyration of the net, and topological compactness, defined from the number of vertex connections, defined as the vertexes in the net which share two faces but, in turn, do not share an edge. The authors catalogued a set of nets for the dodecahedron, icosahedron and truncated octahedron, and selected three for each polyhedron: the most compact, the least compact and the intermediate value for the compactness criterion. For each net, 50 samples were experimentally self-folded under similar experimental conditions. The final folded structures were classified into three groups based on the amount of defects. They found that the amount of defects is minimized when the net has a maximum number of vertex connections. Thus, topological compactness is a more efficient criterion than geometrical compactness.

Dodd et al. [26] have shown that the folded configuration depends not only on the chosen net but also on the experimental conditions. The authors remarked that, since edge binding is nonspecific, the ground state is not, in general, unique, and may not correspond to the desired folded structure, due to the formation of structural defects. To understand those defects, they classified the contacts into native or nonnative, i.e. edge bindings that are present in the desired folded structure and those which are not. Using Markov models to define the folding pathways between the unfolded and folded states, and Langevin molecular dynamics to simulate the folding dynamics, the authors showed that the folding propensity varies with temperature, with its efficiency peaking at an optimal value. For low temperatures, the folding pathway occurs mainly through nonnative contacts, with low values of flux. On the contrary, for higher temperatures, native contacts pathways are favoured. The authors explained that this preference is due to a balanced trade-off between entropy and enthalpy, whereby the nets fold locally, one face folding at a time. This reduces the fewest degrees of freedom per state transition, thereby maximizing its conformational entropy at each step.

Both these aforementioned works elucidated guiding principles for optimizing the folding efficiency for self-folding kirigami, but they require an exhaustive search of the entire set of nets of a given polyhedron, and, as previously mentioned for more complex geometries, this number can get too large to be computable. To overcome this challenge, random methods have been proposed but they fail to provide a unique and deterministic solution [19]. In addition, it has been shown that the fraction of nets which have a maximum number of vertex connections decays exponentially with the number of edges of the polyhedron [19]. This finding reinforces the need for a deterministic method since chances of randomly picking these nets fastly decay with the number of faces of the polyhedron.

Araújo et al. [19] proposed a deterministic and systematic method to identify nets with maximum vertex connections. By using concepts from Graph Theory, the authors mapped the geometry of each respective kirigami structure into two graphs (a collection of nodes connected by links): a shell graph, where the vertices and edges of the polyhedron correspond to the shell graph nodes and links, respectively, and a face graph, where each face of the polyhedron is assigned a node of the face graph, and each link connects a pair of adjacent faces. Then, the corresponding net can be obtained by cutting along the edges, which corresponds to removing the associated links in the shell graph, under the constraint that the face graph remains connected. The set of removed links along with the nodes associated with the vertices define a cut sub-graph that is a spanning tree (a sub-graph that includes every node of the corresponding super-graph and has no loops) of the shell graph. The vertex connections are then the nodes of the cut sub-graph with only one incident link, called leaves. The problem of finding nets with a maximum number of vertex

connections is then equivalent to finding the so-called maximum leaf spanning tree (MLST) of the shell graph. With this deterministic method it is possible to obtain a complete list of nets for optimal self-folding of polyhedra and enables prescription of additional design criteria.

Overall, the study of self-folding microstructures has focused on identifying the nets that minimize the number of defects. However, for some practical applications, it is important to minimize the folding time, since, as previously mentioned, one of the current challenges of kirigami at the microscale is to speed up the folding process. To this end, Melo et al. [13] studied the folding time dependence on the number of faces of a polyhedron, using as a prototypical geometry a regular pyramid with N lateral faces. By showing that the motion of each lateral face is consistent with a random (Brownian) process, the authors derived analytical expressions for the folding time. They found that the time for the first binding event decreases with the logarithm of N, while the time for the remaining sequential binding events grows logarithmically with N. Since the folding process evolves through a set of events, the folding time is a non-monotonous function of N, with a minimum occurring for five lateral faces. The authors tested and verified the analytical predictions by running independent Langevin molecular dynamics simulations, using a coarse-grained numerical model to describe the net, starting from a flat template. This work shows that the geometry of the net can be used to set the time of the folding process, presenting an additional accessible control strategy in conjugation with previous ones based on manipulation of temperature and pH of the surrounding medium [28]. It additionally shows that, even for simple geometries like a regular pyramid, the folding time depends non trivially on the topology of the net. This prompt us into investigating how this time may depend on other parameters as well, such as the slant angles of the polyhedron or the initial configuration of the net.

To gain insight into the temporal statistics of self-folding kirigami at the microscale, in this work we present a lattice model to simulate the stochastic self-folding of a prototypical net of a regular pyramid with N lateral faces. This new approach is computationally fast compared to previous methods, such as the Langevin molecular dynamics. In chapter 1 we present the details of the lattice model and its algorithmic implementation. The results of this work are discussed in two chapters. In chapter 2 we show statistical results whose outcome could be predicted analytically, to validate the numerical modeling. In chapter 3 results from the statistical analysis using the lattice model are presented and discussed, namely, the temporal dependency with the geometry of the net, such as the number of lateral faces and slant angle, and with the net initial configuration, as well as other properties concerning the foldability and interactivity between the panels. We further investigate the temporal and folding efficiency for cases where defects might be formed.

Chapter 1

Model

To study the dynamics of self-folding driven by thermal fluctuations at the microscale, we consider a regular pyramid. This type of geometry has already been used in experimental research [28, 29]. For most of the work, we will assume that the lateral faces can only rotate in one side of the plane, to suppress misfolding and focus on the study of the folding time. This can be realized experimentally, for example, by pinning the base into a substrate [28, 29].

Below, in section 1.1, we first present a lattice model to characterize the self-folding process of nets into regular pyramids, and in section 1.2, we discuss the algorithmic implementation of this model.

1.1 A lattice model for self-folding kirigami at the microscale

For a regular pyramid with N lateral faces, due to thermal fluctuations, each lateral face will jiggle, rotating in a stochastic fashion around the edge that connects the lateral face to the base. The lateral faces move until they bind pairwise and irreversibly when they touch. We assume that binding can occur between vertices of the faces and, therefore, two faces can bind together even when they are not adjacent. We will refer to the outer angles of each lateral face $i = \{0, 1, ..., N-1\}$ as $\theta_i = \pi - \alpha_i$, where α_i is the dihedral angle between the lateral face i and the base (see Fig. 1.1a). Previous numerical simulations suggest that the time series of $\theta_i(t)$ is consistent with a 1D Brownian process [13]. For a net with its base pinned to a substrate, the motion of the lateral faces is limited to the domain $\theta_i \in [0, \pi]$. If the folding process occurs in suspension, however, the lateral faces can rotate through the whole domain $\theta_i \in [-\pi, \pi]$. We will refer to the former as a pinned system, and the latter as a suspended one. In both cases, the boundaries are reflective. A system with all faces with $\theta_i = 0$ corresponds to a planar net,

with all faces being in the same plane. A face with $\theta_i = \pi$ touches the base from above, while one with $\theta_i = -\pi$ touches the base from below. The outer angle of the lateral faces at which the net is in its folded state is defined as the closing angle ϕ, corresponding to the outer slant angle of the respective pyramid (see Fig. 1.1b).

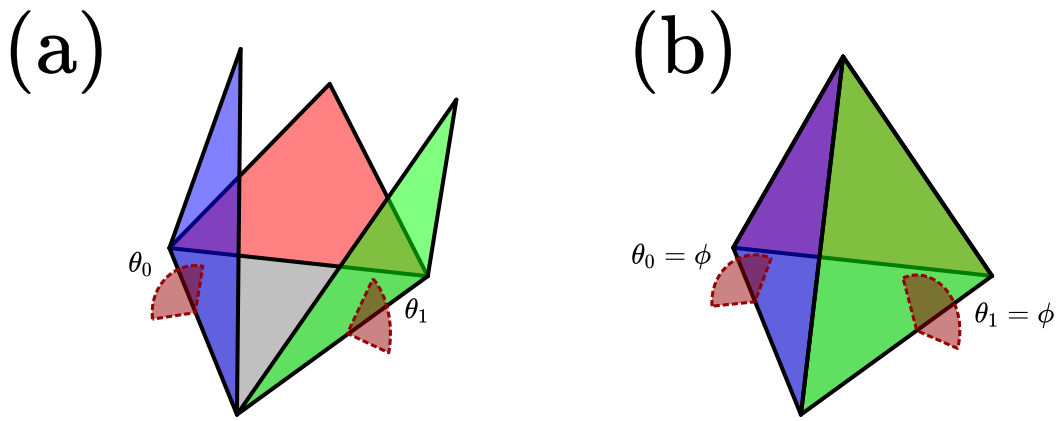

Figure 1.1: (a) Schematic depicting a net for a regular pyramid with three lateral faces, coloured blue, green and red. The position of face $i = \{0, 1, 2\}$ is described by its outer angle θ_i. The state of the folding at each time t is specified by stating the values of the three outer angles $\{\theta_0(t), \theta_1(t), \theta_2(t)\}$. (b) The folded structure with all three lateral faces bound together, with all outer angles equal to the closing angle ϕ.

We propose a lattice model to simulate the spontaneous folding of pyramidal nets at the microscale. One of the main advantages of a lattice model is its ability to highlight the elementary processes that govern the dynamics [30]. We consider a one-dimensional lattice composed of L sites, with reflective boundaries. Since the motion of the lateral faces is diffusive, their dynamics are well-described by a random walk [31]. Each lateral face i is assigned a particle walking randomly on this lattice, to represent its angular position relative to the base over time t. Each time a particle hops, t is incremented by an amount of Δt. Since Δt is inversely proportional to the hopping rate [30], its value should decrease as the number of sites L increases, so that the time scale is independent of the chosen discretization. An outer angle θ is mapped into a site n of the lattice according to

$$n(\theta) = \left\lfloor \frac{L-1}{\theta_u - \theta_l} \cdot (\theta - \theta_l) + 0.5 \right\rfloor \equiv \left\lfloor \frac{\theta - \theta_l}{\Delta\theta} + 0.5 \right\rfloor \quad , \tag{1.1}$$

where $\lfloor\rfloor$ is the floor function, θ_u and θ_l are, respectively, the uppermost and lowermost values of the outer angles, and $\Delta\theta = {}^{(\theta_u - \theta_l)}/_{(L-1)}$ is the lattice spacing. For a pinned system, $\theta_u = \pi$ and $\theta_l = 0$, while for a suspended one, $\theta_u = \pi$ and $\theta_l = -\pi$.

Accordingly, the lattice sites n are mapped back into an outer angle using

$$\theta(n) = n\frac{\theta_u - \theta_l}{L-1} + \theta_l = n\Delta\theta + \theta_l \quad . \tag{1.2}$$

For a single particle i, on the lattice and without particle-particle interactions, the probability $P_n(t)$ for a random walker to be at the lattice site n at some time t obeys the following equation [31]

$$P_n(t+\Delta t) = pP_{n-1}(t) + qP_{n+1}(t) \quad , \tag{1.3}$$

where p ($q = 1-p$) is the probability of hopping to the right (left). The first term in the right-hand side of Eq. (1.3) accounts for the increase in $P_n(t)$ due to a possible hop from $n-1$ to n, while the second term accounts for the increase in $P_n(t)$ due to a possible hop from $n+1$ to n.

More generally however, a particle may have a probability $1-p-q$ of staying in the site n during the time interval Δt, in which case $p+q \leq 1$ and

$$P_n(t+\Delta t) = pP_{n-1}(t) + qP_{n+1}(t) + (1-p-q)P_n(t) \quad , \tag{1.4}$$

where the extra term accounts for the aforementioned possibility.

If $\Delta\theta$ and Δt are small enough, which can be achieved by making the number of sites L arbitrarily large, we can Taylor expand up to first order in time around t on the left-hand side term,

$$P_n(t) + \frac{\partial P_n(t)}{\partial t}\cdot(t+\Delta t - t) + ... = pP_{n-1}(t) + qP_{n+1}(t) + (1-p-q)P_n(t) \quad . \tag{1.5}$$

Rearranging, dividing all terms by Δt, and taking the continuum time limit, we get

$$\frac{\partial P_n(t)}{\partial t} = k_r P_{n-1}(t) + k_l P_{n+1}(t) - kP_n(t) \quad , \tag{1.6}$$

where $k_r = {}^{q}/_{\Delta t}$, $k_l = {}^{p}/_{\Delta t}$ and $k = k_r + k_l = {}^{(p+q)}/_{\Delta t}$ are the hopping rates from $n-1$ to n, from $n+1$ to n and the total hopping rate from n to $n \pm 1$, respectively.

To go from discrete to continuous space, we can reinterpret $P_n(t) \to P(\theta, t)$ as the probability density of finding a particle at $\theta(n)$ at time t, and Taylor expand up to first order around θ the first two terms on

the right-hand side,

$$\begin{aligned}\frac{\partial P(\theta,t)}{\partial t} &= k_r \cdot \left[P(\theta,t) + \frac{\partial P(\theta,t)}{\partial \theta} \cdot (\theta - \Delta\theta - \theta) + ...\right] \\ &+ k_l \cdot \left[P(\theta,t) + \frac{\partial P(\theta,t)}{\partial \theta} \cdot (\theta + \Delta\theta - \theta) + ...\right] \\ &- kP(\theta,t) \quad . \end{aligned} \tag{1.7}$$

We see that the zeroth term cancels out, and we get the transport equation,

$$\frac{\partial P(\theta,t)}{\partial t} + v\frac{\partial P(\theta,t)}{\partial \theta} = 0 \quad , \tag{1.8}$$

where $v = (k_r - k_l)\,\Delta\theta = (p-q)\cdot({}^{\Delta\theta}/_{\Delta t})$ is the bias velocity.

This first order approximation does not capture the diffusive properties, so we go back to Eq. (1.4), dividing once again all terms by Δt, and Taylor expand both in time and space up to second order,

$$\begin{aligned}&\frac{1}{\Delta t}P(\theta,t) + \frac{\partial P(\theta,t)}{\partial t} + \frac{\Delta t}{2}\frac{\partial^2 P(\theta,t)}{\partial t^2} + ... = \\ &= k_r \cdot \left[P(\theta,t) - \Delta\theta\frac{\partial P(\theta,t)}{\partial \theta} + \frac{\Delta\theta^2}{2}\frac{\partial P^2(\theta,t)}{\partial \theta^2} + ...\right] \\ &+ k_l \cdot \left[P(\theta,t) + \Delta\theta\frac{\partial P(\theta,t)}{\partial \theta} + \frac{\Delta\theta^2}{2}\frac{\partial P^2(\theta,t)}{\partial \theta^2} + ...\right] + \left(\frac{1}{\Delta t} - k\right)P(\theta,t) \quad , \end{aligned} \tag{1.9}$$

and again the zeroth term cancels out, so we obtain

$$\frac{\partial P(\theta,t)}{\partial t} + \frac{\Delta t}{2}\frac{\partial^2 P(\theta,t)}{\partial t^2} = -v\frac{\partial P(\theta,t)}{\partial \theta} + k\frac{\Delta\theta^2}{2}\frac{\partial^2 P(\theta,t)}{\partial \theta^2} \quad . \tag{1.10}$$

We can recast the second order time derivative term into a second order space derivative using the transport equation (1.8),

$$\begin{aligned}\frac{\partial^2 P(\theta,t)}{\partial t^2} &= \frac{\partial}{\partial t}\left[-v\frac{\partial P(\theta,t)}{\partial \theta}\right] = -v\frac{\partial}{\partial \theta}\left[\frac{\partial P(\theta,t)}{\partial t}\right] = \\ &= -v\frac{\partial}{\partial \theta}\left[-v\frac{\partial P(\theta,t)}{\partial \theta}\right] = v^2\frac{\partial^2 P(\theta,t)}{\partial \theta^2} \quad , \end{aligned} \tag{1.11}$$

where we made use of Schwarz's theorem to interchange the order of differentiation in the second step.

We can use this equality to recast Eq. (1.10) into the form of the convection-diffusion equation,

$$\frac{\partial P(\theta,t)}{\partial t} + v\frac{\partial P(\theta,t)}{\partial \theta} = D\frac{\partial^2 P(\theta,t)}{\partial \theta^2} \quad , \tag{1.12}$$

where $D = (\Delta\theta^2/2\Delta t)\cdot\left(p+q-(p-q)^2\right)$ is the diffusion coefficient. If there is zero probability for the particle to stay in the same place for a time interval of Δt, i.e. $p+q=1$, then

$$D = 2pq\frac{\Delta\theta^2}{\Delta t} \quad . \tag{1.13}$$

For an unbiased walk, $p = q = 1/2$ and $v = 0$. In this case, Eq. (1.12) simplifies to the diffusion equation,

$$\frac{\partial P(\theta,t)}{\partial t} = D\frac{\partial^2 P(\theta,t)}{\partial \theta^2} \quad , \tag{1.14}$$

with $D = \Delta\theta^2/2\Delta t$.

The folding process occurs through a series of irreversible binding events between the lateral faces when the corresponding particles in the lattice model meet at the closing angle site $n(\phi)$. Let $N_n(t)$ be the number of particles in site n at time t. The specific site $n(\phi)$ acts as a sticky site when $N_{n(\phi)}(t) > 1$, i.e. it is occupied by more than one particle. We assume that the binding is not point-wise, but can occur in a small region $\phi \pm \Delta\phi/2$ of the angular space [13], where $\Delta\phi \equiv \pi/180$ is the size of this region. More generally then, assuming $\Delta\phi$ is a multiple of the lattice spacing $\Delta\theta \leq \Delta\phi$, the trapping region in the lattice would be a set of sites

$$\left\{n\left(\phi+\frac{-\Delta\phi+\Delta\theta}{2}\right), n\left(\phi+\frac{-\Delta\phi+2\Delta\theta}{2}\right), ..., n(\phi), ..., n\left(\phi+\frac{\Delta\phi-2\Delta\theta}{2}\right), n\left(\phi+\frac{\Delta\phi-\Delta\theta}{2}\right)\right\} \tag{1.15}$$

in which all of them act as sticky sites if

$$\sum_{n=n\left(\phi+\frac{-\Delta\phi+\Delta\theta}{2}\right)}^{n\left(\phi+\frac{\Delta\phi-\Delta\theta}{2}\right)} N_n(t) > 1 \quad , \tag{1.16}$$

i.e. the sum of all particles currently at sites $n \in [n(\phi - \Delta\phi/2 + \Delta\theta/2), n(\phi + \Delta\phi/2 - \Delta\theta/2)]$ is larger than unity. However, for simplicity, we will choose the number of lattices sites L accordingly, so that this region always corresponds to a singular site. If the lateral faces are only allowed to fold into one side, only one of such region exists, which is located at $n(\theta = \phi)$. If they can fold to both sides, two sites can act as sticky sites, one in the upper side, at $n(\theta = \phi)$, and another in the lower side, at $n(\theta = -\phi)$, and

binding can occur at either side. Thus, given a pyramidal net with more than three lateral faces, defects are possible in a suspended system, corresponding to some faces binding at at one side and the others at the opposite side. The time at which the i-th binding event occurs will be denoted as the i-th binding time t_i. For a suspended system there can be at least $N-2$ of these binding times, if lateral faces bind at both lower and upper side (an example of a defective folding), or at most $N-1$, if all N lateral faces bind only at one side. For a pinned system there are always $N-1$ binding times.

Considering a pinned system, the physical mechanism governing the first binding event is fundamentally different from the remaining ones [13]. For the first binding event to occur, two lateral faces need to have their outer angle equal to the closing angle $\phi \pm {}^{\Delta\phi}/_{2}$ at the same time. Thus, the first binding event can be mapped into a set of Brownian processes in 2D, each defined by the coordinates (θ_j, θ_k) of the outer angles of lateral faces j and k and a trap at $(\phi \pm {}^{\Delta\phi}/_{2}, \phi \pm {}^{\Delta\phi}/_{2})$. The first binding occurs when one of these Brownian processes first hits the trap. For the subsequent binding events, the lateral faces can bind one-by-one when they reach an outer angle of ϕ, therefore consistent with a set of Brownian processes in 1D. The folding process is complete when the last of these 1D Brownian processes reach ϕ. For a suspended system, the first binding event is modified by having two traps, one at $(-\phi \pm {}^{\Delta\phi}/_{2}, -\phi \pm {}^{\Delta\phi}/_{2})$ and another at $(\phi \pm {}^{\Delta\phi}/_{2}, \phi \pm {}^{\Delta\phi}/_{2})$. Because of this, subsequent binding events are no longer equivalent to 1D Brownian processes, but rather a mixed process where lateral faces can either bind one-by-one by reaching the closing angle where the first two pair of lateral faces previously met, or another pair can bind if they reach the other unoccupied closing angle. If the latter case occurs, the remaining binding events are then again 1D Brownian processes, but with two traps rather than one.

In view of this distinction, we will focus on two particular binding times. The first binding time t_1, which we will redefine as T_F, and the folding time T, the moment at which the folding process is completed. For convenience, we will also define the last binding time $T_L = T - T_F$.

For a suspended system, the region between the two critical angles, $\theta \in [-\phi + {}^{\Delta\phi}/_{2}, \phi - {}^{\Delta\phi}/_{2}]$, is a transparent region, in which the lateral faces can rotate freely, so the particles can coexist in a given site and pass each other without interacting, while the regions $\theta \in [-\pi, -\phi - {}^{\Delta\phi}/_{2}[\cup]\phi + {}^{\Delta\phi}/_{2}, \pi]$ are interacting regions, where if two particles are at the same site simultaneously they will interact with each other, undergoing an elastic collision. For a pinned system, the notion of transparent and interacting region applies logically only for the domains $\theta \in [0, \phi - {}^{\Delta\phi}/_{2}]$ and $\theta \in]\phi + {}^{\Delta\phi}/_{2}, \pi]$, respectively. These interactions are a first simplified modelling of the interaction between faces during the dynamics.

A schematic illustration of this model for both pinned and suspended systems is presented in Fig. 1.2. The particular case $N = 2$ corresponds to a net whose folded state is an open structure, which can be

obtained for example by removing one of the faces of a tetrahedron or, in general, by removing $N-2$ lateral faces from a N lateral faced regular pyramid. Although this case may not have an apparent practical purpose, it is nonetheless insightful to consider it theoretically, since it is the simplest case where a binding event between lateral faces can still be defined and can thus be instructive for more complex situations.

Using this model, the main focus of our work will be studying the statistics of the folding time T it takes for the polyhedron net to fold back into its 3-dimensional shape through stochastic processes, and, in particular, how the folding time depends on the number of lateral faces N, the closing angle ϕ, the initial conditions and how the propensity to self-fold without defects varies with the geometrical and initial configuration of the kirigami structure.

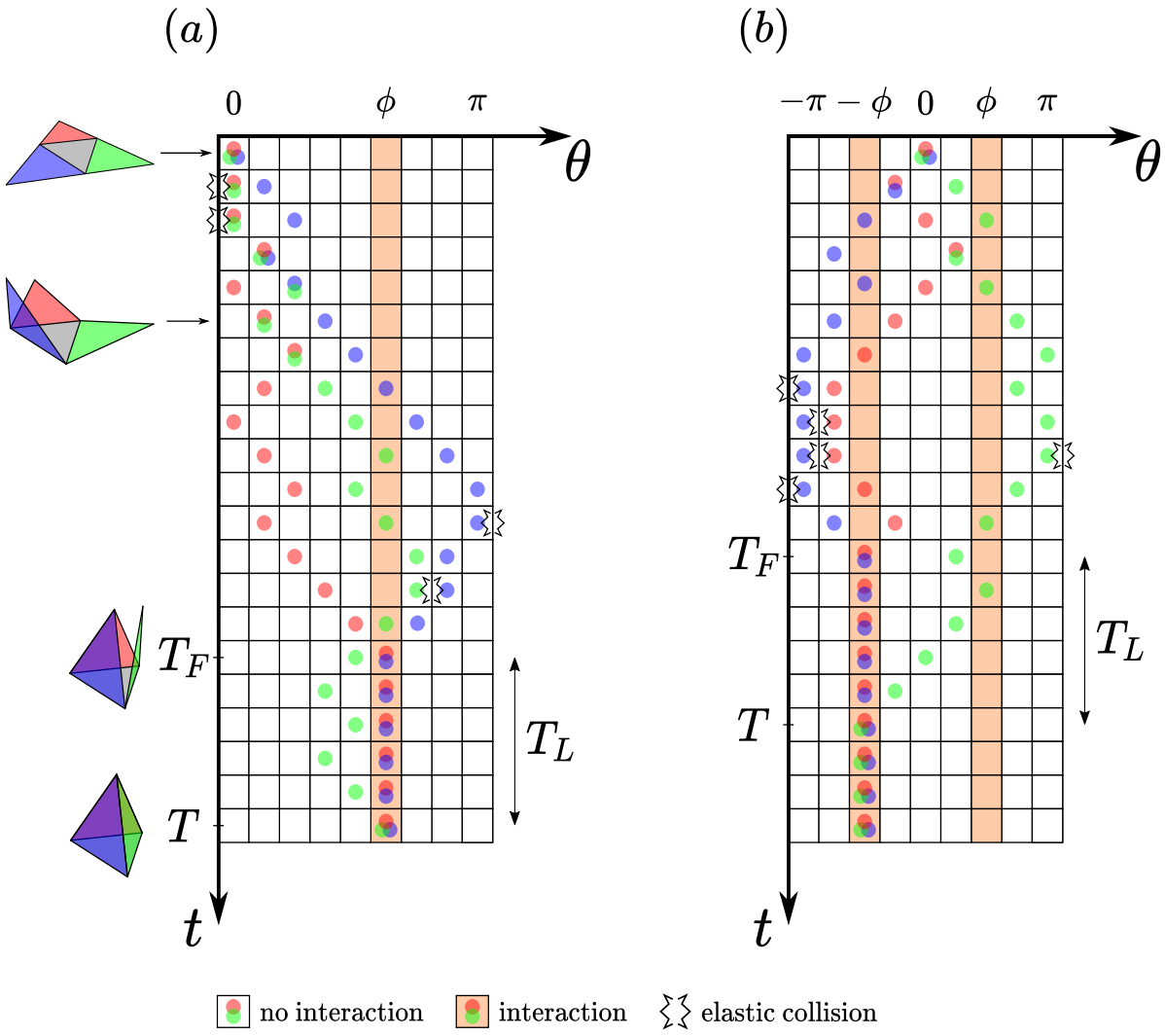

Figure 1.2: (a) Schematic representation of the folding of a pyramidal structure with three lateral faces and closing angle, starting from a flat template, using the lattice model for a pinned system (the lateral faces can only fold to one side). Each lateral face i is assigned a corresponding particle θ_i. The horizontal axis represents the angle of each face at a certain time t, in the domain $\theta \in [0, \pi]$ with reflective boundaries, and the vertical axis is the time t elapsed since the faces started to move. The folding process for $N = 3$ is composed by two binding events. The first binding time T_F is the moment at which the first pair of particles get trapped together at ϕ. The folding time T is when all particles meet at the closing angle ϕ, thus completing the folding process. The last binding time T_L is defined to be the time interval between first binding and the folding time, so $T_L = T - T_F$. The possibility for the particles to undergo elastic collisions at the boundaries or between particles when they are located at $\theta \in]\phi, \pi]$ is also depicted. The corresponding pyramid configuration for four time frames ($t = 0$, $0 < t < T_F$, $t = T_F$ and $t = T$) is also presented. (b) A similar scheme but for a suspended system. In this case, the domain is defined in $\theta \in [-\pi, \pi]$, elastic collisions may occur at $\theta \in [-\pi, -\phi[\cup]\phi, \pi]$, and particles may get trapped either at $-\phi$ or ϕ. In the example depicted, particles got trapped at the lower side, at $-\phi$.

1.2 Implementation of the model: the algorithm

The lattice model is implemented using a kinetic Monte Carlo (kMC) approach [30, 32, 33]. Kinetic Monte Carlo is a Monte Carlo method that allows to consider time explicitly and, therefore, a direct comparison to the physical time. The parameters are: L, the number of sites which make up the lattice; N, the number of random walking particles, equal to the number of lateral faces of the regular pyramid whose folding dynamics we want to simulate; ϕ, the closing angle (to have physical meaning in the context of this geometry, ϕ has to be larger than $\pi/2$, because $\phi = \pi/2$ defines a regular pyramid with infinite height); D, the angular diffusion coefficient associated with each lateral face; q, the probability of a particle to hop into a neighbouring right site; p, the probability of a particle to hop into a neighbouring left site, is then calculated as $p = 1 - q$.

The algorithm follows the next steps:

1. Distribute the N particles on a lattice with L sites, numbered from 0 to $L-1$, according to the chosen initial condition (IC), initiate the time variable as $t = 0$, set $N_F = N$, where N_F is the number of free particles that have yet to bind at ϕ, and assign $n(\phi)$ (or both $n(-\phi)$ and $n(\phi)$, in the case of a suspended system) as the site corresponding to the closing angle, according to Eq. (1.1),
$$n(\phi) = \left\lfloor \frac{L-1}{\theta_u - \theta_l} \cdot (\phi - \theta_l) + 0.5 \right\rfloor = \left\lfloor \frac{\phi - \theta_l}{\Delta\theta} + 0.5 \right\rfloor \quad ; \tag{1.17}$$

2. Generate an integer random number $r_N \in [0, N_F - 1]$ to choose one of the N_F particles from the list of active particles, the particles that are still performing the random walk;

3. Generate a uniformly distributed random number $r_w \in]0, 1]$ to decide if the randomly chosen particle hops into a neighbouring right site (if $r_w < q$) or if it hops into a neighbouring left site (if $r_w > q$, or similarly $r_w < p$);

4. Move the randomly chosen particle presently located at site n to either site $n-1$ or $n+1$ according to the output of r_w, unless one of the following situations arises: (1) the particle is at site 0 and tries to hop into the left; (2) the particle is at site $L-1$ and tries to hop into the right; (3) the particle tries to hop into a site $n(\phi) < n \leq L-1$, for a pinned system, or $0 \leq n < n(-\phi) \vee n(\phi) < n \leq L-1$, for a suspended system, and that site is already occupied by another particle; if any of these previous conditions is met, the randomly chosen particle remains in the site n;

5. Increment the time variable t by an amount of Δt to account for the time spent by the hopping (or at least its attempt) of the randomly chosen particle; when there are still free particles ($N_F > 0$), the value of Δt is calculated according to the definition of the diffusion coefficient (1.13) as follows

$$\Delta t = 2pq\frac{\Delta\theta^2}{D}\frac{1}{N_F} \quad , \tag{1.18}$$

where the additional factor N_F is introduced to account for the fact that the algorithm only moves (or tries to move) one particle per iteration, so the hopping rate of each particle is independent of the number N_F of free particles. For the same physical parameters, the time step value is constant, since we assume discrete time. Here we see that D is just a rescaling parameter, setting the time scale;

6. If the new position of the randomly chosen particle is $n(\phi)$ (or either $n(-\phi)$ or $n(\phi)$, for the case of a suspended system) corresponding to the closing angle, and if $N_{n(\phi)}(t) > 1$ (or either $N_{n(-\phi)}(t) > 1$ if the particle is at $n(-\phi)$ or $N_{n(\phi)}(t) > 1$ if it is at $n(\phi)$, for the suspended system), the site acts as a sticky site and traps the new-coming particle, as well as any other of those particles also currently at that site if they are not yet trapped; if it is the first time particles are getting absorbed at a closing angle, the current value of t corresponds to the first binding time T_F; update the number of particles that are still free, by subtracting the correct amount from N_F and removing from the list of free particles the ones that just got trapped;

7. Repeat steps 2 through 6 until there are no free particles left ($N_F = 0$); when N_F reaches zero, the current value of t is used to compute the last binding time T_L, by subtracting from t the value of T_F; the current value of t itself corresponds to the folding time T, and the folding process is concluded.

For the ICs, the most commonly used throughout this work will be a net starting from a flat (2D) template, which we will refer as a planar IC. Another common IC will be a net starting with its panels at a random outer angle. We implement these ICs in the following way:

Planar allocate all N particles at the zeroth site of the lattice, for a pinned system, so

$$N_n(t=0) = N\delta_{n,0} \quad ; \tag{1.19}$$

if it is a suspended system, all N particles are placed in the middle of the lattice, calculated as

$a_m = \lfloor L/2 \rfloor$, in which case

$$N_n\left(t=0\right) = N\delta_{n,a_m} \quad ; \tag{1.20}$$

Random for each of the N particles, an integer random number $r_L \in [0, L_F - 1]$ is generated, where L_F is the number of sites where a particle may still be placed, to choose a site from the list S_n of available sites where the particle is then allocated; before generating any value of r_L, $L_F = L$ and $S_n = \{0, 1, ..., L-1\}$; for each output i of r_L, if the randomly selected site a_i corresponding to the list element $S_n\left(r_L^{(i)}\right)$ is a closing angle site or is in the interacting region (if $a_i \geq n\left(\phi\right)$ for a pinned system, or $a_i \leq n\left(-\phi\right) \vee a_i \geq n\left(\phi\right)$, for a suspended one), subtract one unit from L_F, and update the list S_n of available sites by removing the one that just got chosen, $S_n \rightarrow S_n \setminus S_n\left(r_L^{(i)}\right)$; with the set of randomly selected sites $\{a_1, a_2, ..., a_N\}$, the N particles are distributed according to

$$N_n\left(t=0\right) = \sum_{i=1}^{N} \delta_{n,a_i} \quad . \tag{1.21}$$

All the values of every random variable are generated using the Mersenne Twister pseudo-random number generator [34], and before they are used to assign a value to any random variable used in the lattice model algorithm, such as r_N or r_w, the first 10^8 pseudo-random numbers generated are always discarded, to reduce possible correlations due to the seed.

The results of four simulations using this model, for random walks with a planar IC in a pinned and suspended systems, for several number of particles, are presented in Fig. 1.3.

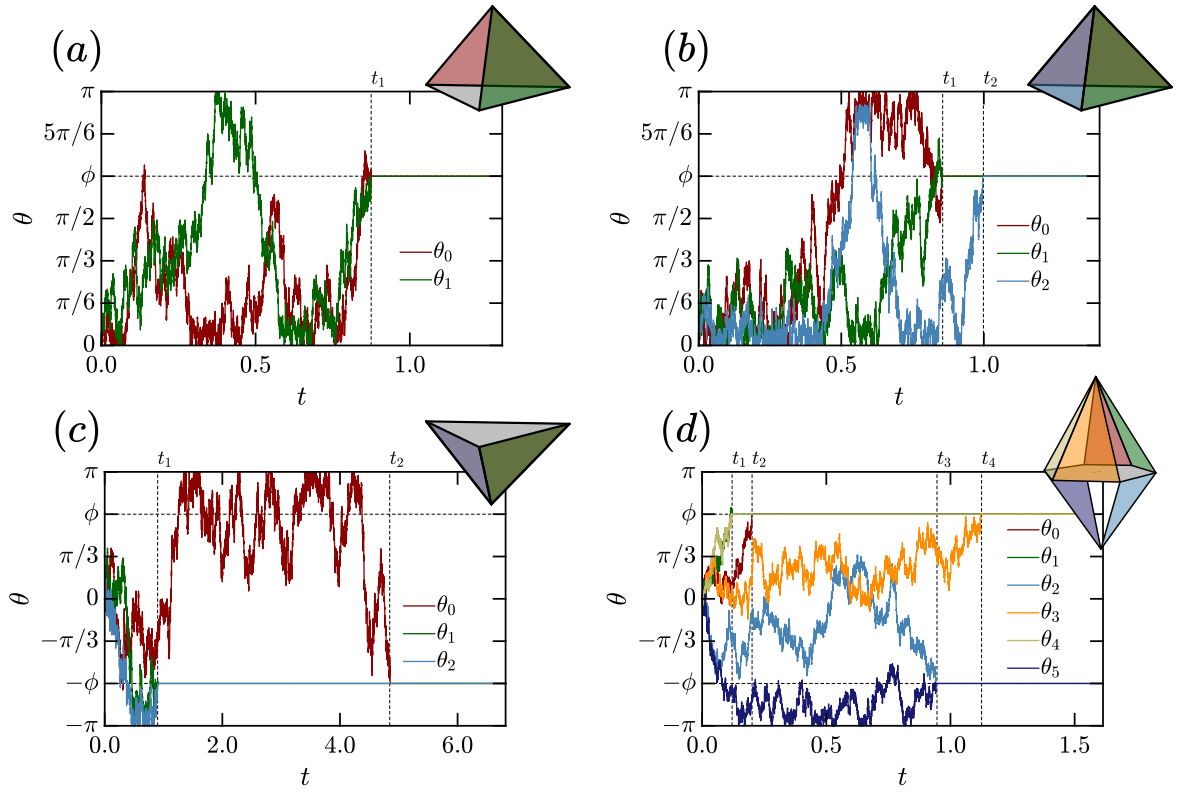

Figure 1.3: Simulation results of several random walks with a planar IC using the lattice model and kMC approach, for (a) two and (b) three particles in a pinned system, as well as (c) three and (d) six particles in a suspended system. For pinned and suspended systems, the lattice had $L = 181$ and $L = 361$ sites, respectively, with corresponding domains $\theta \in [0, \pi]$ and $\theta \in [-\pi, \pi]$. The value of the closing angle was $\phi = {}^{2\pi}/_{3}$ (represented with the horizontal black dashed lines) for all the random walks. The time is in units of Brownian time, defined as the average time it takes for a diffusive particle to travel over a region of length π. t_i ($i = \{1, 2, 3, ...\}$) is the i-th binding time, the time in which there is a particle absorption at a closing angle (marked with the vertical black dashed lines). The corresponding folded configuration of the pyramid is schematized in the top right corner of each graph. The folded configuration corresponding to the random walk (c) is symmetrically equivalent to the one in (b). Random walk (d) corresponds to a folding process where defects are possible, and, in fact, the folded configuration ends up being a misfolded pyramid, where lateral faces have bound on both sides.

Chapter 2

Validation of the model

In this section we present the results to validate and test the numerical model.

In this thesis, unless otherwise specified, the random walk simulations were performed with the following standard conditions (SC):

- A domain of $\theta \in [0, \pi]$ for a pinned system, or $\theta \in [-\pi, \pi]$ for a suspended one;
- A lattice of size $L = 181$ sites for a pinned system, and $L = 361$ sites for a suspended one, so that the lattice spacing always corresponds to $\Delta\theta = {}^{\pi}/_{180}$, and the closing angle always occupies one site in the lattice;
- Equal probabilities of hopping either to the right or to the left (i.e. $p = q = {}^{1}/_{2}$);
- A diffusion coefficient of reference $D = {}^{1}/_{2} \cdot ({}^{\pi}/_{180})^2 \equiv D_0$, which, considering the SC for $\Delta\theta$, p and q, by virtue of Eq. (1.18), corresponds to a time step interval of $\Delta t = 1$ au (arbitrary units of time) when $N_F = 1$;
- A closing angle of $\phi = {}^{2\pi}/_{3}$ (this corresponds to a square regular pyramid in which the slant height of the lateral faces is equal to the distance in the base between the opposite lateral faces, which in turn is equal to twice the in-radius);

We report the results of the mean $\overline{x}$ of some random variable x, whose outcomes for N_S independent samples were $\{x_1, x_2, ..., x_{N_S}\}$, in the standard way (more details in appendix section A.1)

$$\overline{x} = \langle x \rangle \pm \delta x \quad , \tag{2.1}$$

where $\langle x \rangle$ is the best estimate of $\overline{x}$, computed as

$$\langle x \rangle = \frac{1}{N_S} \sum_{i=1}^{N_S} x_i \quad , \tag{2.2}$$

and δx is a measure of the uncertainty in our best estimate of $\overline{x}$ (that is, the uncertainty in the value of $\langle x \rangle$), calculated as

$$\delta x = \sqrt{\frac{\langle x^2 \rangle - \langle x \rangle^2}{N_S}} \quad , \tag{2.3}$$

where

$$\left\langle x^2 \right\rangle = \frac{1}{N_S} \sum_{i=1}^{N_S} x_i^2 \quad . \tag{2.4}$$

All the histograms presented in this work were normalized so that it directly represents the distribution of the respective random variable. The data sets of the histograms were divided into 10 subsets, each with 10% of the total samples, so that for each bin we have 10 samples from where we can compute the best estimate of the bin height and its respective uncertainty by means of the standard form (2.1).

2.1 Probability of the first binding event to occur at either the lower or upper side for an unbiased walk

If the net of the regular pyramid starts self-folding from either a flat template (planar IC) or with the outer angles of its lateral faces randomly distributed (random IC), we should expect that, for an unbiased walk where $p = q = 1/2$, no matter the number of lateral faces N, the first binding event is as likely to happen either at the lower (at $\theta = -\phi$) or at the upper (at $\theta = \phi$) side, since no direction is favoured.

We test this by defining the probability P_{upper} of the first binding event occurring at the closing angle $\theta = \phi$ (rather than at $\theta = -\phi$), and estimate its value through a Bernoulli trial,

$$P_{\text{upper}} \simeq \frac{B_{\text{upper}}}{N_S} \quad \text{, for large } N_S, \tag{2.5}$$

where B_{upper} is the number of times the first binding event occurred at $\theta = \phi$ out of a total of N_S trials. If we were to measure quantity (2.5) for a large number of trials N_S we should expect its value to approach $1/2$.

We could just as well have defined the analogous probability P_{lower} of the first binding event occurring at

the lower closing angle $\theta = -\phi$ instead, the choice is completely arbitrary, but verifying that just one of them is close to $1/2$ is enough, considering that analysing both quantities does not convey new information since they are complementary, related by $P_{\text{lower}} = 1 - P_{\text{upper}}$, i.e. for each time we know that a first binding event occurred at a upper or lower side we know for certain that it did not happen at the opposite side and vice-versa. A statistical analysis of quantity (2.5) for planar and random ICs for up to $N = 30$ particles is presented in Fig. 2.1.

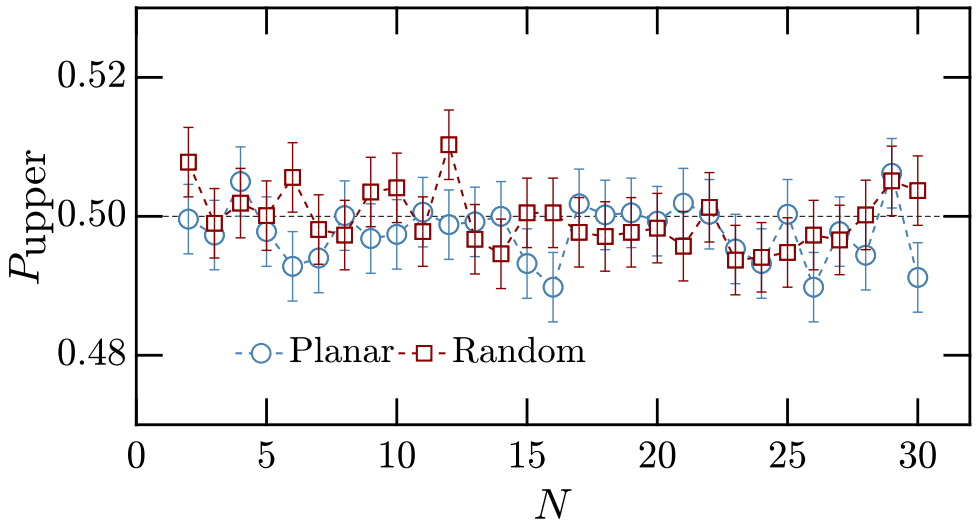

Figure 2.1: Simulation results in a suspended system for the probability (2.5) for various numbers of particles (or lateral faces) $N \in [2, 30]$, for a planar IC (circles) and random IC (squares). The simulation was performed under SC. Results are averages from $N_S = 10^4$ independent trials.

Results in Fig. 2.1 are as expected for both IC, with P_{upper} always close to $1/2$ independently of the number of particles N, with no point deviating from this value more than twice its uncertainty, showing that the algorithm outputs unbiased results when $q = p = 1/2$.

2.2 Comparison of simulations results with a binomial distribution

To estimate the probability $P_n(t)$ that a random walker in a one-dimensional lattice, which can hop into the right with probability p and to the left with probability $q = 1 - p$, starting from a site we define without

loss of generality as the zeroth site, is at some site n after making t hops we could solve Eq. (1.3). Here, for simplicity, we will consider the probability $P(r,t)$ that after t hops the random walker made r hops to the right (ending up at site n), and we just have to count the total number of walks that take a total of r hops to the right [31]. The number of hops l the random walker made to the left is then related to r through

$$\begin{cases} r+l=t \\ \\ r-l=n \end{cases} . \tag{2.6}$$

If there is no restriction in the order of the t hops, the total number of walks of any size and orientation is $t!$, but if we constrain the random walk to make r hops to the right it implies that of these $t!$ arrangements of hops, r hops have to be to the right and l to the left. The number of distinct ways of doing this is ${}^{t!}/_{(r!l!)}$. Furthermore, since each hop to the right or to the left occurs with probabilities p and q, respectively, we still have to multiply ${}^{t!}/_{(r!l!)}$ by p^r, the probability that the walk made r hops to the right, and q^l, the probability that the walk made l hops to the left, to get our probability $P(r,t)$ of a walk that made r hops to the right out of a total of t hops,

$$P(r,t)=\frac{t!}{r!l!}p^r q^l=\frac{t!}{r!\cdot(t-r)!}p^r q^{t-r}=\begin{pmatrix} t \\ \\ r \end{pmatrix}p^r q^{t-r} \quad , \tag{2.7}$$

where $\binom{t}{r}={}^{t!}/_{(r!\cdot(t-r)!)}$ is formally known as a Binomial coefficient, and Eq. (2.7) is a Binomial distribution.

Using relations (2.6), we find that $r={}^{(t+n)}/_2$ and $l={}^{(t-n)}/_2$, and we can reformulate Eq. (2.7) into our original probability $P_n(t)$ as a function of n and t,

$$P(r,t)=P_n(t)=\frac{t!}{\left(\frac{t+n}{2}\right)!\cdot\left(\frac{t-n}{2}\right)!}p^{\frac{t+n}{2}}q^{\frac{t-n}{2}} \quad . \tag{2.8}$$

In the limit of large $t \gg n$ we can use Stirling's approximation to reduce the factorials in Eq. (2.7) into analytical functions, and find that the binomial distribution reduces into a normal (or Gaussian) distribution

in this limit (see appendix B),[1]

$$P(r,t) = P_n(t) \simeq \frac{1}{\sqrt{2\pi tpq}} e^{-\frac{(n-t\cdot(p-q))^2}{8tpq}} \quad \text{, for large } t \gg n. \tag{2.9}$$

To test our algorithm using these distributions, we ran N_S independent simulations where a single particle performed an unbiased random walk of precisely t hops, starting at site zero, and performed a Bernoulli trial, counting the number of times $B_{n,t}$ that the particle finished at site n. The division $B_{n,t}/N_S$, for large N_S, should converge to the same results as those predicted by $P_n(t)$, for a given t.

Results from this analysis are presented in Fig. 2.2 for several values of n, for two values of t.

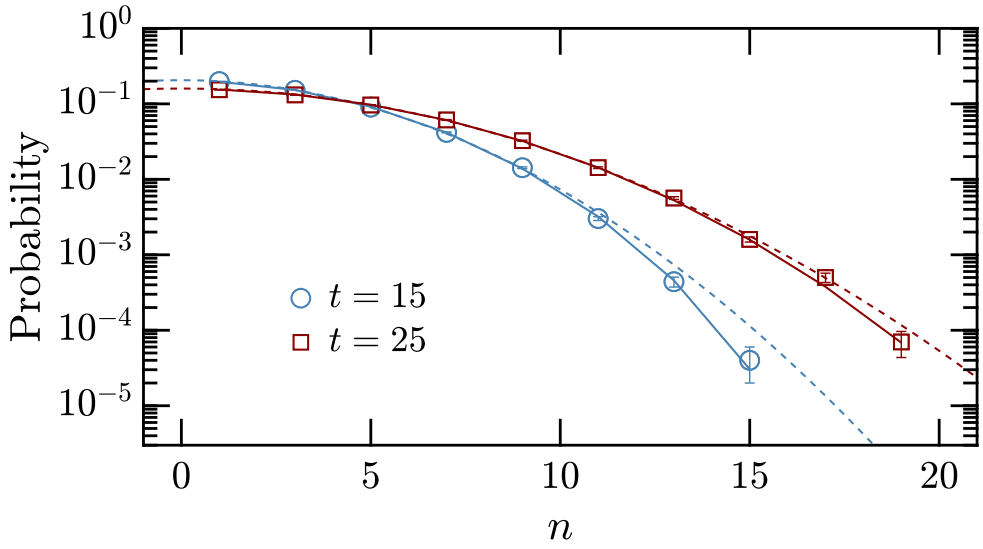

Figure 2.2: Simulation results for the probability that a particle which starts its random walk at site zero is at site n after a walk with $t = 15$ hops (circles) and $t = 25$ hops (squares), along with the respective theoretical values (as indicated by the same colour) predicted by the binomial distribution function (2.8) (continuous curves) and its normal limit (2.9) (dashed curves), both equations using $p = q = 1/2$. Numerical results are averages from $N_S = 10^4$ independent trials.

As we can see from Fig. 2.2 that the numerical results follow the expected distributions. We also note that the normal distribution (dashed curves in Fig. 2.2) start to deviate from the true values of the probability $P_n(t)$ as the approximation becomes less and less valid as n approaches the value of t, as expected.

[1] More precisely, $P_n(t)$ is only approximately equal to the expression (2.9) if n and t have the same parity, otherwise $P_n(t) \equiv 0$ if n and t have opposite parities, since it is impossible to arrive at an even numbered site n after an odd number of t hops, and vice-versa. The same logic applies to Eq. (2.8). Additionally, $P_n(t) \equiv 0$ if $n > t$, since it is also impossible to reach a site n which is farther than the total number of hops t performed.

2.3 Diffusion and first passage times

In this section, we test our simulation by considering a single particle diffusing in a 1D line with reflective lower boundary and an absorbing upper boundary. In our model, this corresponds to a lattice where $\phi = \theta_u = \pi$, and $n(\phi)$ is always a trap. This can simulate a 'toy' system made up of two triangular plates attached at one of the edges, with one plate pinned down and the other freely rotating around the binding edge due to thermal fluctuations. Once the rotating panel hits the other one, they stick. Our particle simulates the time series of the outer angle $\theta = \pi - \alpha$, where α is the dihedral angle between the panels. The diffusion time t_D it takes from the moment the rotating panel starts moving from an initial outer angle $\theta(t = 0) = a$ until it first hits the other one is an example of a first passage time (FPT) [35].

We ran several simulations with this lattice configuration to extract the diffusion time t_D, from which we then compute the mean $\langle t_D \rangle$. Before presenting the numerical results, we derive the theoretical prediction for the average diffusion time $\langle t_D \rangle$. The following derivations can be compared with the numerical results and serve as a preluding theoretical insight for our analysis of kirigami at the microscale.

We start by noticing that the probability density $P(\theta, t)$ of finding the rotating panel with an outer angle θ at time t obeys the diffusion equation (1.14),

$$\frac{\partial P(\theta, t)}{\partial t} = D \frac{\partial^2 P(\theta, t)}{\partial \theta^2} \quad , \tag{2.10}$$

where D is the angular diffusion coefficient of the rotating panel.

The IC that the panel starts with an outer angle $\theta(t = 0) = a$ corresponds to

$$P(\theta, t = 0) = \delta(\theta - a) \quad . \tag{2.11}$$

The boundary conditions at $\theta = 0$ and $\theta = \phi$ are

$$\left. \frac{\partial P(\theta, t)}{\partial \theta} \right|_{\theta=0} = 0 \quad , \tag{2.12}$$

$$P(\theta = \phi, t) = 0 \quad , \tag{2.13}$$

where the first condition is due to the reflective lower boundary, while the second one accounts for the absorbing upper boundary.

Using separation of variables, $P(\theta, t) = \Theta(\theta) T(t)$, and applying the initial and boundary conditions,

the solution for $P(\theta, t)$ can be found (see appendix C)

$$P(\theta, t) = \frac{2}{\phi} \sum_{m=0}^{\infty} e^{-Dk_m^2 t} \cos(k_m a) \cos(k_m \theta) \quad , \tag{2.14}$$

with eigenvalues $k_m = (\pi/\phi) \cdot (m + 1/2)$.

Since the IC (2.11) is normalized, the integral of the probability density function (PDF) (2.14) over the whole angular domain corresponds to what is known as the survival probability $S(t)$ [35], corresponding in this specific situation to the probability that the outer angle θ has not yet reached the absorbing boundary at ϕ by time t,

$$S(t) = \int_0^{\phi} P(\theta, t)\, d\theta \quad . \tag{2.15}$$

The probability (2.15) is related to the PDF $f(t \mid a)$ of obtaining a diffusion time $t_D = t$, for an initial outer angle $\theta = a$, by

$$S(t) = \int_t^{\infty} f(t' \mid a)\, dt' \quad . \tag{2.16}$$

From the fundamental theorem of Calculus we know that

$$\int_t^{\infty} f(t' \mid a)\, dt' = F(\infty \mid a) - F(t \mid a) \quad , \tag{2.17}$$

where $F(t \mid a)$ is the antiderivative of $f(t \mid a)$. Furthermore, $F(\infty \mid a) = 0$ since $f(t \mid a)$ is a PDF and thus must be normalized. Joining Eqs. (2.16) and (2.17) and differentiating both sides in time (using $\partial F(t|a)/\partial t = f(t \mid a)$, by definition), we have the fundamental relation between the distribution of a FPT and the associated survival probability function,

$$f(t \mid a) = -\frac{\partial S(t)}{\partial t} \quad . \tag{2.18}$$

From Eq. (2.15) we see that the only way for the survival probability $S(t)$ to change over time is if the probability density $P(\theta, t)$ is 'draining out' through the absorbing boundary. Thus the change over time of the survival probability must be equivalent to the outgoing flux of the angular probability density at $\theta = \phi$,

$$\frac{\partial S(t)}{\partial t} = D \frac{\partial P(\theta, t)}{\partial \theta}\bigg|_{\theta=\phi} \quad . \tag{2.19}$$

Using Eqs. (2.18) and (2.19), the PDF $f\left(t \mid a\right)$ is then given by

$$f\left(t \mid a\right) = -\, D\frac{\partial P\left(\theta, t\right)}{\partial \theta}\bigg|_{\theta=\phi} = \frac{2D}{\phi}\sum_{m=0}^{\infty}\left(-1\right)^{m} k_m e^{-Dk_m^2 t}\cos\left(k_m a\right) \quad , \tag{2.20}$$

where we used $\sin\left(\pi \cdot \left(m + {}^1\!/_2\right)\right) = \left(-1\right)^m$, since $m \in \mathbb{N}_0$.

If the value of a is random, characterized by some PDF $f_0\left(a\right)$, the PDF for the diffusion time is modified into $f_R\left(t\right)$, calculated by

$$f_R\left(t\right) = \int_0^{\phi} f\left(t \mid a\right) f_0\left(a\right) da \quad . \tag{2.21}$$

For a random initial outer angle with a uniform distribution over the whole domain $a \in [0, \phi]$, the PDF $f_0\left(a\right)$ is

$$f_0\left(a\right) = \frac{1}{\phi} \quad , \tag{2.22}$$

and Eq. (2.21) acquires the form

$$f_R\left(t\right) = \int_0^{\phi} f\left(t \mid a\right)\frac{1}{\phi} da = \frac{2D}{\phi^2}\sum_{m=0}^{\infty} e^{-Dk_m^2 t} \quad , \tag{2.23}$$

where we used the following result for the integral

$$\int_0^{\phi}\cos\left(k_m a\right) da = \frac{\sin\left(k_m \phi\right)}{k_m} = \frac{\sin\left(\pi \cdot \left(m + \frac{1}{2}\right)\right)}{k_m} = \frac{\left(-1\right)^m}{k_m} \quad , \tag{2.24}$$

and the fact that $\left(-1\right)^{2m} = 1$, since $2m$ is always an even number.

The average diffusion time $\langle t_D \rangle$ for a given IC is then found by integrating the product of t with the respective PDF over the whole positive time domain,

$$\langle t_D \rangle_a = \int_0^{\infty} t f\left(t \mid a\right) dt \quad , \tag{2.25}$$

$$\langle t_D \rangle_R = \int_0^{\infty} t f_R\left(t\right) dt \quad , \tag{2.26}$$

where the first equation is for a defined initial outer angle $\theta\left(t = 0\right) = a$, so $P\left(\theta, t = 0\right) = \delta\left(\theta - a\right)$, while the second gives the expected value when the possible initial outer angle is uniformly distributed over the whole angular space, in which case $P\left(\theta, t = 0\right) = {}^1\!/_{\phi}$.

We proceed now to derive the expression for the average diffusion time $\langle t_D \rangle$ for three cases: one starting at an initial position $a = 0$, at a uniformly distributed random initial position, and at a general defined initial position a.

Defined initial position $a = 0$

We start by computing Eq. (2.25) with $a = 0$

$$\langle t_D \rangle_0 = \int_0^\infty t f\left(t \mid 0\right) dt = \frac{2D}{\phi} \sum_{m=0}^{\infty} \left(-1\right)^m k_m \int_0^\infty t e^{-Dk_m^2 t} dt \quad , \tag{2.27}$$

where the 0 subscript in $\langle t_D \rangle_0$ indicates that the rotating panel starts at $\theta\left(t = 0\right) = 0$.

The integral in (2.27) is of the form

$$\int_0^\infty t e^{-st} dt = \frac{1}{s^2} \quad \text{, for } s > 0 \quad , \tag{2.28}$$

whose result can be proved using integration by parts. In this case $s = Dk_m^2$, and so

$$\langle t_D \rangle_0 = \frac{2}{\phi D} \sum_{m=0}^{\infty} \frac{\left(-1\right)^m}{k_m^3} = \frac{2\phi^2}{\pi^3 D} \sum_{m=0}^{\infty} \frac{\left(-1\right)^m}{\left(m + \frac{1}{2}\right)^3} \quad . \tag{2.29}$$

Looking now only at the summation, making the change of variable $m + 1/2 \rightarrow M$ (so $m \rightarrow M - 1/2$)

$$\sum_{m=0}^{\infty} \frac{\left(-1\right)^m}{\left(m + \frac{1}{2}\right)^3} = \sum_{M=\frac{1}{2}}^{\infty} \frac{\left(-1\right)^{M-\frac{1}{2}}}{M^3} = \left(-1\right)^{-\frac{1}{2}} \frac{1}{2^{-3}} \sum_{2M=1;2}^{\infty} \frac{\left(-1\right)^{\frac{1}{2}2M}}{\left(2M\right)^3} \quad , \tag{2.30}$$

where $\sum_{M=i;\Delta M}^{\infty}$ means a summation over the index M starting at i in steps of size ΔM (if the step size $\Delta M = 1$, this index is omitted). Making another change of variable, $2M \rightarrow N$,

$$\left(-1\right)^{-\frac{1}{2}} \frac{1}{2^{-3}} \sum_{2M=1;2}^{\infty} \frac{\left(-1\right)^{\frac{1}{2}2M}}{\left(2M\right)^3} = -8i \sum_{N=1;2}^{\infty} \frac{\left(-1\right)^{\frac{1}{2}N}}{N^3} = -8i \sum_{N=1;2}^{\infty} \frac{i^N}{N^3} = -8i \sum_{N=1;2}^{\infty} \frac{\left(\varepsilon i\right)^N}{N^3} \quad , \tag{2.31}$$

where in the last step a real number ε was introduced, such that $1 - \varepsilon = 0^+$. The motivation for the insertion of ε will become clear right away. First notice that

$$\mathrm{Li}_3\left(\varepsilon i\right) = \sum_{n=1}^{\infty} \frac{\left(\varepsilon i\right)^n}{n^3} = \varepsilon \cdot \left(\frac{i}{1} - \frac{1}{8} - \frac{i}{27} + \frac{1}{64} + \frac{i}{125} - \frac{1}{216} - \frac{i}{343} + \frac{1}{512} + ...\right) \quad , \tag{2.32}$$

$$\mathrm{Li}_3\left(-\varepsilon i\right) = \sum_{n=1}^{\infty} \frac{\left(-\varepsilon i\right)^n}{n^3} = \varepsilon \cdot \left(-\frac{i}{1} - \frac{1}{8} + \frac{i}{27} + \frac{1}{64} - \frac{i}{125} - \frac{1}{216} + \frac{i}{343} + \frac{1}{512} + ...\right) \quad , \tag{2.33}$$

where $\mathrm{Li}_s\left(z\right)$ is the polylogarithm of order s and argument z [36], which has a non-integral representation when $|z| < 1$ (and it was because of this condition that ε was introduced in (2.31)), in

which case it is also referred as Jonquière's function, $\mathrm{Li}_s(z) = \sum_{n=1}^{\infty} {}^{z^n}/_{n^s}$. In Eqs. (2.32) and (2.33), the fact that $\varepsilon^n = \varepsilon$ (since $\varepsilon = 1^-$) was used to factor the constant out of the summation. Second, we have

$$\mathrm{Li}_3(\varepsilon i) - \mathrm{Li}_3(-\varepsilon i) = \varepsilon \cdot \left(2\frac{i}{1} + 0 - 2\frac{i}{27} + 0 + 2\frac{i}{125} + 0 - 2\frac{i}{343} + 0 + \ldots\right) = 2\sum_{N=1;2}^{\infty} \frac{(\varepsilon i)^N}{N^3} \quad , \tag{2.34}$$

so we reach the conclusion that

$$\sum_{N=1;2}^{\infty} \frac{(\varepsilon i)^N}{N^3} = \frac{1}{2} \cdot \left(\mathrm{Li}_3(\varepsilon i) - \mathrm{Li}_3(-\varepsilon i)\right) \quad . \tag{2.35}$$

Plugging the equality (2.35) in (2.31)

$$\sum_{m=0}^{\infty} \frac{(-1)^m}{\left(m+\frac{1}{2}\right)^3} = -8i \sum_{N=1;2}^{\infty} \frac{(\varepsilon i)^N}{N^3} = -8i\frac{1}{2} \cdot \left(\mathrm{Li}_3(\varepsilon i) - \mathrm{Li}_3(-\varepsilon i)\right) = 4i \cdot \left(\mathrm{Li}_3(-\varepsilon i) - \mathrm{Li}_3(\varepsilon i)\right) \quad , \tag{2.36}$$

and going back to (2.29) with the result (2.36) we arrive at

$$\langle t_D \rangle = \frac{8(\theta^*)^2}{\pi^3 D} i \cdot \left(\mathrm{Li}_3(-\varepsilon i) - \mathrm{Li}_3(\varepsilon i)\right) \quad . \tag{2.37}$$

We can use the following identity of the polylogarithm

$$\mathrm{Li}_3(-z) - \mathrm{Li}_3\left(-z^{-1}\right) = -\frac{1}{6}\ln^3(z) - \frac{1}{6}\pi^2 \ln(z) \quad , \tag{2.38}$$

with $z = \varepsilon i$ to simplify the expression (2.37). Knowing $-i^{-1} = i$ and $\ln(i) = {}^{i\pi}/_2$, Eq. (2.38) becomes (applying the limit value 1^- to ε in the last step)

$$\mathrm{Li}_3(-z) - \mathrm{Li}_3\left(-z^{-1}\right) = \mathrm{Li}_3(-\varepsilon i) - \mathrm{Li}_3(\varepsilon i) = -\frac{1}{6}\ln^3(i) - \frac{1}{6}\pi^2 \ln(i) \quad , \tag{2.39}$$

which simplifies to

$$\mathrm{Li}_3(-\varepsilon i) - \mathrm{Li}_3(\varepsilon i) = -\frac{\pi^3}{16} i \quad . \tag{2.40}$$

Substituting this last result in Eq. (2.37) allows us to express $\langle t_D \rangle_0$ in the following simpler form

$$\langle t_D \rangle_0 = \frac{\phi^2}{2D} \quad . \tag{2.41}$$

We see that, if the rotating panel starts with a zero outer angle, the average diffusion time for the rotating panel to hit the pinned one is $\langle t_D \rangle_0 = {}^{\pi^2}/_{2D}$, since $\phi = \pi$. More generally, Eq. (2.41) tells us that, in the line, the average time $\langle t \rangle$ to travel a distance ϕ for a diffusing particle with diffusivity D follows a power law,

$$\langle t\left(\phi\right) \rangle = A\phi^k \quad , \tag{2.42}$$

with parameters $A = {}^1/_{2D}$ and $k = 2$.

All time quantities in this work are reported in units of Brownian time, given by Eq. (2.41) with $\phi = \pi$ and $D = D_0$.

Random initial position

This time we use Eq. (2.26)

$$\langle t_D \rangle_R = \int_0^\infty t f_R\left(t\right) dt = \frac{2D}{\phi^2} \sum_{m=0}^{\infty} \int_0^\infty t e^{-Dk_m^2 t} dt \quad , \tag{2.43}$$

and using the result of the same integral as in the previous case,

$$\langle t_D \rangle_R = \frac{2}{\phi^2 D} \sum_{m=0}^{\infty} \frac{1}{k_m^4} = \frac{2\phi^2}{\pi^4 D} \sum_{m=0}^{\infty} \frac{1}{\left(m + \frac{1}{2}\right)^4} \quad . \tag{2.44}$$

Making the same change of variables in the summation as in the last case, $m + {}^1/_2 \to M$ and then $2M \to N$,

$$\sum_{m=0}^{\infty} \frac{1}{\left(m + \frac{1}{2}\right)^4} = \sum_{M=\frac{1}{2}}^{\infty} \frac{1}{M^4} = \frac{1}{2^{-4}} \sum_{2M=1;2}^{\infty} \frac{1}{\left(2M\right)^4} = 16 \sum_{N=1;2}^{\infty} \frac{1}{N^4} \quad , \tag{2.45}$$

and noticing that

$$16 \sum_{N=1;2}^{\infty} \frac{1}{N^4} = 16 \left(\sum_{n=1}^{\infty} \frac{1}{n^4} - \sum_{N=2;2}^{\infty} \frac{1}{N^4} \right) = 16 \left(\sum_{n=1}^{\infty} \frac{1}{n^4} - \frac{1}{16} \sum_{\frac{N}{2}=1}^{\infty} \frac{1}{\left(\frac{N}{2}\right)^4} \right) \quad . \tag{2.46}$$

We make one more change of variable ${}^N/_2 \to n$ in the second summation term

$$\sum_{m=0}^{\infty} \frac{1}{\left(m + \frac{1}{2}\right)^4} = 16 \left(\sum_{n=1}^{\infty} \frac{1}{n^4} - \frac{1}{16} \sum_{n=1}^{\infty} \frac{1}{n^4} \right) = 15 \sum_{n=1}^{\infty} \frac{1}{n^4} = 15\zeta\left(4\right) \quad , \tag{2.47}$$

where $\zeta\left(s\right) = \sum_{n=1}^{\infty} {}^1/_{n^s}$ is the Riemann Zeta function of a (generally complex) variable s. Knowing that $\zeta\left(4\right) = {}^{\pi^4}/_{90}$ [37] and plugging the result (2.47) in (2.44) we arrive at the simplified expression

for $\langle t_D \rangle$ in the case of a uniformly distributed random initial position

$$\langle t_D \rangle_R = \frac{\phi^2}{3D} = \frac{2}{3} \langle t_D \rangle_0 \quad . \tag{2.48}$$

This result shows that, on average, the time it takes for the rotating panel to hit the pinned one should be $3/2$ faster if it starts at a random outer angle, rather than from a flat template.

General initial position a

We start again by Eq. (2.25), this time with a general value a,

$$\langle t_D \rangle_a = \int_0^\infty t f(t \mid a)\, dt = \frac{2D}{\phi} \sum_{m=0}^{\infty} (-1)^m k_m \cos(k_m a) \int_0^\infty t e^{-D k_m^2 t} dt \quad , \tag{2.49}$$

and, plugging in the result of the usual integral,

$$\langle t_D \rangle_a = \frac{2\phi^2}{\pi^3 D} \sum_{m=0}^{\infty} \frac{(-1)^m}{\left(m+\frac{1}{2}\right)^3} \cos\left(\frac{a\pi}{\phi} \cdot \left(m+\frac{1}{2}\right)\right) \quad . \tag{2.50}$$

Using the exponential representation of the cosine function, $\cos(x) = 1/2 \cdot (e^{ix} + e^{-ix})$, Eq. (2.50) can be redefined as

$$\langle t_D \rangle_a = \frac{2\phi^2}{\pi^3 D} \sum_{m=0}^{\infty} \frac{(-1)^m}{\left(m+\frac{1}{2}\right)^3} \frac{1}{2} \cdot \left(e^{i\frac{\pi}{2\phi}a} e^{i\frac{\pi}{\phi}am} + e^{-i\frac{\pi}{2\phi}a} e^{-i\frac{\pi}{\phi}am}\right) \quad . \tag{2.51}$$

Performing some manipulations, we can express Eq. (2.51) in the following form

$$\langle t_D \rangle_a = \frac{\phi^2}{\pi^3 D} \cdot \left[e^{i\frac{\pi}{2\phi}a} \sum_{m=0}^{\infty} \frac{\left(-e^{i\frac{\pi}{\phi}a}\right)^m}{\left(m+\frac{1}{2}\right)^3} + e^{-i\frac{\pi}{2\phi}a} \sum_{m=0}^{\infty} \frac{\left(-e^{-i\frac{\pi}{\phi}a}\right)^m}{\left(m+\frac{1}{2}\right)^3} \right] \quad , \tag{2.52}$$

which is equivalent to

$$\langle t_D \rangle_a = \frac{\phi^2}{\pi^3 D} \cdot \left[e^{i\frac{\pi}{2\phi}a} \Phi\left(-e^{i\frac{\pi}{\phi}a}, 3, \frac{1}{2}\right) + e^{-i\frac{\pi}{2\phi}a} \Phi\left(-e^{-i\frac{\pi}{\phi}a}, 3, \frac{1}{2}\right) \right] \quad , \tag{2.53}$$

where $\Phi(z, s, \alpha) = \sum_{m=0}^{\infty} z^m/(m+\alpha)^s$ is the Lerch transcendent function [36]. Note that for

$a = 0$, recalling relations (2.31) and (2.40), we have

$$\begin{aligned}\langle t_D \rangle_{a=0} &= \frac{\phi^2}{\pi^3 D} \cdot \left[\Phi\left(-1, 3, \frac{1}{2}\right) + \Phi\left(-1, 3, \frac{1}{2}\right) \right] = \\ &= \frac{2\phi^2}{\pi^3 D} \Phi\left(-1, 3, \frac{1}{2}\right) = \frac{2\phi^2}{\pi^3 D} \sum_{m=0}^{\infty} \frac{(-1)^m}{\left(m + \frac{1}{2}\right)^3} = \frac{\phi^2}{2D} \quad , \end{aligned} \tag{2.54}$$

which gives the expected result for the case starting at zero, as supposed.

A plot of Eq. (2.53) is shown in Fig. 2.3 with $a \in [0, \phi]$.

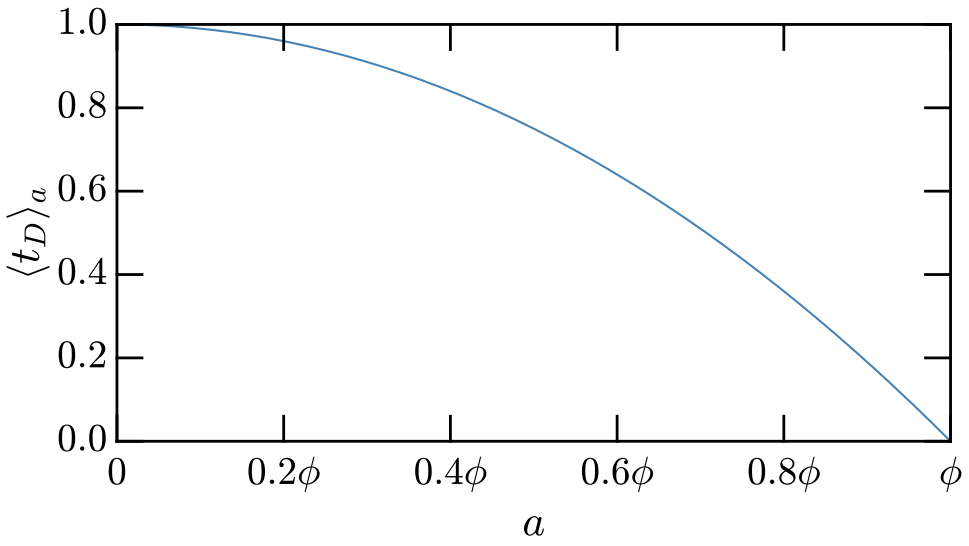

Figure 2.3: Plot of Eq. (2.53) in units of Brownian time in the domain $a \in [0, \phi]$.

With this set of equations, we can compare to numerical results. We present in Fig. 2.4 the histograms of the diffusion time distribution for a random walk with one particle in a pinned system, for both planar and random ICs, in a lattice of size $L = 181$ sites and closing angle $\phi = \theta_u = \pi$. For these simulations, we redefine step 6 of the algorithm presented in section 1.2, so that the minimum number of particles needed at a closing angle site in order for it to act as a sticky site is now zero. In other words, the closing angle site $n(\phi)$ was always active.

We compare the numerical results with the respective PDFs $f(t \mid a)$ and $f_R(t)$ for each IC. To calculate the theoretical points for each histogram bin, we integrate the appropriate PDF (Eq. (2.20) for the planar IC and Eq. (2.23) for the random IC) between the lower and upper edge values of the respective bin.

A table with the average values of the diffusion time $\langle t_D \rangle$ computed from the histograms is also presented in 2.1, along with the values predicted by Eqs. (2.41) and (2.48). In the last column we list the values of

the relative error ΔX, calculated using the following equation

$$\Delta X = \frac{X_E}{X_T} - 1 \quad , \tag{2.55}$$

with X_E and X_T being the experimental (or numerical) value and theoretical value, respectively.

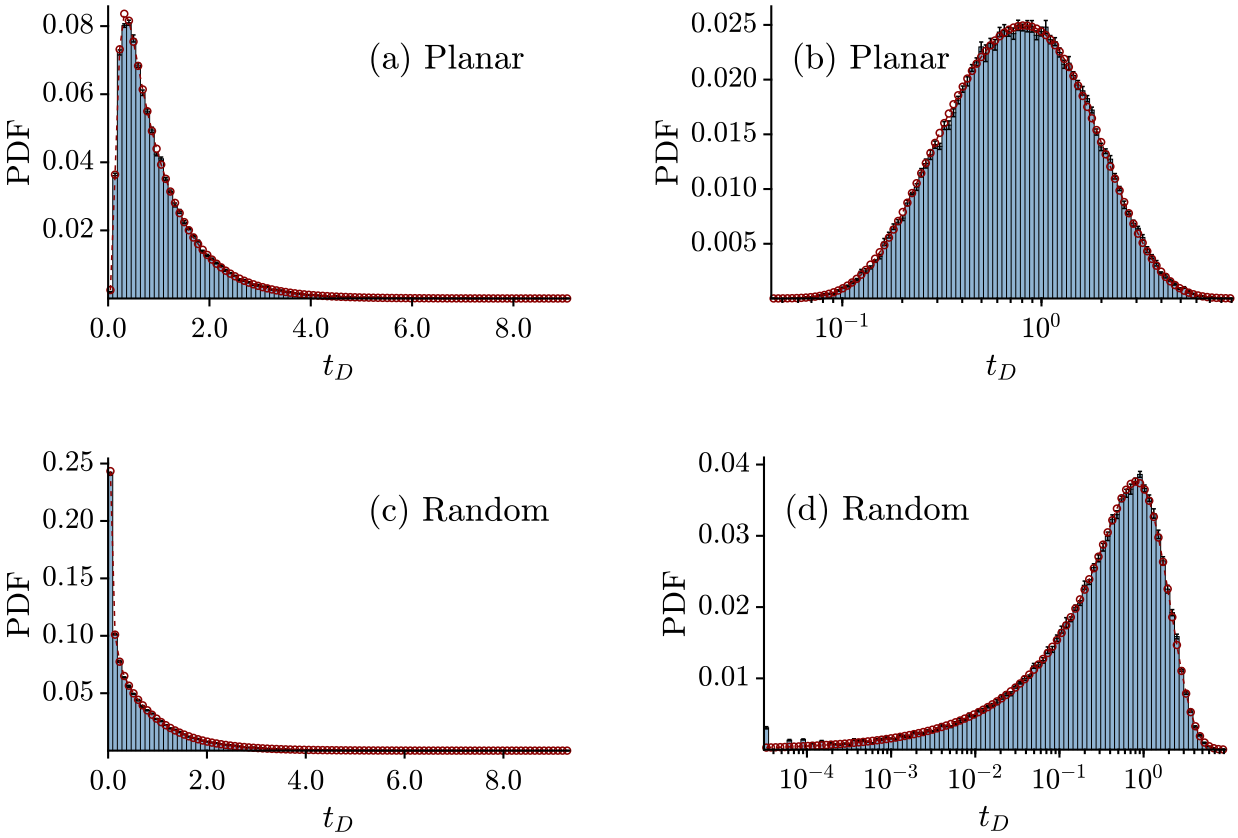

Figure 2.4: Normalized histograms with 100 bins each made from $N_S = 10^5$ independent samples obtained for the diffusion time t_D for a planar (a) and random (c) IC. Histograms (b) and (d) show the same results as (a) and (c), respectively, but with logarithmic binning. Times are in units of Brownian time. Simulations were performed with one particle in a pinned system with SC, except the sticky site was now positioned at the upper boundary, at $\phi = \pi$, and was always active. The theoretical points (in red) are the expected PDFs of the diffusion time for a planar and random IC, as given by Eqs. (2.20) with $a = 0$ and (2.23), respectively.

Table 2.1: Table with the simulated average values of the diffusion times $\langle t_D \rangle$ in Brownian time units, for a planar and random IC computed from the histograms of Fig. 2.4. The theoretical values of these time averages predicted by Eqs. (2.41) and (2.48), as well as the respective relative error, are also presented.

Initial condition	Simulated $\langle t_D \rangle$	Theoretical $\langle t_D \rangle$	Relative error ΔX (%)
Planar - Eq.(2.41)	1.011±0.003	1	1.10
Random - Eq.(2.48)	0.675±0.003	2/3=0.666...	1.25

From Fig. 2.4 we see that the histograms are well described by the respective expressions for the PDFs

(2.20) with $a = 0$ and (2.23). The numerical values for the averages $\langle t_D \rangle_0$ and $\langle t_D \rangle_R$ also show good agreement with theoretical predictions, considering their relative error ΔX, presented in the last column of table 2.1, is close to 1% for both ICs.

2.4 Convergence of the relative error for increasing lattice size and number of simulations

In this section we analyse how the relative error ΔX, as defined in Eq. (2.55), for the planar IC, behaves as we increase the number of sites L and the sample size N_S.

Equations (2.41) and (2.48) were derived assuming a system with continuous time and space, so they should only predict the values with good accuracy if the number of lattice sites L is large enough. We should expect that the relative error ΔX between the numerical and theoretical values of the average diffusion time $\langle t_D \rangle$ should converge to a small value, ideally zero. We choose the planar IC average diffusion time $\langle t_D \rangle_0$, with a expected value given by Eq. (2.41), as our test case.

The results for the absolute value of ΔX as a function of L are presented in Fig. 2.5. Given an uncertainty σ_E in the numerical value X_E, the uncertainty $\sigma_{\Delta X}$ in ΔX is then calculated by error propagation,

$$\sigma_{\Delta X} = \sigma_E \left| \frac{\partial \Delta X}{\partial X_E} \right| = \frac{\sigma_E}{|X_T|} = \frac{1}{|X_T|} \sqrt{\frac{\langle X_E^2 \rangle - \langle X_E \rangle^2}{N_S}} \quad . \tag{2.56}$$

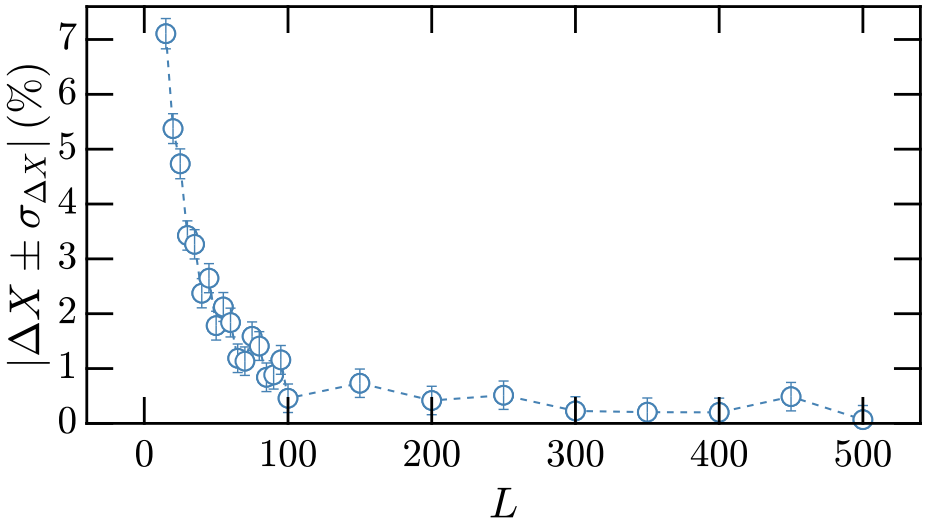

Figure 2.5: Absolute value of the relative error ΔX, Eq. (2.55), between the simulated $\langle t_D \rangle_0$ and the value predicted by Eq. (2.41) as a function of the number of lattice sites $L \in [5, 500]$. The uncertainty $\sigma_{\Delta X}$ in the values of ΔX was computed by propagating the uncertainty σ_E associated with the numerical value for the average diffusion time, using Eq. (2.56). Apart from L, the remaining parameters were identical to the ones used in the numerical analysis of section 2.3. The averages for the diffusion times, from which we then compute ΔX and $\sigma_{\Delta X}$, were calculated from $N_S = 10^5$ independent samples.

From Fig. (2.5) we see that the absolute value of the relative error ΔX converges rather quickly, at about $L = 100$, oscillating around 0.5% for larger values of L. This seems to indicate that, for a sample size of $N_S = 10^5$, the systematic error introduced by the discretization of the space is reduced to a minimum starting at these values of L. The remaining discrepancy may be due to random error sources, such as the pseudo randomness associated with the random number generator.

Next we test if our statistical averages converges for an increasing number of independent samples N_S. In Fig. 2.6 we plot the absolute value of relative error ΔX, again for a planar IC, as function of the sample size N_S, for several values of L.

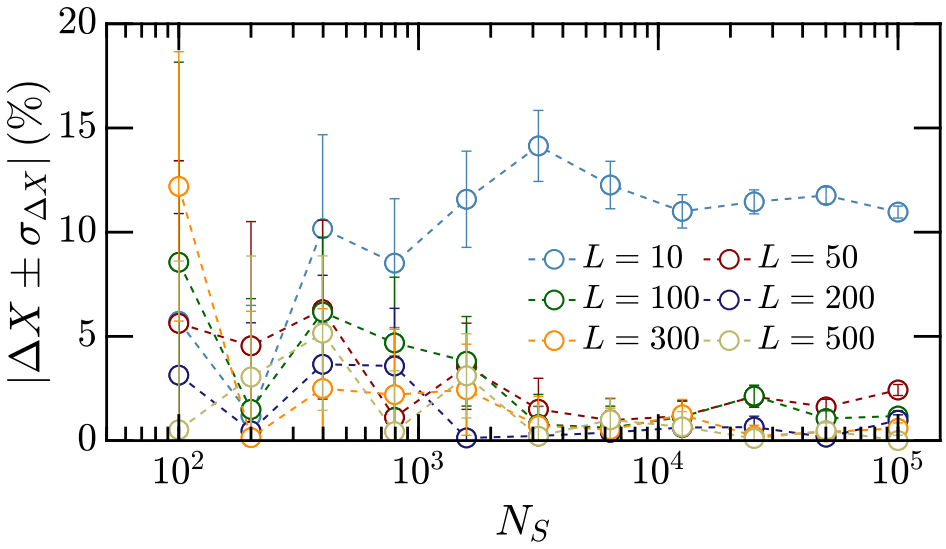

Figure 2.6: Absolute value of the relative error ΔX, Eq. (2.55), between the simulated $\langle t_D \rangle_0$ and the value predicted by Eq. (2.41) as a function of the number of independent samples $N_S \in [10^2, 10^5]$, for several values of the number of lattice sites L. The uncertainty $\sigma_{\Delta X}$ in the values of ΔX was computed, as in the case of Fig. 2.5, using Eq. (2.56). For each L, all simulations were performed under conditions identical to the ones used in the analysis presented in Fig. 2.5.

For better readability, we plot the absolute value of the best estimate of ΔX and its uncertainty $\sigma_{\Delta X}$ in separate graphs in Fig. 2.7.

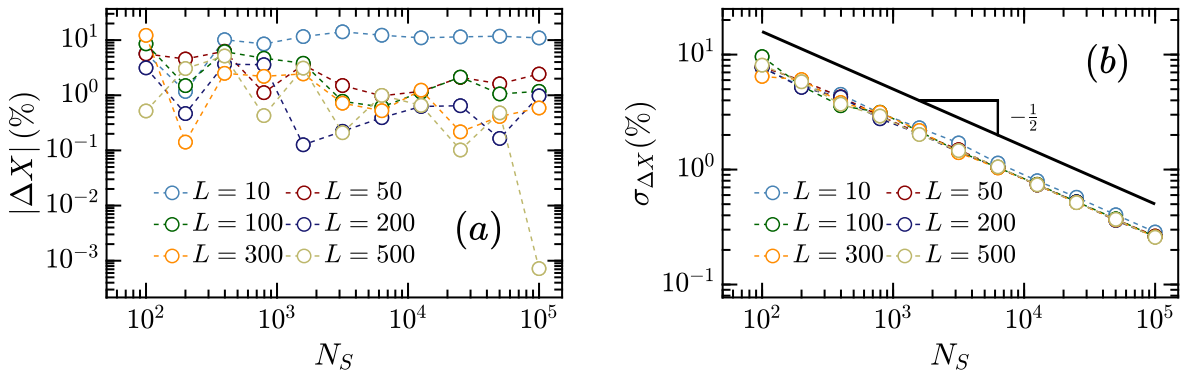

Figure 2.7: Same results of Fig. 2.6 but with the absolute value of the best estimate of ΔX (a) and its uncertainty $\sigma_{\Delta X}$ (b) plotted in separate graphs. The black line with slope $-1/2$ in (b) is a visual aid to indicate that the data varies as expected from Eq. (2.56) as the sample size grows.

These results indicate that, as we increase the number of samples N_S, the averages of our statistical quantity do converge to a defined value, as we can infer from the decreasing uncertainty $\sigma_{\Delta X}$ as N_S grows. In Fig. 2.7.b, we clearly see that our uncertainty in our best estimate for the average of our statistical quantity varies with N_S as $\sigma_{\Delta X} \propto 1/\sqrt{N_S}$, in accordance with Eq. (2.56) for the standard error.

In general, the analysis of this section shows that, indeed, by increasing the number of lattice sites L we can improve the accuracy of our results, by reducing the error due to the discretization, while our precision

improves as we increase the number of samples N_S and the effect of the law of large numbers gets progressively more preponderant.

Chapter 3

Folding statistics

In this chapter we study the folding statistics of regular pyramid nets. In section 3.1 we analyse the binding times as a function of the number of lateral faces N. We derive a set of equations and compare them with numerical results. In section 3.2 we analyse the binding times as a function of the closing angle ϕ. In section 3.3 we quantify the effect of the particle-particle interactions present in the interacting region. Additionally, we study how the functional dependence of the binding times on N changes if binding can only occur between neighbouring lateral faces. In section 3.4 we study how the binding times behave for different ICs. In section 3.5 we quantify the propensity for the nets to fold into the target structure, by numerically computing the fraction of samples that self-fold free of defects.

3.1 Dependence of the binding times on the number of faces

In this section we study the functional dependence of the average folding time $\langle T \rangle$ with the number of lateral faces N. As previously stated, the dynamics of the folding can be decomposed into two distinguishable sets of Brownian processes. The first binding event, occurring at the first binding time T_F, is well described by a set of competing 2D Brownian processes. In turn, the subsequent binding events are each described by a 1D Brownian process. The last binding event concludes at the folding time T. The time interval between the first and last binding event is defined as the last binding time $T_L = T - T_F$. We now proceed to study the functional dependency on N of each one of these times separately. We consider only a pinned system to focus on the folding time.

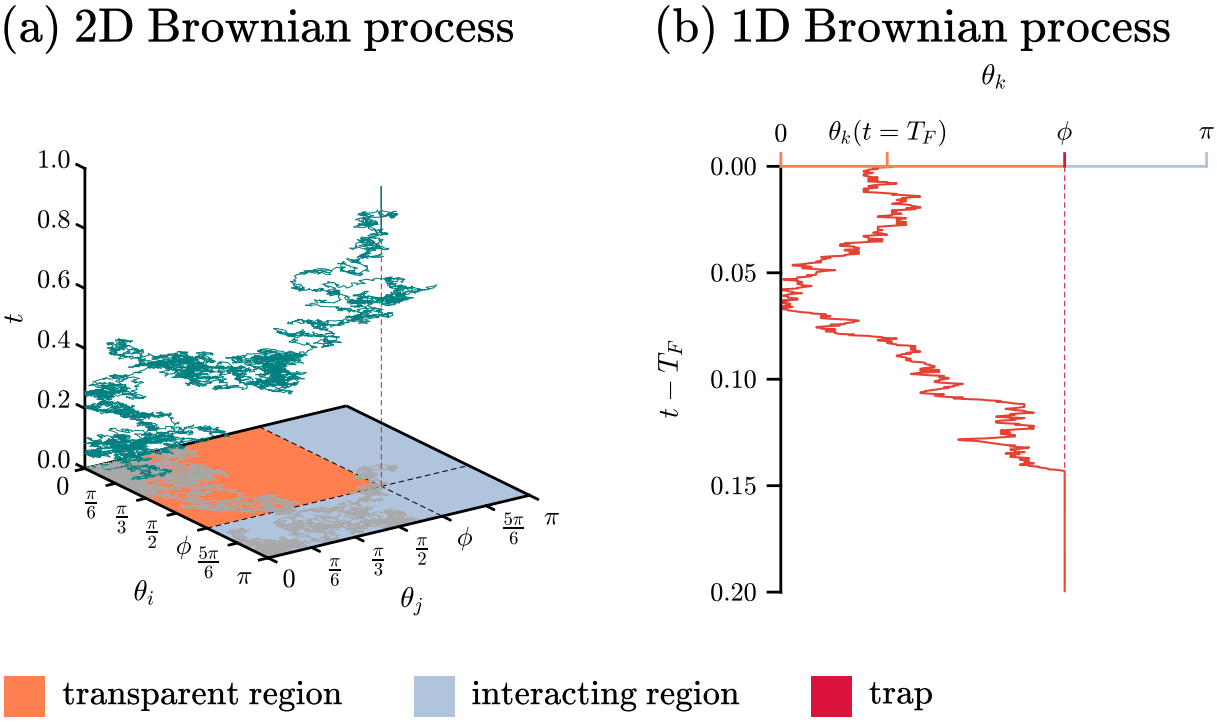

Figure 3.1: (a) Configuration of the 2D Brownian process for the first binding event in a pinned system (for a planar IC). A pair of particles i and j, $\forall i, j$, diffuse with coordinates (θ_i, θ_j) in a square domain of size $\pi \times \pi$ until they hit the trap at $(\phi \pm {}^{\Delta\phi}/_{2}, \phi \pm {}^{\Delta\phi}/_{2})$. The trapping region (coloured red) is a square domain with side length equal to the closing angle region size $\Delta\phi$. The trajectory of the pair is projected onto the plane $\theta_i\theta_j$ (coloured grey). (b) After the first binding event, the subsequent ones can be mapped to a 1D Brownian process. The remaining particles k, $\forall k \backslash \{i, j\}$, diffuse with coordinates θ_k, starting at θ_k $(t = T_F)$, in a line of length π until they reach the trap at $\phi \pm {}^{\Delta\phi}/_{2}$. The regions at the right of the trapping region are the interacting region (light blue), where particle-particle interactions may occur.

3.1.1 First binding time as a function of N

The first binding event is defined by a set of competing 2D Brownian processes, each with angular coordinates (θ_i, θ_j) of the pair of lateral faces i and j and a trap at $(\phi \pm {}^{\Delta\phi}/_{2}, \phi \pm {}^{\Delta\phi}/_{2})$ (Fig. 3.1a). The number of 2D Brownian processes is equal to the number of possible pairs of lateral faces. For N lateral faces, since we assume that binding can occur between any pair of faces, independently of whether they are adjacent or not, this number is given by the binomial coefficient $\binom{N}{2} = {}^{N\cdot(N-1)}/_{2}$. Associated with the 2D Brownian processes we have a set $\left\{t_F^{(1)}, t_F^{(2)}, ..., t_F^{(N\cdot(N-1)/2)}\right\}$ of FPTs for reaching the trap. These times are a set of random variables, characterized by some PDF $f_F(t)$. The first binding time is set by the fastest of the ${}^{N\cdot(N-1)}/_{2}$ FPTs, so

$$T_F = \min\left\{t_F^{(1)}, t_F^{(2)}, ..., t_F^{(N\cdot(N-1)/2)}\right\} \quad . \tag{3.1}$$

The probability $S_F(t)$ that some particular measurement drawn from $f_F(t)$ yields a value larger than t is

$$S_F(t) = \int_t^\infty f_F(t')\,dt' \quad . \tag{3.2}$$

From the theory of order statistics [38], we can estimate the PDF $F_F(t)$ that some measurement $t_F^{(i)}$ picked from the set of the $N\cdot(N-1)/2$ FPTs has a value of t, and is no larger than the remaining $N\cdot(N-1)/2-1$ times,

$$F_F(t) = \frac{N\cdot(N-1)}{2} f_F(t)\cdot[S_F(t)]^{\frac{N\cdot(N-1)}{2}-1} \quad . \tag{3.3}$$

The PDF (3.3) is the distribution for the first binding time T_F. Provided that $t_F^{(i)} = t$, the term with square brackets is the probability that $t_F^{(i)}$ is the fastest of the $N\cdot(N-1)/2$ FPTs, thus $t_F^{(i)} = T_F$ by definition, while the coefficient $N\cdot(N-1)/2$ accounts for the fact that any of those times may be the fastest one.

The average first binding time is then computed by multiplying the PDF (3.3) by t and integrating over the whole positive time domain,

$$\langle T_F(N)\rangle = \frac{N\cdot(N-1)}{2}\int_0^\infty t f_F(t)\cdot\left[\int_t^\infty f_F(t')\,dt'\right]^{\frac{N\cdot(N-1)}{2}-1} dt \quad . \tag{3.4}$$

For a planar IC, calculations in previous works [39] suggest that Eq. (3.4) should be inversely proportional to the logarithm of the number of 2D Brownian processes,

$$\langle T_F(N)\rangle \propto \frac{1}{\ln\left(\frac{N\cdot(N-1)}{2}\right)} \quad . \tag{3.5}$$

We expect the average first binding time to decrease logarithmically with the number of lateral faces, when the folding starts from a flat template. It is noteworthy that we assumed that the 2D Brownian processes are completely uncorrelated. For example, for three lateral faces, the three 2D Brownian processes with coordinates (θ_0,θ_1), (θ_0,θ_2) and (θ_1,θ_2) are seemingly correlated by the coordinates shared between them. Despite this, it has been shown that the dynamics are still consistent with Brownian behaviour [13], and correlations can therefore be neglected. Furthermore, this approach also ignores the particle-particle interactions that may occur in the interacting region. We will see in the later section 3.3, however, that these interactions have little influence on the results.

We present in Fig. 3.2 the numerical results for $\langle T_F\rangle$ as a function of N for a planar IC. We also fit the data to the following equation

$$y(x) = \tau_{F_0} + \tau_F x \quad , \tag{3.6}$$

with $x = {}^{1}/_{\ln(N\cdot(N-1)/2)}$, where τ_F and τ_{F_0} are fitting parameters. The point $N = 2$ was excluded from the fitting since the function is undefined for that value.

To fit the data we use the method of least squares (more details can be found in appendix section A.2). For a set of N_P data points $\{(x_1, y_1), (x_2, y_2), ..., (x_{N_P}, y_{N_P})\}$ where the values x_i are known exactly and y_i each have an associated uncertainty σ_i, we estimate the fitting parameters a_0 and a_1 that give the straight line that best fits the data,

$$y_i = a_0 + a_1 x_i \quad , \tag{3.7}$$

by computing

$$a_0 = \frac{\sum_{i=1}^{N_P}\left(\frac{x_i^2}{\sigma_i^2}\right)\sum_{i=1}^{N_P}\left(\frac{y_i}{\sigma_i^2}\right) - \sum_{i=1}^{N_P}\left(\frac{x_i}{\sigma_i^2}\right)\sum_{i=1}^{N_P}\left(\frac{x_i y_i}{\sigma_i^2}\right)}{\Delta} \quad , \tag{3.8}$$

$$a_1 = \frac{-\sum_{i=1}^{N_P}\left(\frac{x_i}{\sigma_i^2}\right)\sum_{i=1}^{N_P}\left(\frac{y_i}{\sigma_i^2}\right) + \sum_{i=1}^{N_P}\left(\frac{1}{\sigma_i^2}\right)\sum_{i=1}^{N_P}\left(\frac{x_i y_i}{\sigma_i^2}\right)}{\Delta} \quad , \tag{3.9}$$

where

$$\Delta = \sum_{i=1}^{N_P}\left(\frac{1}{\sigma_i^2}\right)\sum_{i=1}^{N_P}\left(\frac{x_i^2}{\sigma_i^2}\right) - \left(\sum_{i=1}^{N_P}\frac{x_i}{\sigma_i^2}\right)^2 \quad . \tag{3.10}$$

The uncertainties σ_{a_0} and σ_{a_1} of a_0 and a_1, respectively, are calculated by error propagation from the uncertainties σ_i of the corresponding y_i values, so

$$\sigma_{a_0} = \sqrt{\frac{\sum_{i=1}^{N_P}\frac{x_i^2}{\sigma_i^2}}{\Delta}} \quad , \tag{3.11}$$

$$\sigma_{a_1} = \sqrt{\frac{\sum_{i=1}^{N_P}\frac{1}{\sigma_i^2}}{\Delta}} \quad . \tag{3.12}$$

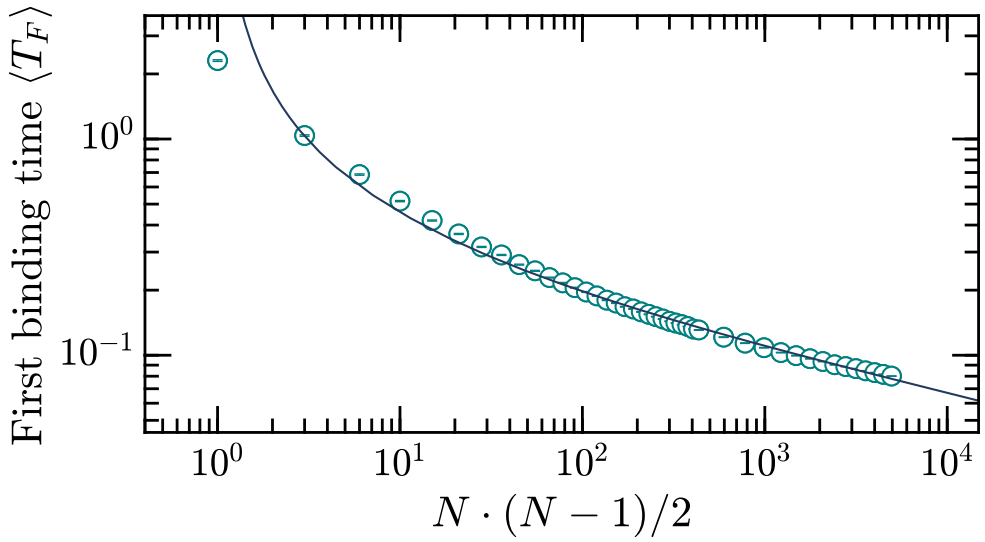

Figure 3.2: Numerical results for the average first binding time $\langle T_F \rangle$ as a function of the number of lateral faces $N \in [2, 100]$ for a planar IC. The solid line is given by Eq. (3.6) with fitting parameters $\tau_{F_0} = -0.0636 \pm 0.0003$ and $\tau_F = 1.203 \pm 0.002$, obtained from a least square fit of the simulation data. The simulations were performed with SC. Results are averages from $N_S = 10^4$ independent samples.

These results show that the average first binding time decays logarithmically with the number of lateral faces for a planar IC, as predicted from Eq. (3.5).

3.1.2 Last binding time as a function of N

After the first binding, the remaining $N-2$ lateral faces only have to reach the closing angle ϕ one by one to attach. Thus, we have a set of $N-2$ 1D Brownian processes, with an associated set $\left\{t_L^{(1)}, t_L^{(2)}, ..., t_L^{(N-2)}\right\}$ of FPTs for reaching $\phi \pm {}^{\Delta\phi}/_2$ (Fig. 3.1b). These times are again a set of random variables, drawn from some PDF $f_L(t)$. The set in non-decreasing order of these FPTs, added by the first binding time T_F, is the set of binding times $\{t_2, t_3, ..., t_{N-2}\}$ excluding t_1. The last binding time is set by the longest of these $N-2$ FPTs,

$$T_L = \max\left\{t_L^{(1)}, t_L^{(2)}, ..., t_L^{(N-2)}\right\} \quad . \tag{3.13}$$

As before, the probability $S_L(t)$ that a measurement drawn from $f_L(t)$ yields a value larger than t is

$$S_L(t) = \int_t^\infty f_L(t')\, dt' \quad . \tag{3.14}$$

Naturally, the probability that the value is not larger than t is $1 - S_L(t)$.

Proceeding in a similar way to the first binding time, from the theory of order statistics, we estimate the distribution $F_L(t)$ of the last binding time to be

$$F_L(t) = (N-2)\, f_L(t) \cdot [1 - S_L(t)]^{(N-2)-1} \quad , \tag{3.15}$$

where now, provided that some measurement $t_L^{(i)}$ picked from the set of FPTs has a value of t, the term with square brackets is the probability that $t_L^{(i)}$ is the slowest of the $N-2$ FPTs, thus $t_L^{(i)} = T_L$ by definition.

The average last binding time $\langle T_L \rangle$ is then computed by

$$\langle T_L(N) \rangle = (N-2) \int_0^\infty t f_L(t) \cdot \left[1 - \int_t^\infty f_L(t')\, dt'\right]^{(N-2)-1} dt \quad . \tag{3.16}$$

If we assume that the $N-2$ lateral faces all have an outer angle $\theta_i < \phi$ at the first binding moment, the PDF $f_L(t)$ would be the distribution of a first-passage time as discussed in section 2.3. Its specific form depends on the distribution of the angles θ_i at $t = T_F$. For a planar IC, if the first two lateral faces attach quickly, we expect that the remaining $N-2$ still have an outer angle close to $\theta_i = 0$. In this case, $f_L(t)$ would be approximately given by Eq. (2.20) with $a = 0$. If the first binding event takes too long, however, we assume that the distribution of the outer angles of the remaining $N-2$ lateral faces is close to a uniform one, in which case $f_L(t)$ could be approximated by Eq. (2.23). In both cases, the eigenmodes of the PDFs decay exponentially with time, and in the long-time limit the eigenmode with $m = 0$ dominates [35]. Thus we approximate the PDF $f_L(t)$ to an exponential distribution,

$$f_L(t) \simeq \frac{1}{\tau_L} e^{-\frac{t}{\tau_L}} \quad , \tag{3.17}$$

where $\tau_L = {}^{4\phi^2}/_{\pi^2 D}$ is the characteristic decay time of the eigenmode $m = 0$.

Plugging PDF (3.17) in (3.16) and computing the integral inside the square brackets we get

$$\langle T_L(N) \rangle \simeq \frac{N-2}{\tau_L} \int_0^\infty t e^{-\frac{t}{\tau_L}} \cdot \left[1 - e^{-\frac{t}{\tau_L}}\right]^{(N-2)-1} dt \quad . \tag{3.18}$$

Making the change of variable $u = 1 - e^{-\frac{t}{\tau_L}}$ we have

$$\langle T_L(N) \rangle \simeq -(N-2)\, \tau_L \int_0^1 \ln(1-u)\, u^{(N-2)-1} du \quad . \tag{3.19}$$

We can use the Mercator series,

$$\ln(1+x) = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} x^m \quad \text{, for } x \in]-1,1], \tag{3.20}$$

with $x = -u$, to perform the integration,

$$\langle T_L(N) \rangle \simeq (N-2)\,\tau_L \sum_{m=1}^{\infty} \frac{1}{m} \int_0^1 u^{(N-2)-1+m} du = \tau_L \sum_{m=1}^{\infty} \frac{N-2}{m \cdot (N-2+m)} \quad . \tag{3.21}$$

Notice that

$$\frac{N-2}{m \cdot (N-2+m)} = \frac{N-2+m-m}{m \cdot (N-2+m)} = \frac{1}{m} - \frac{1}{N-2+m} \quad , \tag{3.22}$$

and so

$$\langle T_L(N) \rangle \simeq \tau_L \sum_{m=1}^{\infty} \left(\frac{1}{m} - \frac{1}{N-2+m} \right) \quad . \tag{3.23}$$

Expanding the summation, defining $\mathcal{N} \equiv N-2$,

$$\sum_{m=1}^{\infty} \left(\frac{1}{m} - \frac{1}{\mathcal{N}+m} \right) = 1 - \frac{1}{\mathcal{N}+1} + \frac{1}{2} - \frac{1}{\mathcal{N}+2} + \frac{1}{3} - \frac{1}{\mathcal{N}+3} + \frac{1}{4} - \frac{1}{\mathcal{N}+4} + \ldots \quad , \tag{3.24}$$

one can check that, for a given $\mathcal{N}$, this series is equivalent to summing the term $1/m$ up to that $\mathcal{N}$, as all other terms cancel out. Therefore,

$$\langle T_L(N) \rangle \simeq \tau_L \sum_{m=1}^{\infty} \left(\frac{1}{m} - \frac{1}{N-2+m} \right) = \tau_L \sum_{m=1}^{N-2} \frac{1}{m} \quad . \tag{3.25}$$

Using the definition of the Euler-Mascheroni constant γ,

$$\gamma = \lim_{N \to \infty} \left[-\ln(N-2) + \sum_{m=1}^{N-2} \frac{1}{m} \right] \quad , \tag{3.26}$$

we can reformulate Eq. (3.25) into

$$\langle T_L(N) \rangle \simeq \tau_L \cdot [\ln(N-2) + \gamma] \quad \text{, for large } N. \tag{3.27}$$

From this equation we expect that, for a large number of lateral faces, the average last binding time should grow logarithmically as N increases. We present the numerical results of $\langle T_L \rangle$ as a function of N in Fig. 3.3. Alongside the numerical data we also plot Eq. (3.27).

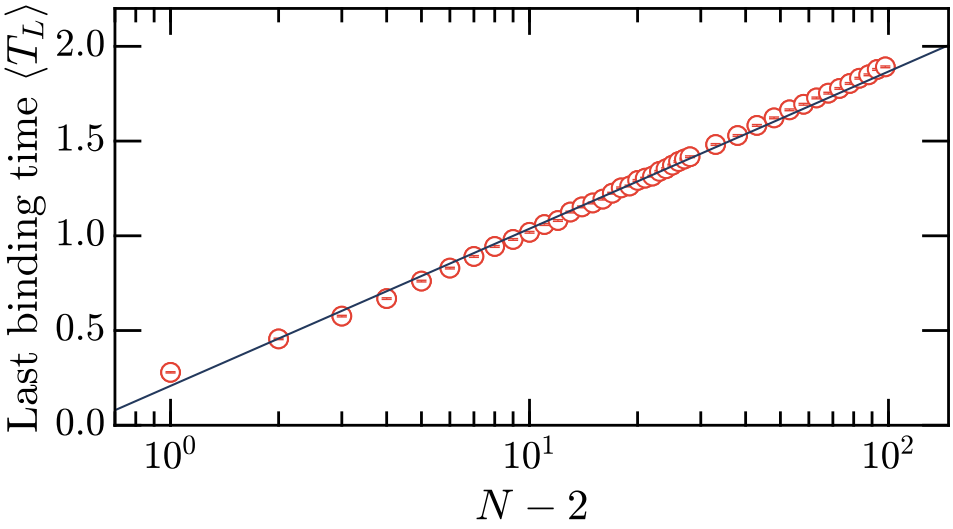

Figure 3.3: Numerical results for the average last binding time $\langle T_L \rangle$ as a function of the number of lateral faces $N \in [3, 100]$ for a planar IC. The solid line is given by Eq. (3.27) in units of Brownian time, where $\tau_L = {}^{4\phi^2}/_{\pi^2 D}$, with $\phi = {}^{2\pi}/_{3}$ and $D = D_0$. The simulations were performed with SC. Results are averages from $N_S = 10^4$ independent samples.

Numerical results seem to agree with the theoretical predictions, with the data following the line given by Eq. (3.27).

3.1.3 Folding time as a function of N

The folding time T is the time it takes for all lateral faces to bind. By definition, it is the sum of the first and last binding times. Here we present the numerical results of its average $\langle T \rangle$ as a function of N in Fig. 3.4. Since $\langle T \rangle = \langle T_F \rangle + \langle T_L \rangle$, we expect that the functional dependence of $\langle T \rangle$ with N is given by the sum of Eqs. (3.6) and (3.27),

$$\langle T\,(N) \rangle \simeq \frac{\tau_F}{\ln\left(\frac{N\cdot(N-1)}{2}\right)} + \tau_L \ln\left(N-2\right) + \tau_0 \quad , \tag{3.28}$$

where $\tau_0 = \tau_{F_0} + \gamma\tau_L$, and τ_F, τ_{F_0} and τ_L are the same parameters as those introduced in the previous subsections 3.1.1 and 3.1.2. We plot Eq. (3.28) alongside the data in Fig. 3.4.

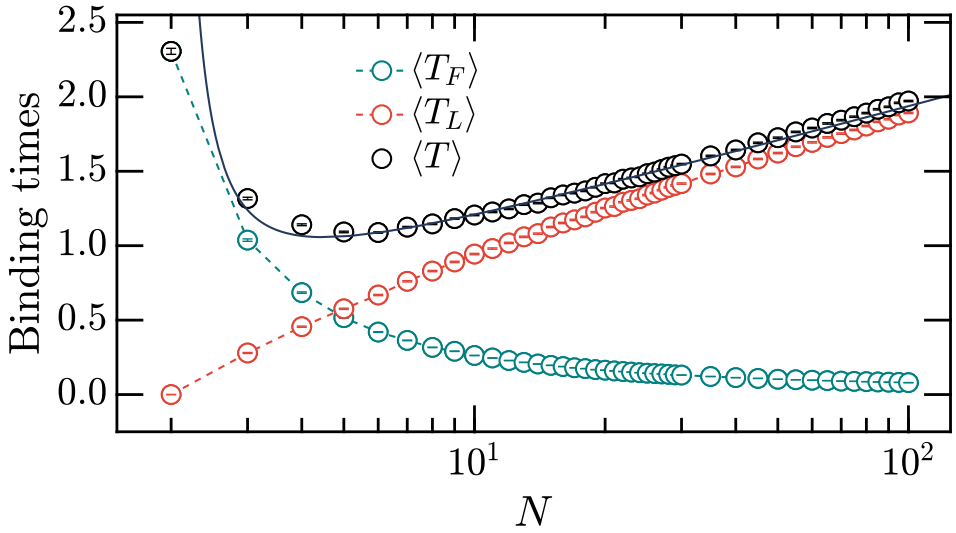

Figure 3.4: Numerical results for the average binding times $\langle T_F \rangle$, $\langle T_L \rangle$ and average folding time $\langle T \rangle$ as a function of the number of lateral faces $N \in [2, 100]$ for a planar IC. The results for $\langle T_F \rangle$ and $\langle T_L \rangle$ are the same as those presented in Figs. (3.2) and (3.3), respectively. The solid line is given by Eq. (3.28) in units of Brownian time, where the parameters τ_F, τ_{F_0} and τ_L have the same values as those used in Figs. (3.2) and (3.3). The simulations were performed with SC. Results are averages from $N_S = 10^4$ independent samples.

From Fig. (3.4) we see that the contrasting limiting behaviour of the first and last binding times lead to a minimum around $N = 5$ and $N = 6$. This shows that there is an optimal number of lateral faces for fast self-folding. While the theoretical formula is too simplified to accurately predict the minimum, it does however describe well the data for larger values of N. The numerical results are consistent with previous works based on Langevin molecular dynamics simulations [13].

3.1.4 Folding time for a random IC

In this subsection we compare the average first and last binding times, as well as the total folding time, between a planar and random IC. We plot the numerical results for these three mean times for a random IC in Fig. 3.5, alongside the results for the planar case, the same data presented in Fig. 3.4.

For the average first binding time $\langle T_F(N) \rangle_R$ with a random IC, we assume that the FPT distribution $f_F(t)$ in 2D is approximately exponential,

$$f_F(t) \simeq \frac{1}{\tau_{F,R}} e^{-\frac{t}{\tau_{F,R}}} \quad , \tag{3.29}$$

with a characteristic decay time $\tau_{F,R}$.

With this distribution, Eq. (3.4) reads, after computing the integral inside the square brackets, which

evaluates to $e^{-t/\tau_{F,R}}$,

$$\langle T_F(N)\rangle_R \simeq \frac{N\cdot(N-1)}{2\tau_{F,R}}\int_0^\infty te^{-t\frac{N\cdot(N-1)}{2\tau_{F,R}}}dt \quad . \tag{3.30}$$

The integral in Eq. (3.30) is of the form (2.28) with $s = {}^{N\cdot(N-1)}/_{2\tau_{F,R}}$, and so the average first binding time for a random IC should depend on the number of lateral faces as

$$\langle T_F(N)\rangle_R \simeq \frac{2\tau_{F,R}}{N\cdot(N-1)} \quad , \tag{3.31}$$

decaying as N^{-2} for a large number of lateral faces.

To corroborate this, we use the linear least squares method to fit the data for the average first binding time with a random IC presented in Fig. 3.5a to the following model,

$$y(x) = \tau_{F_0,R} + \tau_{F,R}x \quad , \tag{3.32}$$

with $x = {}^{2}/_{N\cdot(N-1)}$ and linear fitting parameters $\tau_{F_0,R}$ and $\tau_{F,R}$.

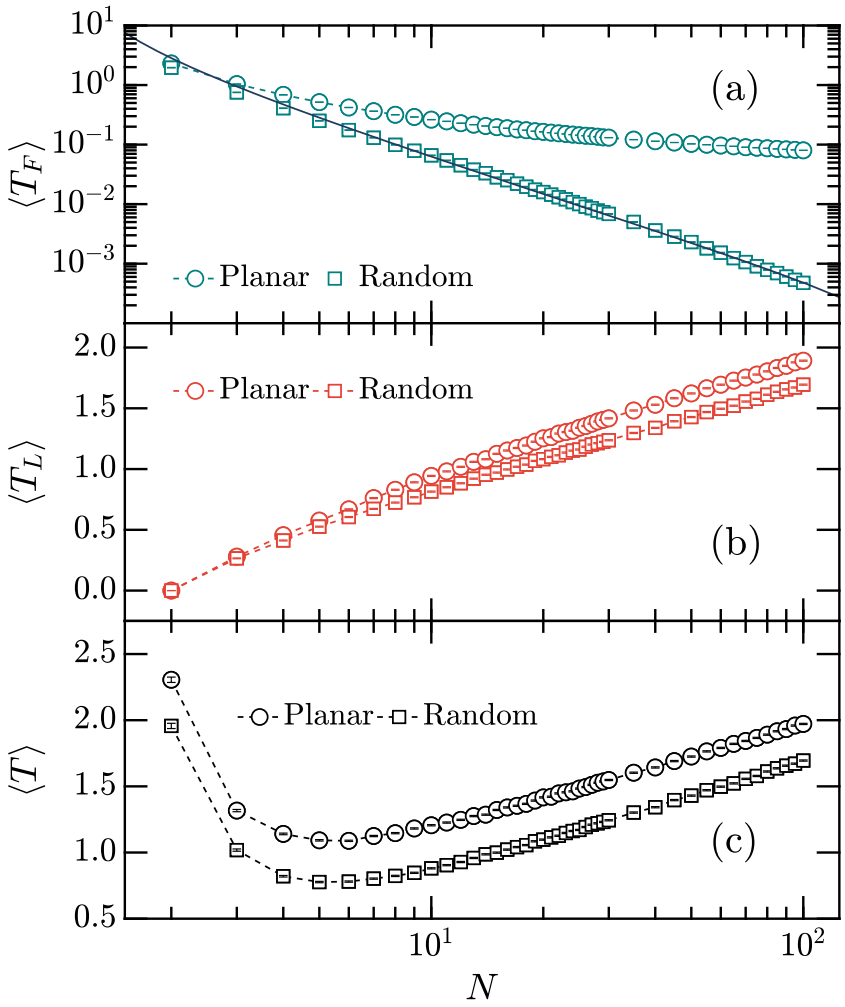

Figure 3.5: Numerical results for the average binding times $\langle T_F \rangle$ (a), $\langle T_L \rangle$ (b) and average folding time $\langle T \rangle$ (c) as a function of the number of lateral faces $N \in [2, 100]$ for a random IC (square points), alongside the respective counterpart for the planar IC (circle points, the same data plotted in Fig. 3.4). The solid line in (a) is given by Eq. (3.32) in units of Brownian time, with fitted parameters $\tau_{F_0,R} = -(92 \pm 4) \cdot 10^{-6}$ and $\tau_{F,R} = 2.835 \pm 0.007$ obtained from a least square fit of the numerical data. The simulations were performed with SC. Results are averages from $N_S = 10^4$ independent samples.

We see from Fig. 3.5a that the average first binding time varies with N as expected from Eq. (3.31), decaying with the square of N for a large number of lateral faces, in contrast with the slower logarithmic decay for the planar IC. The average last binding time is also slightly faster for all N for a random IC when compared with the planar one. Despite these functional differences of the average binding times between the two ICs, the distinction of the total average folding time $\langle T \rangle$ between those same ICs seems to be only quantitative, with overall faster times for the random IC, while the minimum remains unshifted.

3.2 Dependence of the binding times on the closing angle

In this section we study how the closing angle ϕ influences the average folding time. Changing the value of these parameter could be achieved in practice, for example, by varying the slant height of the lateral faces. We assume however that the angular diffusion coefficient of these faces remains independent of ϕ. In Fig. 3.6a we plot the numerical results for the average folding time $\langle T \rangle$ as a function of N for three different values of ϕ. As we have seen in section 3.1, as N grows, the contribution of the last binding time $\langle T_L \rangle$ for the total folding time $\langle T \rangle$ dominates over the first binding time $\langle T_F \rangle$. In this regime, if we assume that, after the first binding, all $N-2$ particles are at the left of the trapping region, so $\theta_i\,(t = T_F) < \phi$ (considering $\Delta\phi \simeq 0$), and distributed approximately uniformly over the domain $\theta \in [0, \phi[$, we should expect $\langle T \rangle$ to scale with ϕ^2, in accordance with result (2.48) for the average time a diffusing particle in a line with domain $\theta \in [0, \phi]$ takes to reach ϕ, starting with a uniform spatial distribution. In Fig. 3.6b, we rescale the data of Fig. 3.6a by the square of the closing angle, ϕ^2.

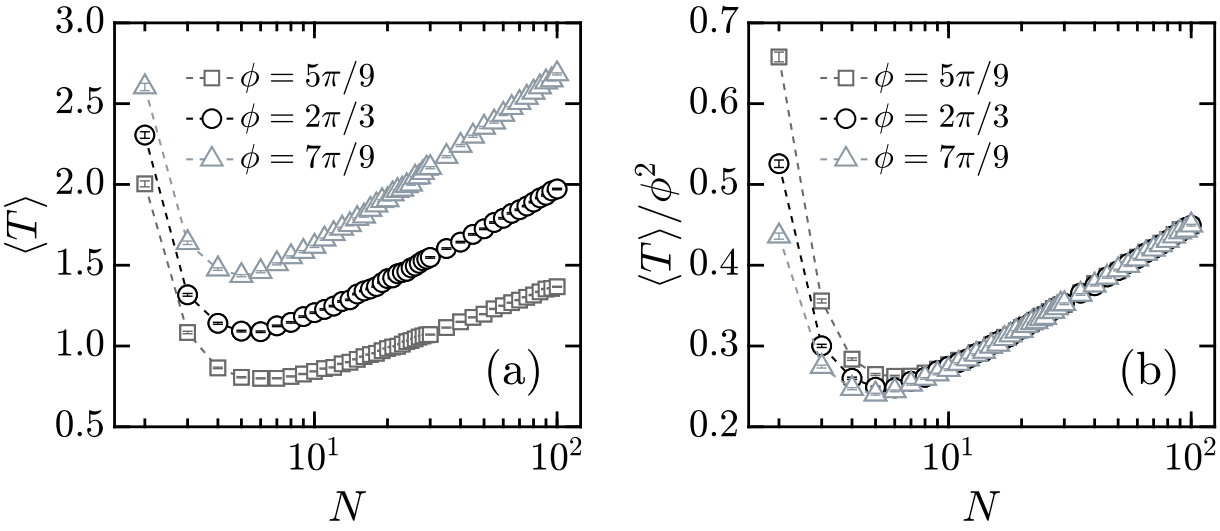

Figure 3.6: (a) Numerical results for a planar IC of the average folding time $\langle T \rangle$ as a function of the number of lateral faces $N \in [2, 100]$ for three values of the closing angle $\phi = \{5\pi/9, 2\pi/3, 7\pi/9\}$. (b) The same data as in (a) but rescaled by the square of the closing angle ϕ^2. A data collapse is observed for larger values of N indicating that $\langle T \rangle$ scales as ϕ^2 in this regime. Apart from the closing angle, the simulations were performed with SC. Results are averages from $N_S = 10^4$ independent samples.

From Fig. 3.6b we see that, for sufficiently large $N \gtrsim 10$, all three curves with different ϕ overlap, indicating that the average folding time $\langle T \rangle$ has the expected scaling with this parameter in this regime. The case is different for a low value of N, where overlapping is no longer observed, therefore the power law relation of $\langle T \rangle$ with ϕ^2 is no longer valid.

To better understand these functional relationships, we plot in Fig. 3.7 the numerical results for the average

first (a) and last (b) binding times, as well as the folding time (c), as a function of ϕ, for several values of N. We additionally fit the data sets $\{\phi_i, y_i\}$ of Figs. 3.7a and 3.7b (where y_i are the numerical values of the respective average binding time) to a power law,

$$y_i = A\phi_i^k \quad , \tag{3.33}$$

with parameters A and k. To estimate these parameters, we use again the method of least squares, by finding the linear regression that best fits the data sets $\{x_i, y_i'\}$,

$$y_i' = a_0 + a_1 x_i \quad , \tag{3.34}$$

with $y_i' = \ln(y_i)$, $x_i = \ln(\phi_i)$, and where $a_0 = \ln(A)$ and $a_1 = k$ are the fitting parameters, given by Eq. (3.8) and (3.9), respectively. The uncertainties σ_i of the values y_i propagate to the uncertainties σ_i' of y_i' as

$$\sigma_i' = \sigma_i \left| \frac{\partial y_i'}{\partial y_i} \right| = \frac{\sigma_i}{y_i} \quad . \tag{3.35}$$

The uncertainty σ_A of the parameter A is then calculated from σ_{a_0} using

$$\sigma_A = \sigma_{a_0} \left| \frac{\partial A}{\partial a_0} \right| = \sigma_{a_0} e^{a_0} \quad . \tag{3.36}$$

The obtained values for the parameters A and k are gathered in Table 3.1.

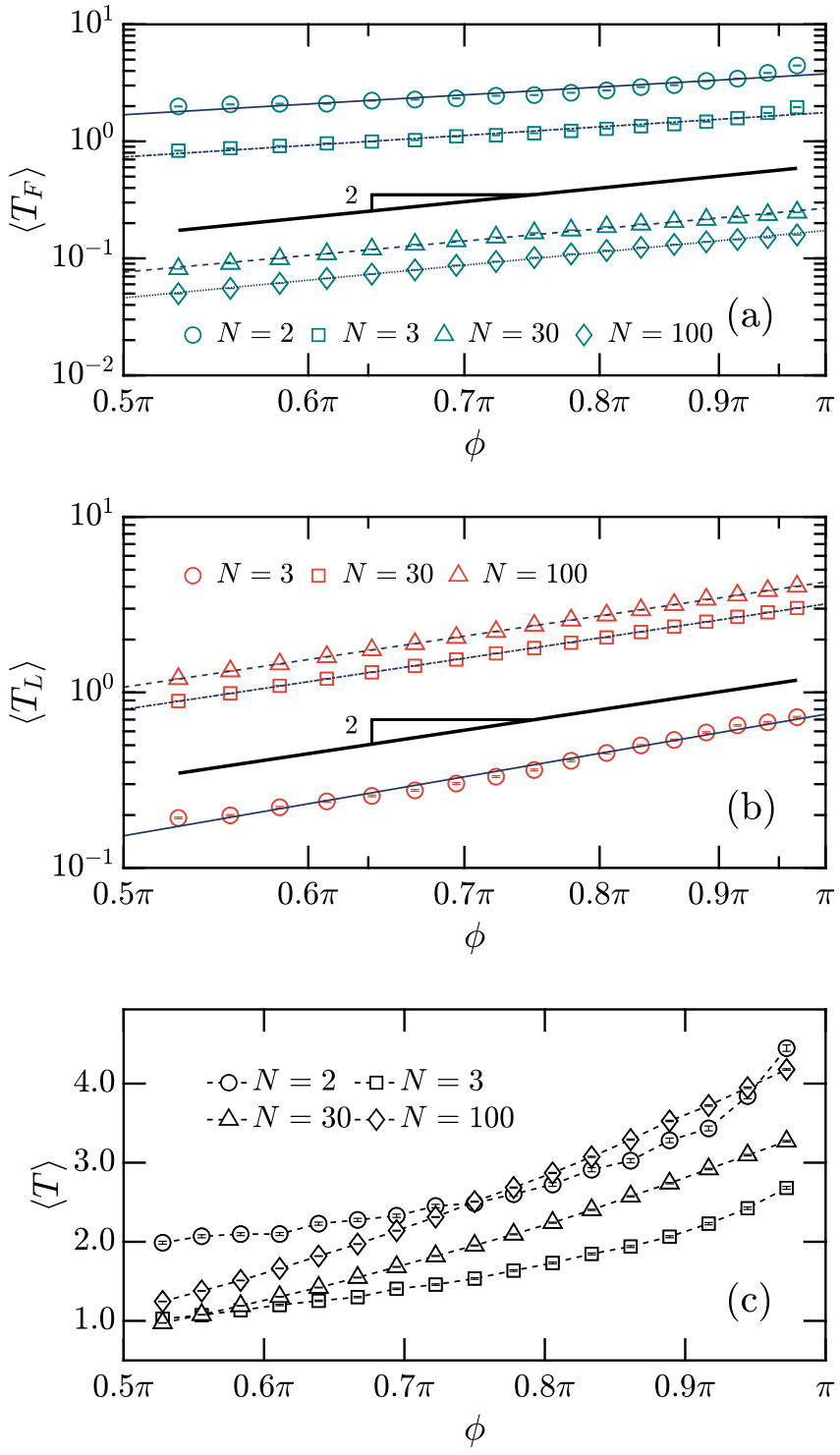

Figure 3.7: Numerical results for a planar IC of the average first (a) and last (b) binding times, as well as (c) the folding time, as a function of the closing angle $\phi \in [{}^{19\pi}/_{36}, {}^{35\pi}/_{36}]$, for four different values of $N = \{2, 3, 30, 100\}$. The graph in (b) does not have data for $N = 2$ since there is only one binding time (the first one, T_F) in a system with just two lateral faces. The solid and dashed blue lines in (a) and (b) are given by Eq. (3.33) with parameters A and k, whose values are presented in Table 3.1, obtained by a least square fit of the data. The black line segment with slope two in (a) and (b) is a visual aid to indicate how the data would vary should it follow a power law of the form (3.33) with $k = 2$. The data in (a) and (b) is plotted in a log-log scale, while in (c) it is linear. Apart from the closing angle, the simulations were performed with SC. Results are averages from $N_S = 10^4$ independent samples.

Table 3.1: Values of the parameters of the power law (3.33) that best predict the data from Figs. 3.7a and 3.7b, obtained from a least square fit. The values of A are in Brownian time units.

Binding time	A	k
$\langle T_F\,(N=2)\rangle$	1.00±0.01	1.16±0.01
$\langle T_F\,(N=3)\rangle$	0.416±0.004	1.26±0.01
$\langle T_F\,(N=30)\rangle$	0.0335±0.0001	1.817±0.004
$\langle T_F\,(N=100)\rangle$	(1937±4)×10^{-5}	1.912±0.003
$\langle T_L\,(N=3)\rangle$	0.0539±0.0005	2.30±0.01
$\langle T_L\,(N=30)\rangle$	0.324±0.001	1.997±0.004
$\langle T_L\,(N=100)\rangle$	0.436±0.001	1.991±0.003

From Fig. 3.7a we see that, for low N, the average first binding time $\langle T_F\rangle$ does not exhibit a power law relation with the closing angle ϕ. As we increase N, this relationship seems to tend to linearity (in a log-log scale), with the parameter k also approaching a value close to two, as seen in Table 3.1. In Fig. 3.7b the results show that, for low N, the average last binding time $\langle T_L\rangle$ also does not follow a power law relation with ϕ, indicating that, after the first binding, the mean first passage time (MFPT) for the $N-2$ remaining particles of the lattice model does not follow the functional prediction of the average diffusion time in a line, as given by Eq. (2.41). This implies that the assumption $\theta_i\,(t=T_F)<\phi$ may not be applicable when N is low. Thus the 1D Brownian processes are not well described by a singular domain ranging from $\theta\in[0,\phi]$ in this regime, and the probability of particles being found at $\theta_i\,(t\geq T_F)>\phi$ must also be considered. On the other hand, for high values of N, the aforementioned assumption does seem to be valid, with the data displaying linearity (in a log-log scale), characteristic of the power law relation. Furthermore, the value of the parameter k is also consistent with Eq. (2.41). For $N=100$, the deviation of k from the exact value two may be due to a greater probability for particles to reach values of $\theta_i\,(t\leq T_F)>\phi$, since there is a greater number of lateral faces. This would imply an optimal value in the range $3<N<100$ where the assumption $\theta_i\,(t=T_F)<\phi$ is most valid, and where the 1D Brownian processes are therefore well described by Eq. (2.41). For the average folding time $\langle T\rangle$, we see in Fig. 3.7c that by varying ϕ we can further control the average time it takes for the folding process to complete. Namely, while in Fig. 3.4 the results seemed to indicate that the folding time for $N=2$ was the longest for all the data points obtained, even considering up to $N=100$, here the results in Fig. 3.7c show a crossover value for ϕ between the cases $N=2$ and $N=100$, in which $\langle T\,(N=2)\rangle$ is faster than $\langle T\,(N=100)\rangle$.

In the following two subsections we try to elucidate the functional dependence of the average first and last

binding times with the closing angle for low values of N.

3.2.1 First binding time as a function of ϕ

As we have seen, the average first binding time for two lateral faces with outer angles θ_0 and θ_1 is a MFPT associated with a 2D Brownian process. The dynamics of this process can be approximated by a 2D random walk in a lattice of $L \times L$ sites and reflective boundaries [40, 41]. The position $(n(\theta_0(t)), n(\theta_1(t)))$ of the random walk in this lattice is defined by the values of the outer angles $\theta_0(t)$ and $\theta_1(t)$. For a large enough number of sites L, the average first binding time $\langle T_F(\phi, N=2)\rangle$ for a given ϕ should be given by the MFPT to reach the target $(n(\phi), n(\phi))$. An exact spatiotemporal analytical solution for this MFPT in this domain configuration has already been obtained by Giuggioli [41]. To derive it the author first obtained the occupation probability $P_{\vec{a}}(\vec{n}, t)$ for a 2D domain with reflective boundaries of finding the random walk at site $\vec{n} = (n_1, n_2)$ at time t, provided it started at $\vec{a} = (a_1, a_2)$. n_1, n_2 are integers which represent a lattice site, while a_1, a_2 are integers which identify the starting one. This probability can then be linked to the first passage probability $F_{\vec{a}}(\vec{n}, t)$ for the random walk, which started at $\vec{a}$, to reach $\vec{n}$ for the first time at time t, by means of the renewal equation [31, 35],

$$P_{\vec{a}}(\vec{n}, t) = \delta_{t,0}\delta_{\vec{n},\vec{a}} + \sum_{t'=0}^{t} F_{\vec{a}}(\vec{n}, t') P_{\vec{n}}(\vec{n}, t - t') \quad . \tag{3.37}$$

The term with the Kronecker delta functions account for the initial condition that the walk starts at $\vec{a}$ while the convolution term takes into account the possibility for the walk to first reach $\vec{n}$ at a time $t' \leq t$ with probability $F_{\vec{a}}(\vec{n}, t')$, and to return to $\vec{n}$ after the time interval $t - t'$ with probability $P_{\vec{n}}(\vec{n}, t - t')$. If we multiply Eq. (3.37) on both sides by z^t, with z being a (generally complex) variable, and sum over all values of t we can turn the equation into an algebraic one by making use of the following generating functions,

$$\tilde{P}_{\vec{a}}(\vec{n}, z) = \sum_{t=0}^{\infty} P_{\vec{a}}(\vec{n}, t) z^t \quad , \tag{3.38}$$

$$\tilde{F}_{\vec{a}}(\vec{n}, z) = \sum_{t=0}^{\infty} F_{\vec{a}}(\vec{n}, t) z^t \quad , \tag{3.39}$$

which gives

$$\tilde{P}_{\vec{a}}(\vec{n}, z) = \delta_{\vec{n},\vec{a}} + \tilde{P}_{\vec{n}}(\vec{n}, z) \tilde{F}_{\vec{a}}(\vec{n}, z) \quad . \tag{3.40}$$

Rearranging in order of $\tilde{F}_{\vec{a}}(\vec{n}, z)$ gives the fundamental relation between the generating functions of the occupation probability and the first passage probability,

$$\tilde{F}_{\vec{a}}(\vec{n}, z) = \frac{\tilde{P}_{\vec{a}}(\vec{n}, z) - \delta_{\vec{n},\vec{a}}}{\tilde{P}_{\vec{n}}(\vec{n}, z)} \quad . \tag{3.41}$$

Finally, the MFPT $\langle t_{\vec{a}\to\vec{n}} \rangle$ for the random walk to first reach $\vec{n}$, provided it started at $\vec{a}$, can be obtained using Eq. (3.41) and the following known relation between the generating function of the first passage probability and the MFPT [42],

$$\langle t_{\vec{a}\to\vec{n}} \rangle = \left. \frac{d\tilde{F}_{\vec{a}}(\vec{n}, z)}{dz} \right|_{z=1} \quad . \tag{3.42}$$

Thus, for our lattice model with L sites, for two lateral faces with an identical starting outer angle of $\theta_i(t=0) = a$, we approximate the average first binding time $\langle T_F(\phi, N=2) \rangle$ by the MFPT (3.42) with $\vec{a} = (n(a), n(a))$ and $\vec{n} = (n(\phi), n(\phi))$. The expression for $\langle T_F(\phi, N=2) \rangle$ then consists in a nested sum

$$\langle T_F(\phi, N=2) \rangle \simeq 8\Delta t \cdot \left[\sum_{k_1=1,k_2=1}^{L-1} K_{k_1,k_2}(\phi) + \sum_{k=1}^{L-1} K_{k,0}(\phi) \right] \quad , \tag{3.43}$$

where

$$\begin{aligned} K_{k_1,k_2}(\phi) = \Bigg\{ & \cos^2\left(\left(n(\phi) + \frac{1}{2} \right) \frac{\pi k_1}{L} \right) \cos^2\left(\left(n(\phi) + \frac{1}{2} \right) \frac{\pi k_2}{L} \right) \\ & - \cos\left(\left(n(a) + \frac{1}{2} \right) \frac{\pi k_1}{L} \right) \cos\left(\left(n(a) + \frac{1}{2} \right) \frac{\pi k_2}{L} \right) \times \\ & \times \cos\left(\left(n(\phi) + \frac{1}{2} \right) \frac{\pi k_1}{L} \right) \cos\left(\left(n(\phi) + \frac{1}{2} \right) \frac{\pi k_2}{L} \right) \Bigg\} \cdot \left[2 - \cos\left(\frac{\pi k_1}{L} \right) - \cos\left(\frac{\pi k_2}{L} \right) \right]^{-1} \quad . \end{aligned} \tag{3.44}$$

To account for the time the particle spends hopping, Eq. (3.43) has been additionally multiplied by Δt, as given by Eq. (1.18) with $N_F = 2$, in comparison with the solution presented in [41]. Furthermore the site variable n has also been added one unit to account for the different site numbering convention, since in [41] the sites start at $n = 1$, whereas in our work they start at $n = 0$. We plot Eq. (3.43) for a planar IC ($a = 0$) along with the numerical data of $\langle T_F(\phi, N=2) \rangle$ (presented in Fig. 3.7a) in Fig. 3.8a. Eq. (3.43) clearly predicts the correct functional dependence of $\langle T_F(N=2) \rangle$ with ϕ, for a planar IC. The small deviations are due to the possible particle-particle interactions in the interacting region $\theta \in]\phi, \pi]$. This is verified by Fig. 3.8b, where we again plot Eq. (3.43) along with the data for $\langle T_F(N=2) \rangle$, but now the simulation was performed without the particle-particle interactions. To turn these interactions off,

we simply ignore condition three of step four in the algorithm presented in section 1.2.

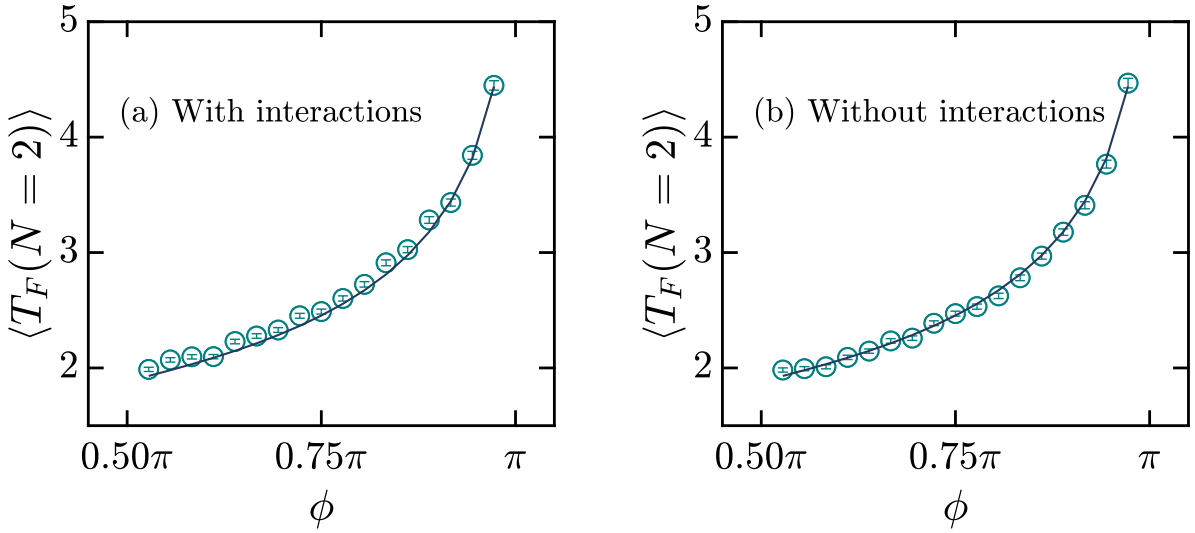

Figure 3.8: (a) Plot of Eq. (3.43) (solid line) in units of Brownian time, with $L = 181$, $a = 0$ and $D = D_0$, for $\phi \in [{}^{19\pi}/_{36}, {}^{35\pi}/_{36}]$ along with the numerical data of $\langle T_F(\phi, N=2)\rangle$ as presented in Fig. 3.7a (circle points). (b) Plot of the same Eq. (3.43) (solid line) with the same parameters as in (a), but now compared to the data obtained for $\langle T_F(\phi, N=2)\rangle$ where the particle-particle interactions were turned off. Apart from the closing angle, the simulations were performed with SC. Numerical results are averages from $N_S = 10^4$ independent samples.

For $N = 3$, by extension, the dynamics of the first binding should be consistent with a 3D lattice random walk in a box with $L\times L\times L$ sites and reflective boundaries. The position $(n(\theta_0(t)), n(\theta_1(t)), n(\theta_2(t)))$ of the random walker in this 3D lattice is now defined by the outer angles $\theta_0(t)$, $\theta_1(t)$ and $\theta_2(t)$ of the three lateral faces. The average first binding time $\langle T_F(N=3)\rangle$, for a large L, should then be similar to the MFPT to reach any of the set of $3L-1$ targets $(n(\theta_0(t)), n(\phi), n(\phi))$, $(n(\phi), n(\theta_1(t)), n(\phi))$ or $(n(\phi), n(\phi), n(\theta_2(t)))$, where $n(\theta_0(t)), n(\theta_1(t)), n(\theta_2(t)) \in [0, L-1] \setminus n(\phi)$. Although the general analogue of Eq. (3.43) was derived in [41] for arbitrary dimensions, it is only applicable when there is a single target. Therefore the analytical dependence of $\langle T_F(N=3)\rangle$ with ϕ remains an open question.

3.2.2 Last binding time as a function of ϕ

Results for $N = 3$ presented in Fig. 3.7b suggest that the assumption $\theta_i(t = T_F) < \phi$, where i is any of the lateral faces that did not attach at the first binding, fails for a low number of lateral faces. This prompt us to also account for the probability density function $P(\theta_i > \phi, t)$ in the region at the right of ϕ in the calculation of the MFPT distribution for the last binding time. We consider a domain consisting of a region of size ϕ with a lower reflective boundary and an absorbing upper one, connected at the right

border to another region of size $\theta_u - \phi$ with the boundary types interchanged. Here, for simplicity, we assume $\Delta\phi \simeq 0$, and consider the simplest case $N = 3$. We define $P_\phi(\theta_i, t)$ as the probability density of finding the outer angle of the remaining particle i with a value θ_i after an elapsed time t since the first binding time T_F, which should obey the 1D diffusion equation (1.14). Since for $N = 3$ there is only one lateral face still free after the first binding, we drop the subscript i, $\theta_i \equiv \theta$. For this domain configuration, we hypothesize the following general piecewise solution for $P_\phi(\theta, t)$,

$$P_\phi(\theta,t) = P_{<\phi}(\theta,t) + P_{>\phi}(\theta,t) = P_R^A(\theta,t)\,H(\phi-\theta) + P_A^R(\theta,t)\,H(\theta-\phi) \quad , \tag{3.45}$$

where $\theta \in [0, \theta_u]$, $\phi \in]0, \theta_u[$, $P_{<\phi}(\theta,t) = P_R^A(\theta,t)\,H(\phi-\theta)$, $P_{>\phi}(\theta,t) = P_A^R(\theta,t)\,H(\theta-\phi)$, $H(x)$ is the Heaviside step function (with the half-maximum convention),

$$H(x) = \begin{cases} 0 & \text{, if } x < 0 \\ \frac{1}{2} & \text{, if } x = 0 \\ 1 & \text{, if } x > 0 \end{cases} \quad , \tag{3.46}$$

while $P_R^A(\theta,t)$ and $P_A^R(\theta,t)$ are the solutions of the 1D diffusion equation for a system that ranges from 0 to ϕ with a reflective lower boundary and absorbing upper one, and another that ranges from 0 to $\theta_u - \phi$ with an absorbing lower boundary and a reflective upper one, respectively,

$$\begin{cases} P_R^A(\theta,t) = \sum_{m=0}^{\infty} A_m e^{-k_m^2 Dt}\cos(k_m\theta) \\ P_A^R(\theta,t) = \sum_{m=0}^{\infty} B_m e^{-{k'_m}^2 Dt}\sin(k'_m\cdot(\theta-\phi)) \end{cases} \quad , \tag{3.47}$$

with eigenvalues $k_m = {}^{\pi}/_{\phi}\cdot({}^1/_2 + m)$ and $k'_m = ({}^{\pi}/_{(\theta_u-\phi)})\cdot({}^1/_2 + m)$, while A_m and B_m are constants to be defined from the distribution of this outer angle at the moment of the first binding, $P_\phi(\theta, t=0)$. Since the first binding time T_F is a random variable, so is the distribution $P_\phi(\theta, t=0)$, since it depends on T_F. As we are trying to compute the mean of a random variable, we bypass this intricacy by considering instead the average of this distribution,

$$P_F(\theta,\phi) \equiv \langle P_\phi(\theta, t=0)\rangle = \int_0^\infty P(\theta, t_F \mid \text{free})\, f_F(t_F)\, dt_F \quad , \tag{3.48}$$

where $f_F(t')$ is the PDF for the MFPT associated with the 2D Brownian processes, as introduced in

subsection 3.1.1, while $P\left(\theta, t' \mid \text{free}\right)$ is the probability density of the outer angle θ_i having a value of θ at the first binding time $T_F = t_F$, provided that i is the lateral face that does not attach at the first binding. Thus, $P\left(\theta, t' \mid \text{free}\right)$, and by extension $P_F\left(\theta, \phi\right)$, is a conditional probability density. We set ϕ as an argument in $P_F\left(\theta, \phi\right)$ as we expect in advance that this average conditional distribution should in general depend on the closing angle. The IC then reads

$$\begin{cases} P_{<\phi}\left(\theta, t=0\right) = P_F\left(\theta, \phi\right) H\left(\phi - \theta\right) \\ P_{>\phi}\left(\theta, t=0\right) = P_F\left(\theta, \phi\right) H\left(\theta - \phi\right) \end{cases}, \tag{3.49}$$

From these ICs we have

$$\begin{cases} H\left(\phi - \theta\right) \sum_{m=0}^{\infty} A_m \cos\left(k_m \theta\right) = P_F\left(\theta, \phi\right) H\left(\phi - \theta\right) \\ H\left(\theta - \phi\right) \sum_{m=0}^{\infty} B_m \sin\left(k'_m \cdot \left(\theta - \phi\right)\right) = P_F\left(\theta, \phi\right) H\left(\theta - \phi\right) \end{cases}. \tag{3.50}$$

Multiplying the first equation by $\cos\left(k_n \theta\right)$ and the second by $\sin\left(k'_n \cdot \left(\theta - \phi\right)\right)$, with $n \in \mathbb{N}_0$, and integrating both equations in θ over the whole angular domain (while also making the change of variable $\theta - \phi \to x$ in the second equation) we get

$$\begin{cases} \sum_{m=0}^{\infty} A_m \int_0^{\phi} \cos\left(k_n \theta\right) \cos\left(k_m \theta\right) d\theta = \int_0^{\phi} \cos\left(k_n \theta\right) P_F\left(\theta, \phi\right) d\theta \\ \sum_{m=0}^{\infty} B_m \int_0^{\theta_u - \phi} \sin\left(k'_n x\right) \sin\left(k'_m x\right) dx = \int_0^{\theta_u - \phi} \sin\left(k'_n x\right) P_F\left(x + \phi, \phi\right) dx \end{cases}. \tag{3.51}$$

We can make use of the orthogonality of the trigonometric functions (see appendix D),

$$\begin{cases} \int_0^{\phi} \cos\left(k_n \theta\right) \cos\left(k_m \theta\right) d\theta = \frac{\phi}{2} \delta_{m,n} \\ \int_0^{\theta_u - \phi} \sin\left(k'_n x\right) \sin\left(k'_m x\right) dx = \frac{\theta_u - \phi}{2} \delta_{m,n} \end{cases}, \tag{3.52}$$

to redefine the values of A_m and B_m in Eqs. (3.51) in terms of the integral involving $P_F\left(\theta, \phi\right)$,

$$\begin{cases} A_m = \frac{2}{\phi} \int_0^{\phi} \cos\left(k_m \theta\right) P_F\left(\theta, \phi\right) d\theta \equiv \frac{2}{\phi} C_m^{(1)}\left(\phi\right) \\ B_m = \frac{2}{\theta_u - \phi} \int_0^{\theta_u - \phi} \sin\left(k'_m x\right) P_F\left(x + \phi, \phi\right) dx \equiv \frac{2}{\theta_u - \phi} C_m^{(2)}\left(\phi\right) \end{cases}. \tag{3.53}$$

Eq. (3.45) is then cast in the following form

$$P_{\phi}\left(\theta,t\right)=H\left(\phi-\theta\right)\frac{2}{\phi}\sum_{m=0}^{\infty}C_{m}^{(1)}\left(\phi\right)e^{-k_{m}^{2}Dt}\cos\left(k_{m}\theta\right)+ \tag{3.54}$$

$$+H\left(\theta-\phi\right)\frac{2}{l-\phi}\sum_{m=0}^{\infty}C_{m}^{(2)}\left(\phi\right)e^{-{k'_{m}}^{2}Dt}\sin\left(k'_{m}\cdot\left(\theta-\phi\right)\right)\quad . \tag{3.55}$$

The PDF $f_{\phi}\left(t\right)$ of the last binding time T_L for $N=3$ is, as before, given by the flux at the closing angle ϕ, but we need to be aware of a little nuance, which is the fact that the probability concentration 'exits' the left region at ϕ by going from left to right, while it 'exits' the right one at ϕ by going from right to left, so their fluxes must have opposite signs. Instead of computing the derivatives exactly at $\theta=\phi$ (remembering that $^{\partial H(x)}/_{\partial x}=\delta\left(x\right)$), we introduce a small real number $\varepsilon>0$, and immediately take the limit $\varepsilon\to0$, so that we can unambiguously attribute the value one to the Heaviside step functions after the derivation,

$$f_{\phi}\left(t\right)=\lim_{\varepsilon\to0}\left[-D\left.\frac{\partial P_{\theta<\phi}\left(\theta,t\right)}{\partial\theta}\right|_{\theta=\phi-\varepsilon}+D\left.\frac{\partial P_{\theta>\phi}\left(\theta,t\right)}{\partial\theta}\right|_{\theta=\phi+\varepsilon}\right]=$$

$$=D\cdot\left[-P_{R}^{A}\left(\phi,t\right)-\left.\frac{\partial P_{R}^{A}\left(\theta,t\right)}{\partial\theta}\right|_{\theta=\phi}+P_{A}^{R}\left(\phi,t\right)+\left.\frac{\partial P_{A}^{R}\left(\theta,t\right)}{\partial\theta}\right|_{\theta=\phi}\right]\quad . \tag{3.56}$$

Knowing that

$$\begin{cases}P_{R}^{A}\left(\phi,t\right)=0\\ \left.\frac{\partial P_{R}^{A}(\theta,t)}{\partial\theta}\right|_{\theta=\phi}=-\frac{2}{\phi}\sum_{m=0}^{\infty}\left(-1\right)^{m}k_{m}e^{-k_{m}^{2}Dt}C_{m}^{(1)}\left(\phi\right)\\ P_{A}^{R}\left(\phi,t\right)=0\\ \left.\frac{\partial P_{A}^{R}(\theta,t)}{\partial\theta}\right|_{\theta=\phi}=\frac{2}{\theta_{u}-\phi}\sum_{m=0}^{\infty}k'_{m}e^{-{k'_{m}}^{2}Dt}C_{m}^{(2)}\left(\phi\right)\end{cases}\quad , \tag{3.57}$$

we have

$$f_{\phi}\left(t\right)=2D\cdot\left[\frac{1}{\phi}\sum_{m=0}^{\infty}\left(-1\right)^{m}k_{m}e^{-k_{m}^{2}Dt}C_{m}^{(1)}\left(\phi\right)+\frac{1}{\theta_{u}-\phi}\sum_{m=0}^{\infty}k'_{m}e^{-{k'_{m}}^{2}Dt}C_{m}^{(2)}\left(\phi\right)\right]\quad . \tag{3.58}$$

The average of the last binding time $\left\langle T_{L}\left(\phi,N=3\right)\right\rangle$ for three lateral faces is then given by

$$\left\langle T_{L}\left(\phi,N=3\right)\right\rangle=\int_{0}^{\infty}tf_{\phi}\left(t\right)dt\quad . \tag{3.59}$$

The integrals in Eq. (3.59) are of the form (2.28), and after computing them we arrive at the expression for the average last binding time, for $N=3$, which takes into account the region at the right of the closing

angle ϕ,

$$\langle T_L(\phi, N=3)\rangle = \frac{2}{D}\cdot\left[\frac{1}{\phi}\sum_{m=0}^{\infty}\frac{(-1)^m}{k_m^3}C_m^{(1)}(\phi)+\frac{1}{\theta_u-\phi}\sum_{m=0}^{\infty}\frac{1}{{k'_m}^3}C_m^{(2)}(\phi)\right] \quad , \tag{3.60}$$

with parameters

$$\begin{cases} k_m = \frac{\pi}{\phi}\cdot\left(\frac{1}{2}+m\right) \\ k'_m = \frac{\pi}{\theta_u-\phi}\cdot\left(\frac{1}{2}+m\right) \\ C_m^{(1)}(\phi) = \int_0^{\phi}\cos(k_m\theta)\,P_F(\theta,\phi)\,d\theta \\ C_m^{(2)}(\phi) = \int_{\phi}^{\theta_u}\sin\left(k'_m\cdot(\theta-\phi)\right)P_F(\theta,\phi)\,d\theta \end{cases} . \tag{3.61}$$

If we were to assume that $\theta<\phi$ at the first binding moment was always true, and that its distribution is, on average, uniform over the subdomain $\theta\in[0,\phi[$, then $P_F(\theta,\phi)=(1/\phi)\,H(\phi-\theta)$, and so

$$\begin{cases} C_m^{(1)} = \frac{1}{\phi}\int_0^{\phi}\cos(k_m\theta)\,d\theta = \frac{(-1)^m}{\phi k_m} \\ C_m^{(2)} = 0 \end{cases} . \tag{3.62}$$

Plugging these coefficients into Eq. (3.60) would give

$$\langle T_L(\phi, N=3)\rangle = \frac{2}{\phi^2 D}\sum_{m=0}^{\infty}\frac{1}{k_m^4} \quad , \tag{3.63}$$

which is equal to Eq. (2.44), and so the solution of $\langle T_L(\phi,N=3)\rangle$ with these assumptions would be equivalent to Eq. (2.48), the average time it takes for a diffusing particle in a line of length ϕ with a reflective and absorbing boundary, starting with a uniform distribution, to reach the absorbing boundary,

$$\langle T_L(\phi, N=3)\rangle = \langle t_D\rangle_R = \frac{\phi^2}{3D} \quad \text{, for } P_F(\theta,\phi)=(1/\phi)\,H(\phi-\theta). \tag{3.64}$$

On the other hand, if, on average, the angular distribution of the free lateral face at the moment of the first binding were uniform over the whole domain $\theta\in[0,\theta_u]$, so $P_F(\theta,\phi)=1/\theta_u$, we would have

$$\begin{cases} C_m^{(1)} = \frac{1}{\theta_u}\int_0^{\phi}\cos(k_m\theta)\,d\theta = \frac{(-1)^m}{\theta_u k_m} \\ C_m^{(2)} = \frac{1}{\theta_u}\int_{\phi}^{\theta_u}\sin\left(k'_m\cdot(\theta-\phi)\right)d\theta = \frac{1}{\theta_u k'_m} \end{cases} . \tag{3.65}$$

Plugging these results in Eq. (3.60) we get

$$\langle T_L(\phi, N=3)\rangle = \frac{2}{\theta_u D}\cdot\left[\frac{1}{\phi}\sum_{m=0}^{\infty}\frac{1}{k_m^4} + \frac{1}{\theta_u-\phi}\sum_{m=0}^{\infty}\frac{1}{{k'_m}^4}\right] = \frac{2\cdot\left[\phi^3+(\theta_u-\phi)^3\right]}{\pi^4\theta_u D}\sum_{m=0}^{\infty}\frac{1}{\left(\frac{1}{2}+m\right)^4} \quad . \tag{3.66}$$

From Eq. (2.47) we see that the series in the rightmost hand-side of Eq. (3.66) is equal to $\pi^4/6$. Thus, using the value $\theta_u = \pi$ for a pinned system, we get the following result for the average last binding $\langle T_L(\phi, N=3)\rangle$ for three lateral faces, in which the outer angle of the free one has, on average, a uniform angular distribution at the moment of the first binding,

$$\begin{aligned}\langle T_L(\phi, N=3)\rangle &= \frac{1}{3\pi D}\cdot\left[\phi^3+(\pi-\phi)^3\right] = \\ &= \frac{\phi}{\pi}\frac{\phi^2}{3D} + \frac{\pi-\phi}{\pi}\frac{(\pi-\phi)^2}{3D} \quad \text{, for } P_F(\theta,\phi) = 1/\pi.\end{aligned} \tag{3.67}$$

By comparing this result with Eq. (2.48) we notice that this average is just the weighted sum of the two individual MFPTs for a diffusing particle, starting with a uniform distribution in the respective subdomains of length ϕ and $\pi-\phi$, to reach the absorbing boundary. However, the following numerical results will show that the average angular distribution of the free lateral face is not uniform at the moment of the first binding.

Considering the lattice model, we define now $\varepsilon_\phi(n)$ as the average conditional occupation probability of finding the free particle at site n at the moment the other two get trapped at $n(\phi)$,

$$\varepsilon_\phi(n) = \int_0^\infty P_n(t_F \mid \text{free})\, f_F(t_F)\, dt_F \quad , \tag{3.68}$$

where $P_n(t_F \mid \text{free})$ is the probability of finding the random walk i at site n at time $t_F = T_F$, which should satisfy Eq. (1.6), provided that i is the particle not yet trapped after the first binding.

If we divide $\varepsilon_\phi(n)$ by the lattice spacing $\Delta\theta$, and take the limit $L\to\infty$, this fraction should converge to $P_F(\theta,\phi)$,

$$P_F(\theta,\phi) = \lim_{L\to\infty}\frac{\varepsilon_\phi(n)}{\Delta\theta} \quad . \tag{3.69}$$

We ran N_S independent simulations with $N=3$ to estimate $\varepsilon_\phi(n)$ numerically by performing a Bernoulli trial,

$$\varepsilon_\phi(n) \simeq \frac{B_n}{N_S} \quad \text{, for large } N_S, \tag{3.70}$$

where B_n is the number of times out of a total of N_S that the free particle was at site n at the moment

the other two particles met at $n(\phi)$. The results of these simulations for a planar IC, for two values of $\phi = \{5\pi/9, 2\pi/3\}$ are presented in Figs. 3.9a and 3.9b. As we can see, the distribution $P_F(\theta, \phi)$ is far from uniform. We also present independent simulation results for a random IC in Figs. 3.9c and 3.9d to show that for this case $P_F(\theta, \phi)$ is also not uniform. This means that lateral faces starting the folding process with a uniform angular distribution do not ensure that the average distribution of the one still free is also uniform when the first binding occurs. We believe this is due to the conditional nature of this probability density. For $N = 3$, before the first binding, we have three trajectories in the angular space associated with the three outer angles θ_i, $i = \{0, 1, 2\}$. From these three, only the one associated with the lateral face that does not attach at the first binding will contribute to $P_F(\theta, \phi)$. Trajectories in the angular space that have the outer angle θ_i far from ϕ when the first binding is imminent have higher weight in $P_F(\theta, \phi)$ compared to the ones that are closer since intuitively there is a higher probability that i is not the lateral face out of the three that will attach at the first binding.

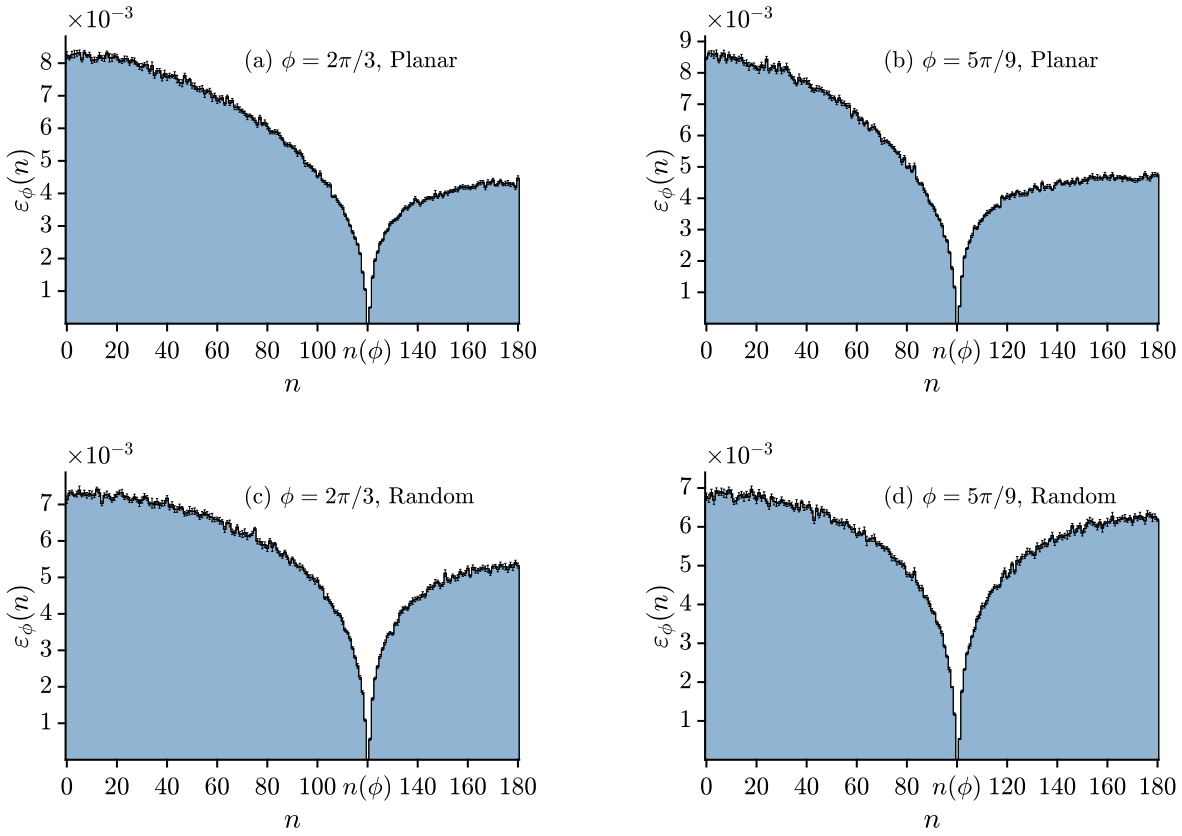

Figure 3.9: (a,b) Histograms with 181 bins each made from $N_S = 10^6$ independent samples of the average conditional occupation probability $\varepsilon_\phi(n)$ of finding the free particle at site n at the moment the other two get trapped at $n(\phi)$, as estimated from Eq. (3.70), for a planar IC, for two values of the closing angle $\phi = \{5\pi/9, 2\pi/3\}$. (c,d) The results for the same probability with the same two values of ϕ, but with a random IC. Apart from the closing angle, the simulations were performed with SC.

The average probability $\langle P(\theta < \phi, T_F) \rangle$ of finding the free particle at the left subdomain $\theta \in [0, \phi[$

when the first binding occurs is related to $P_F(\theta, \phi)$ by

$$\langle P(\theta < \phi, T_F) \rangle = \int_0^{\phi} P_F(\theta, \phi)\, d\theta \quad . \tag{3.71}$$

We estimate this probability numerically by again performing a Bernoulli trial,

$$\langle P(\theta < \phi, T_F) \rangle \simeq \frac{B_{<\phi}}{N_S} \quad \text{, for large } N_S, \tag{3.72}$$

where $B_{<\phi}$ is the number of times out of a total of N_S that the free particle was at a site $n < n(\phi)$ at the moment the other two particles met at $n(\phi)$. From the results presented in Fig. 3.10 we see that the conditional nature of $P_F(\theta, \phi)$ and the planar IC lead to a non-linear average probability $\langle P(\theta < \phi, T_F) \rangle$ as a function of ϕ.

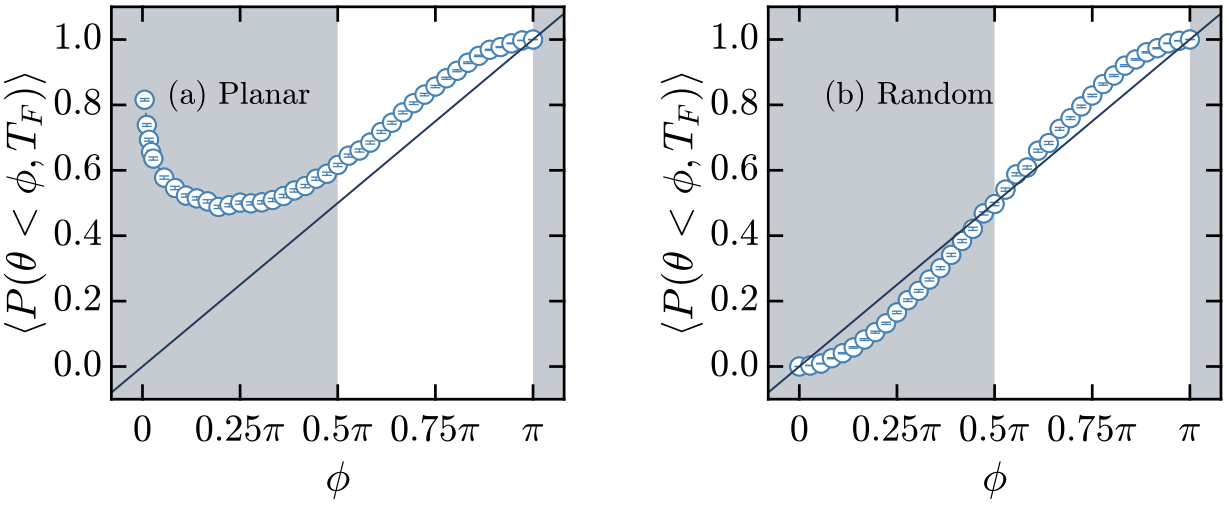

Figure 3.10: Numerical results of the average probability $\langle P(\theta < \phi, T_F) \rangle$ of finding the free particle at the left subdomain $\theta \in [0, \phi[$ when the first binding occurs, as a function of ϕ (with $\phi \in [\pi/180, \pi]$ in (a), and $\phi \in [0, \pi]$ in (b)), as estimated from Eq. (3.72), for a (a) planar and (b) random ICs. The solid line corresponds to the hypothetical case where the probability density of the free particle at the moment of the first binding is, on average, uniform, so $P_F(\theta, \phi) = 1/\pi$, and thus $\langle P(\theta < \phi, T_F) \rangle = \phi/\pi$. Points inside the greyed-out region are geometrically forbidden in the context of this kirigami structure, and are only plotted for theoretical insight. The points half inside this region, $\phi = \pi/2$ and $\phi = \pi$, have geometrical meaning only if they are interpreted as limiting values, approached from the right and the left, respectively. Apart from the closing angle, the simulations were performed with SC. Numerical results are averages from $N_S = 10^4$ independent trials.

We now compare in Table 3.2 the theoretical value predicted by Eq. (3.60) for the average last binding time with $N = 3$ and $\phi = \{5\pi/9, 2\pi/3\}$, for a planar IC, with the respective numerical results presented in Fig. 3.7b. To estimate the coefficients $C_m^{(1)}(\phi)$ and $C_m^{(2)}(\phi)$ in Eq. (3.60) we approximate the integration to

a sum using the trapezoid rule,

$$\begin{cases} C_m^{(1)}(\phi) \simeq \frac{\Delta\theta}{2} \sum_{i=1}^{n(\phi)} \left[\cos\left(k_m \theta\left(i-1\right)\right) \frac{\varepsilon_\phi(i)}{\Delta\theta} + \cos\left(k_m \theta\left(i\right)\right) \frac{\varepsilon_\phi(i+1)}{\Delta\theta} \right] \\ C_m^{(2)}(\phi) \simeq \frac{\Delta\theta}{2} \sum_{i=n(\phi)+1}^{L-1} \left[\sin\left(k'_m \cdot \left(\theta\left(i-1\right)-\phi\right)\right) \frac{\varepsilon_\phi(i)}{\Delta\theta} + \sin\left(k'_m \cdot \left(\theta\left(i\right)-\phi\right)\right) \frac{\varepsilon_\phi(i+1)}{\Delta\theta} \right] \end{cases}, \tag{3.73}$$

where $\theta(i)$ is calculated according to Eq. (1.2), $\varepsilon_\phi(i)$ is the value of the i-th bin of the respective histogram in Fig. 3.9, $\Delta\theta$ is the usual lattice spacing, $n(\phi)+1$ is the number of the bin associated with the trapping site, and L is the usual number of lattice sites, equal to the number of bins used in each of the histograms of Fig. 3.9. To compute the series, since the first eigenmode $m=0$ dominates over the others [35], we consider only a few terms in the summation, truncating at $m=15$. We also compare with the case in which we assume $\theta<\phi$ at the first binding, in which case the average last binding time for $N=3$ should vary with ϕ according to Eq. (3.64). We also present the absolute value of the relative error ΔX between the semi-analytical and numerical values, calculated as before using Eq. (2.55).

Table 3.2: Table with the numerical values of the average last binding time $\langle T_L(\phi, N=3)\rangle$ for three lateral faces, for two values of $\phi=\{5\pi/9, 2\pi/3\}$ (second and sixth circle points, respectively, of Fig. 3.7b), for a planar IC, compared with the semi-analytical prediction of Eq. (3.60) (third column) and the average diffusion time $\langle t_D\rangle_R$ for a uniform random IC (fourth column), given by Eq. (3.64). The series of Eq. (3.60) were truncated at $m=15$, and the coefficients $C_m^{(1)}(\phi)$ and $C_m^{(2)}(\phi)$ were calculated by means of the trapezoid rule, Eqs. (3.73), using the values of the histogram in Fig. 3.9a for $\phi=2\pi/3$, and the values from the one in Fig. 3.9b for $\phi=5\pi/9$. The absolute value of the relative error ΔX, as calculated from Eq. (2.55), between the numerical value and the predictions of Eqs. (3.60) and (3.64) is also presented in the fifth and sixth column, respectively. The parameters used in the equations were $D=D_0$, $\theta_u=\pi$ and $L=181$. All time averages are presented in Brownian time units.

Closing angle	Numerical value	Eq. (3.60)	Eq. (3.64)	$\lvert\Delta X\rvert$ (%) - Eq. (3.60)	$\lvert\Delta X\rvert$ (%) - Eq. (3.64)
$\phi=5\pi/9$	0.199±0.002	≈0.200	50/243≈0.206	0.50	3.27
$\phi=2\pi/3$	0.277±0.003	≈0.274	8/27≈0.296	1.09	6.51

As seen in Table 3.2, the numerical values of $\langle T_L(\phi, N=3)\rangle$ for $\phi=\{5\pi/9, 2\pi/3\}$ agree well with Eq. (3.60), with the semi-analytical predictions within the uncertainty of the simulated values, for both cases. The higher precision of Eq. (3.60) compared with (3.64) is also significant, consistent with the respective relative errors presented in the last two columns. Surprisingly, Eq. (3.64), which assumes $\theta(T_F)<\phi$, has greater precision estimating $\langle T_L(\phi, N=3)\rangle$ for the lower value of ϕ, for which the average probability $\langle P(\theta<\phi, T_F)\rangle$ of finding the free particle at the left subdomain at the first binding is also lower, as seen from Fig. 3.10a.

For higher values of $N>3$, $\langle T_L(\phi, N)\rangle$ can in principle be estimated by Eq. (3.16) with the distribution

$f_L(t)$ given by Eq. (3.60). However, the interpretation of the average conditional probability density $P_F(\theta, \phi)$ needs to be slightly modified to account for the number $N - 2 > 1$ of additional lateral faces still free after the first binding.

3.3 Effects of the degree of interaction on the average folding time

In this section we present numerical results to help elucidate how the types of interactions between lateral faces may affect the average binding times. In subsection 3.3.1 we analyse the effect that the particle-particle interactions in the interacting region have on these mean times. In subsection 3.3.2 we study how these times change if binding is only possible between the edges of lateral faces. Again, we consider only pinned systems.

3.3.1 Effect of the particle-particle interaction on the average binding times

Up until now we have been assuming without great reason that the particle-particle interactions in the sites $n \in]n(\phi), L-1]$ associated with the interacting region have negligible preponderance in the estimations of the average binding times. Here we show that this assumption is valid by comparing numerical results between simulations where these particle-particle interactions are possible and simulations where they are not. As before, to turn off these interactions, we simply ignore condition three of the fourth step in the algorithm presented in section 1.2.

In Fig. 3.11 we present this comparison for $\langle T_F \rangle$, $\langle T_L \rangle$ and $\langle T \rangle$ obtained for simulations in SC as a function of N, for a planar IC. We see that the deviations between the average binding times of these two cases are indeed very small, with only perceptible differences for the lowest values of $N = 2$ and $N = 3$.

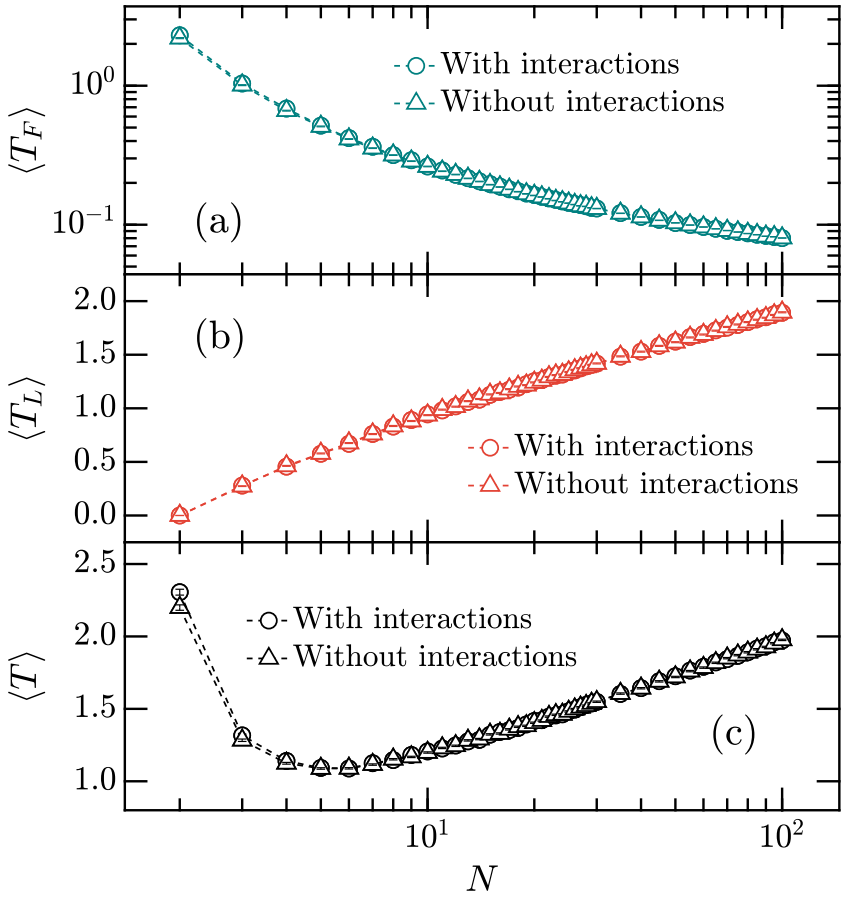

Figure 3.11: Comparison of numerical results for the average binding times $\langle T_F \rangle$ (a), $\langle T_L \rangle$ (b) and average folding time $\langle T \rangle$ (c) as a function of the number of lateral faces $N \in [2, 100]$ for a planar IC between simulations where particle-particle interactions are possible in the interacting region (circle points) and simulations where they are not (triangle points). For both cases, the parameters of the simulations were in SC. Results are averages from $N_S = 10^4$ independent samples.

To show that these deviations are still small even if the interacting region is larger, we plot in Fig. 3.12b the results of a similar analysis but for a closing angle of $\phi = \pi/2$, the value for which the interacting region encompasses the most angular space while ϕ is still geometrically allowed in the context of this kirigami structure if interpreted as a limiting value, approached from the right. Moreover we also compare in Fig. 3.12c and 3.12d these two cases of degrees of interactions but with a random IC, for the same two values of ϕ.

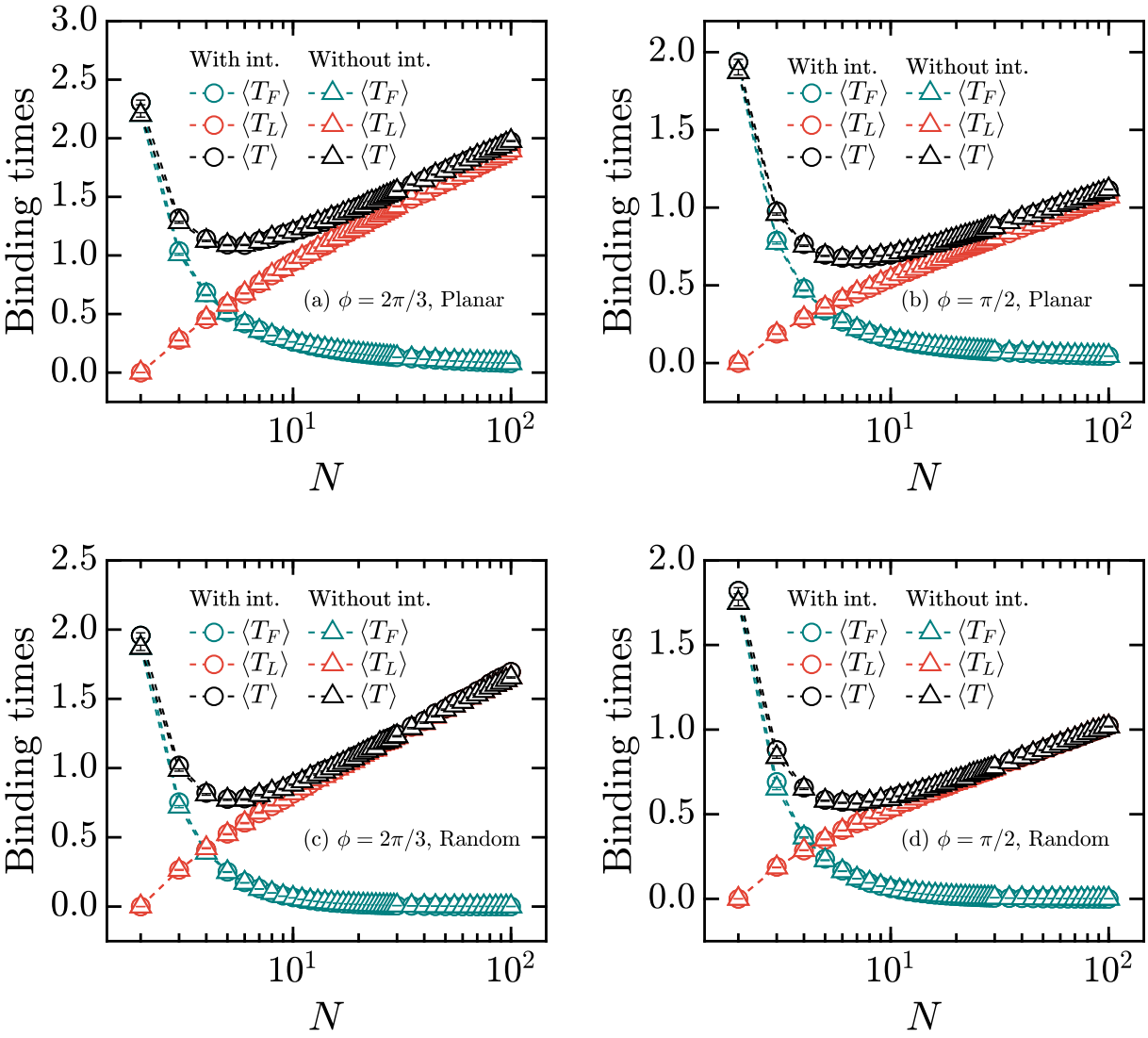

Figure 3.12: Comparison of numerical results for the average first and last binding times, as well as the average folding time, as a function of the number of lateral faces $N \in [2, 100]$, between simulations where particle-particle interactions are possible in the interacting region (circle points) and simulations where they are not (triangle points), for a planar IC (a,b) and a random one (c,d), for two values of the closing angle, $\phi = 2\pi/3$ (a,c) and $\phi = \pi/2$ (b,d). The value $\phi = \pi/2$ is to be interpreted as a limiting value, approached from the right, so that it still retains geometrical meaning for this kirigami structure. For both cases of degree of interaction, apart from the closing angle, the parameters of the simulations were in SC. Results are averages from $N_S = 10^4$ independent samples.

For better readability, we plot in Fig. 3.13 the absolute value of the relative difference $\Delta \langle T \rangle$ between the numerical values of Fig. 3.12 of the average folding time $\langle T \rangle_{\text{int}}$ with particle-particle interactions and the corresponding average folding time $\langle T \rangle_{\text{no int}}$ without them,

$$\Delta \langle T \rangle = \frac{\langle T \rangle_{\text{int}}}{\langle T \rangle_{\text{no int}}} - 1 \quad . \tag{3.74}$$

The uncertainty $\sigma_{\Delta\langle T\rangle}$ in $\Delta \langle T \rangle$ is calculated from the uncertainties σ_{int} and $\sigma_{\text{no int}}$ of $\langle T \rangle_{\text{int}}$ and $\langle T \rangle_{\text{no int}}$,

respectively, by error propagation,

$$\sigma_{\Delta\langle T\rangle} = \sqrt{\left(\sigma_{\text{int}}\frac{\partial\Delta\langle T\rangle}{\partial\langle T\rangle_{\text{int}}}\right)^2 + \left(\sigma_{\text{no int}}\frac{\partial\Delta\langle T\rangle}{\partial\langle T\rangle_{\text{no int}}}\right)^2} = \sqrt{\left(\frac{\sigma_{\text{int}}}{\langle T\rangle_{\text{no int}}}\right)^2 + \left(\frac{\sigma_{\text{no int}}\langle T\rangle_{\text{int}}}{\left(\langle T\rangle_{\text{no int}}\right)^2}\right)^2} \quad . \tag{3.75}$$

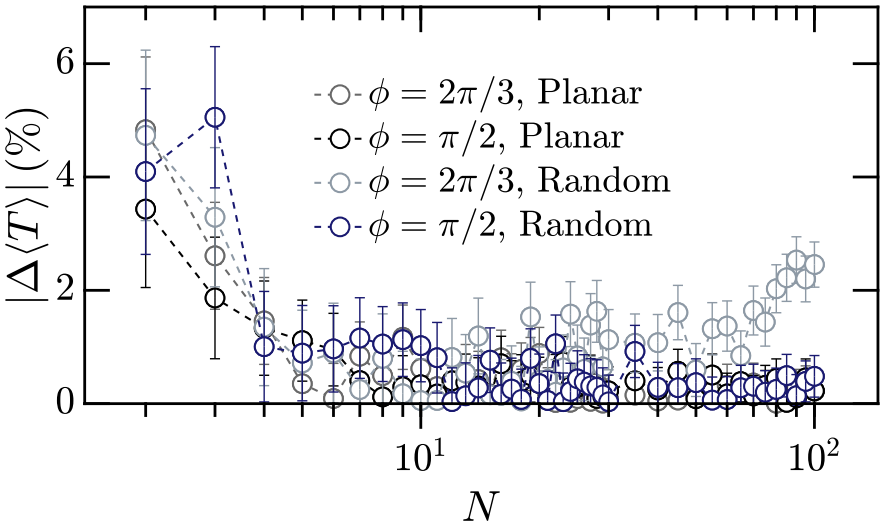

Figure 3.13: Absolute value of the relative difference $\Delta\langle T\rangle$, as calculated from Eq. (3.74), between the numerical values presented in Fig. 3.12 of the average folding time $\langle T\rangle$ with and without particle-particle interactions in the interacting region, for the two values of $\phi = \{\pi/2, 2\pi/3\}$, for both planar and random ICs.

From Fig. 3.13 we see that the relative difference between the average times for folding processes with and without particle-particle interactions in the interacting region are largest for the first few values of N, ranging from six to two percent, independently of the IC and the closing angle, and quickly converge to even smaller values close to zero as N increases. An exception of this converging behaviour is found for the case of a random IC with $\phi = 2\pi/3$, where this relative difference seems to slightly increase again for high values of $N \gtrsim 70$. It is noteworthy that this slight increase for the case of $\phi = 2\pi/3$ implies that, for a random IC, when $N \gtrsim 70$ is large, the particle-particle interactions have a greater preponderance on the average folding time when compared to the case of $\phi = \pi/2$, in which the interacting region encompasses a larger portion of the angular space.

3.3.2 Folding times for edge only binding

Here we address the question of how the average binding times and the total folding time changes if the lateral faces can only bind through their edges to their adjacent face. The indexes j_R and j_L of the right

and left adjacent face of some lateral face i, respectively, are defined by

$$j_R = \begin{cases} i+1 & \text{, if } i < N-1 \\ 0 & \text{, if } i = N-1 \end{cases} , \tag{3.76}$$

$$j_L = \begin{cases} i-1 & \text{, if } i > 0 \\ N-1 & \text{, if } i = 0 \end{cases} . \tag{3.77}$$

We will refer to this type of binding as edge only binding. For the binding where lateral faces may attach through vertices, we will coin it nonspecific binding. To implement the edge only binding in the simulations, we change step six of the algorithm presented in subsection 1.2 in the following way

6.* If the new position of the randomly chosen particle i is $n(\phi)$ (or either $n(-\phi)$ or $n(\phi)$, for the case of a suspended system) corresponding to the closing angle, if $N_{n(\phi)}(t) > 1$ (or either $N_{n(-\phi)}(t) > 1$ if the particle is at $n(-\phi)$ or $N_{n(\phi)}(t) > 1$ if it is at $n(\phi)$, for the suspended system), and if any other particle at that site is either j_R or j_L, as given respectively by Eqs. (3.76) and (3.77), the site acts as a sticky site and traps the new-coming particle, as well as j_R and/or j_L if they are not yet trapped; if it is the first time particles are getting absorbed at a closing angle, the current value of t corresponds to the first binding time T_F; update the number of particles that are still free, by subtracting the correct amount from N_F and removing from the list of free particles the ones that just got trapped;

For a template with N lateral faces which can only attach through the edges, the number of possible pairs of faces that can attach at the first binding is now only N. Thus, the number of competing 2D Brownian processes associated with the first binding event is also just N for a folding process with this type of binding. The average first binding time as a function of N, as given by Eq. (3.4), should be modified to reflect this different type of binding,

$$\langle T_F(N)\rangle_{\text{adj}} = N\int_0^\infty t f_F(t)\cdot\left[\int_t^\infty f_F(t')\,dt'\right]^{N-1} dt \quad , \tag{3.78}$$

where the subscript in $\langle T_F\rangle$ indicates that this is the average first binding time for a folding process where binding occurs only between edges.

As before, we assume that for a planar IC, the asymptotic solution of Eq. (3.78) for large N should be

proportional to the inverse of the logarithm of the number of competing 2D Brownian processes,

$$\langle T_F(N)\rangle_{\text{adj}} \propto \frac{1}{\ln(N)} \quad . \tag{3.79}$$

The last binding event for folding with edge only binding is no longer simply defined by a set of 1D Brownian processes, but rather a mixing of 1D and 2D ones, since the remaining $N-2$ lateral faces can still either bind in pairs or one by one if their adjacent faces have already bound. Estimating the average last binding time under these conditions then becomes a more intricate task.

In Fig. 3.14 we present a comparison between the numerical results of the average first and last binding times, as well as the average folding time, of folding processes with edge only and nonspecific binding. In Fig. 3.14a we also fit the numerical data of the edge only binding case with the following linear model,

$$y(x) = \tau_{F_0,\text{adj}} + \tau_{F,\text{adj}} x \quad , \tag{3.80}$$

using the method of linear least squares, where $x = {}^{1}\!/_{\ln(N)}$, while $\tau_{F_0,\text{adj}}$ and $\tau_{F,\text{adj}}$ are the fitting parameters.

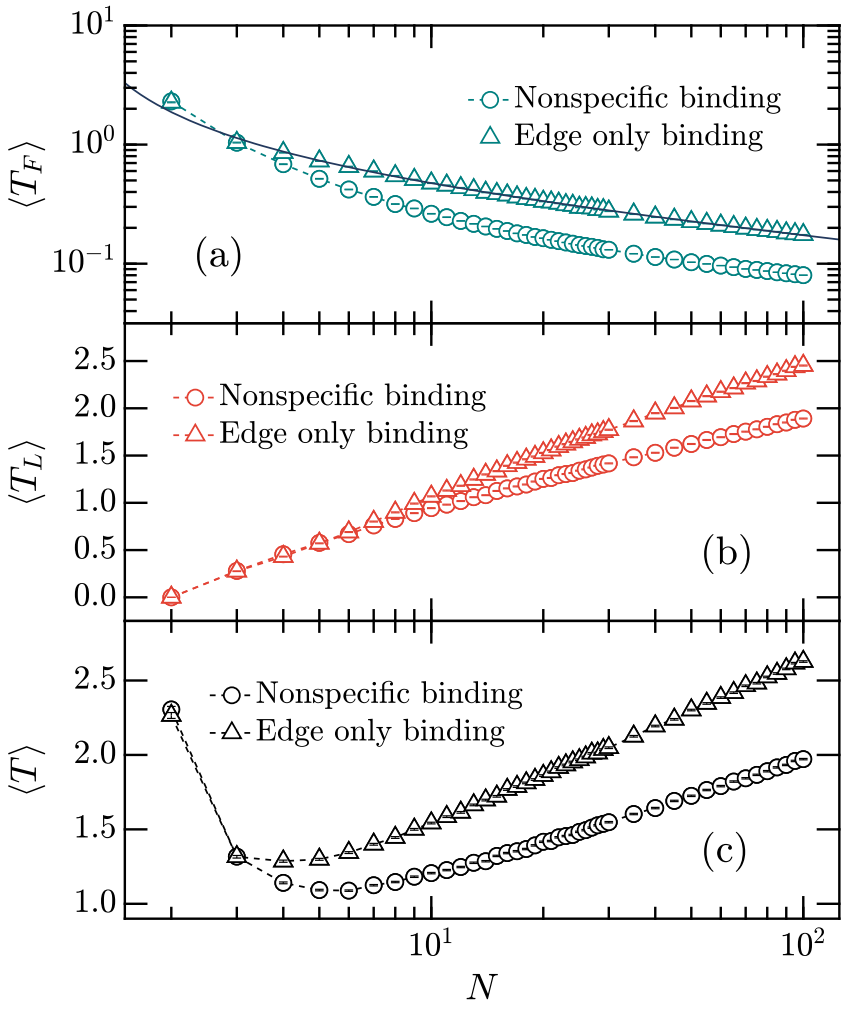

Figure 3.14: Comparison of numerical results for the average binding times $\langle T_F \rangle$ (a), $\langle T_L \rangle$ (b) and average folding time $\langle T \rangle$ (c) as a function of the number of lateral faces $N \in [2, 100]$ for a planar IC between simulations where binding is nonspecific (circle points) and simulations where binding can only occur through edges (triangle points). The solid line in (a) is given by Eq. (3.80), with fitting parameters $\tau_{F_0,\text{adj}} = -0.1278 \pm 0.0009$ and $\tau_{F,\text{adj}} = 1.387 \pm 0.004$ obtained from a linear least square fit of the data, ignoring the first four points $N = \{2, 3, 4, 5\}$. For both cases, the parameters of the simulations were in SC. Results are averages from $N_S = 10^4$ independent samples.

We can see that average first binding time when binding can only occur through edges has a slower decay, following the predicted the proportionality (3.79) well for large N.

For a random IC, if we assume again that the distribution $f_F(t)$ is approximately exponential,

$$f_F(t) \simeq \frac{1}{\tau_{F,R,\text{adj}}} e^{-\frac{t}{\tau_{F,R,\text{adj}}}} \quad , \tag{3.81}$$

with a characteristic decay time $\tau_{F,R,\text{adj}}$, the average first binding time given by Eq. (3.78) reads

$$\langle T_F(N) \rangle_{R,\text{adj}} \simeq \frac{N}{\tau_{F,R,\text{adj}}} \int_0^\infty t e^{-t\frac{N}{\tau_{F,R,\text{adj}}}} dt = \frac{\tau_{F,R,\text{adj}}}{N} \quad . \tag{3.82}$$

We present in Fig. 3.15 the numerical results for the average first and last binding times, as well as the average folding time, from a folding with edge only binding, with a random IC. In Fig. 3.15a we fit the data of both binding type cases to the following power law,

$$y\left(N\right) = AN^{k} \quad , \tag{3.83}$$

with parameters A and k. The fitting is done by the method of linear least squares, by taking the logarithm of Eq. (3.83).

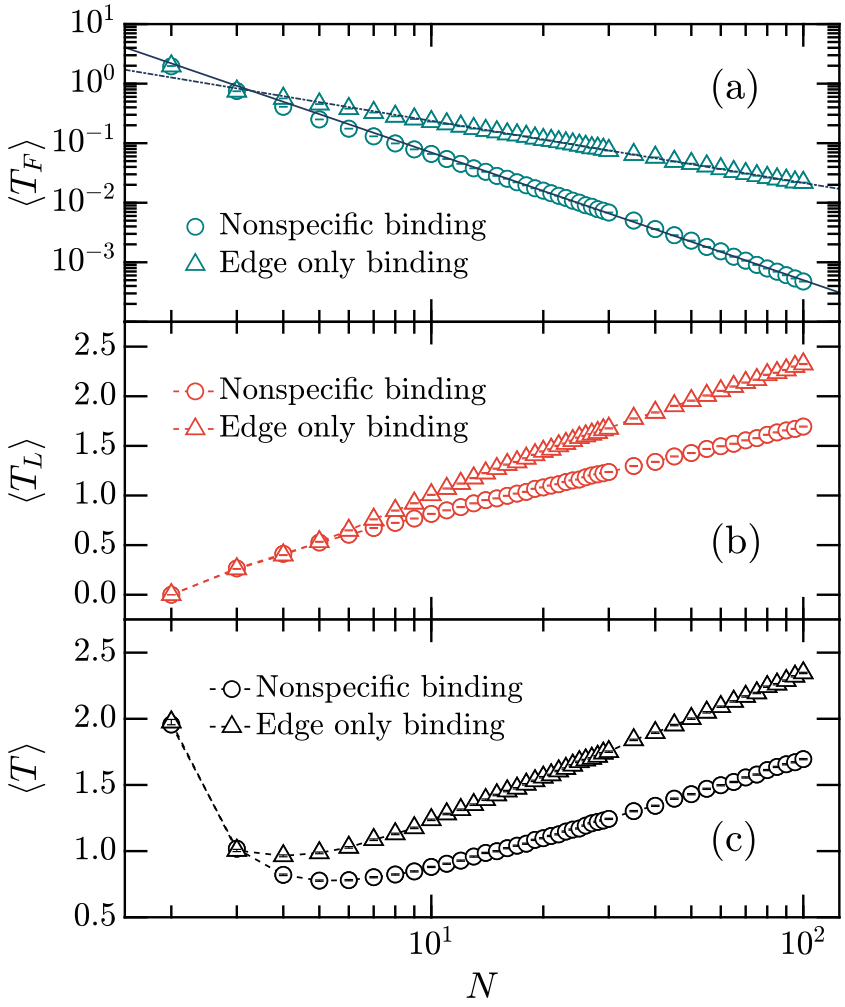

Figure 3.15: Comparison of numerical results for the average binding times $\langle T_F \rangle$ (a), $\langle T_L \rangle$ (b) and average folding time $\langle T \rangle$ (c) as a function of the number of lateral faces $N \in [2, 100]$ for a random IC between simulations where binding is nonspecific (circle points) and simulations where binding can only occur through edges (triangle points). The solid and dashed lines in (a) are given by Eq. (3.83), with fitting parameters (A, k) obtained from a linear least square fit of the data, ignoring the points up to $N = 12$ (exclusive), the value for which the relative difference ${}^{N^2}/_{(N\cdot(N-1))} - 1$ is less than 10%. The values of these parameters are $(2.27 \pm 0.01, -2.143 \pm 0.004)$ and $(0.96 \pm 0.01, -1.042 \pm 0.003)$ for the nonspecific and edge only binding cases, respectively. For both cases, the parameters of the simulations were in SC. Results are averages from $N_S = 10^4$ independent samples.

From Figs. (3.14) and (3.14) we can see that, for a random IC, the average first binding time decays with N as expected when N is large, for both cases. When binding is nonspecific, we have seen that in the asymptotic limit this average time should decay as N^{-2}, while for edge only binding it should be proportional to N^{-1} instead in this same limit. The fitted parameter k of the power law (3.83) seems to corroborate this. The small discrepancies from the integer values of two and one, mainly for the nonspecific binding type, may be due to the fact that the assumed FPT distribution is only an approximation. We also observe a slight shift to the left of the minimum of $\langle T \rangle$ when the faces can only bind between the edges, whilst this minimum also has a higher value. This is due to a slower decay of $\langle T_F \rangle$ with N when lateral faces bind only between edges, as opposed when binding is nonspecific, since $\langle T_L \rangle$ remains similar in this regime of low N between the two binding types.

3.4 Dependence of the average folding time on the initial conditions

We now study how different initial conditions affect the average folding time. Until now we have performed analysis for two ICs, the planar and random one. Here we present four additional ones and compare the average binding and folding times for the six ICs, both in a pinned and suspended system. To prevent misfolding, we consider only the cases with two and three lateral faces. We list all six ICs below, introducing again the planar and random ones for convenience. The first paragraph describes the state of the kirigami structure for that IC, while the second one explains how the IC is implemented in the lattice model. In this section we rename the usual random IC as 'Random 2' to distinguish it from two other ICs which will be introduced here, who also define an initial configuration with lateral faces starting with an outer angle uniformly distributed in the angular space.

Planar the folding starts from a flat template;

all N particles are allocated at the zeroth site of the lattice, for a pinned system, so

$$N_n\left(t=0\right) = N\delta_{n,0} \quad ; \tag{3.84}$$

if it is a suspended system, all N particles are placed in the middle of the lattice, calculated as $a_m = \lfloor L/2 \rfloor$, in which case

$$N_n\left(t=0\right) = N\delta_{n,a_m} \quad ; \tag{3.85}$$

Random 1 one lateral face starts with a zero outer angle, coplanar with the base, and the remaining ones start with an outer angle uniformly distributed over the angular space;

for N particles, the first one is allocated either at the zeroth site, if it is a pinned system, or at the middle site a_m, for a suspended one; for the remaining $N-1$ particles, an integer random number $r_L \in [0, L_F - 1]$ is generated, where L_F is the number of sites where a particle may still be placed, to choose a site from the list S_n of available sites where the particle is then allocated; before generating any value of r_L, $L_F = L$ and $S_n = \{0, 1, ..., L-1\}$; for each output i of r_L, if the randomly selected site a_i corresponding to the list element $S_n\left(r_L^{(i)}\right)$ is $a_i \geq n(\phi)$, for a pinned system, or $a_i \leq n(-\phi) \vee a_i \geq n(\phi)$, for a suspended one, subtract one unit from L_F, and update the list S_n of available sites by removing the one that just got chosen, $S_n \to S_n \backslash S_n\left(r_L^{(i)}\right)$; with the set of randomly selected sites $\{a_1, a_2, ..., a_{N-1}\}$, for a pinned system, the N particles are distributed according to

$$N_n(t=0) = \delta_{n,0} + \sum_{i=1}^{N-1} \delta_{n,a_i} \quad , \tag{3.86}$$

while for a suspended one we have

$$N_n(t=0) = \delta_{n,a_m} + \sum_{i=1}^{N-1} \delta_{n,a_i} \quad ; \tag{3.87}$$

Random 2 each lateral face starts from an outer angle uniformly distributed over the angular space;

each of the N particles are distributed randomly according to the outputs of the integer random variable r_L, as defined in the random 1 IC; with the set of randomly selected sites $\{a_1, a_2, ..., a_N\}$, the N particles are distributed according to

$$N_n(t=0) = \sum_{i=1}^{N} \delta_{n,a_i} \quad ; \tag{3.88}$$

Random 3 one lateral face starts with an outer angle of π, touching the base from above, and the remaining ones start with an outer angle uniformly distributed over the angular space;

for N particles, the first one is allocated at $L-1$, and the remaining $N-1$ are distributed randomly according to the outputs of the integer random variable r_L, as defined in the random IC; for this IC however, before generating any value of r_L, $L_F = L-1$ and $S_n = \{0, 1, ..., L-2\}$ contains all lattice sites except $L-1$; with the set of randomly selected sites $\{a_1, a_2, ..., a_{N-1}\}$,

the N particles are distributed according to

$$N_n\left(t=0\right)=\delta_{n,L-1}+\sum_{i=1}^{N-1}\delta_{n,a_i}\quad ; \tag{3.89}$$

Stacking all lateral faces are stacked on top of each other at upper side of the base, with an outer angle difference between each other of $\Delta\theta$, chosen to match the lattice spacing for the SC; if the number of lateral faces is enough for their cumulative outer angle difference to surpass the value of the closing angle (so $N > \lfloor {}^{(\pi-\phi)}/_{\Delta\theta} \rfloor$), the remaining $N - \lfloor {}^{(\pi-\phi)}/_{\Delta\theta} \rfloor$ particles all have an outer angle of $\phi - \Delta\theta$;

for N particles, from the first ensemble of $L - n\left(\phi\right)$ particles, the i-th one is allocated at site $L-i$, while the remaining $N-\left(L-n\left(\phi\right)\right) \equiv \mathcal{N}$ are all allocated at site $n\left(\phi\right)-1$; the initial distribution then reads

$$N_n\left(t=0\right)=\sum_{i=1}^{N-\mathcal{N}\xi}\delta_{n,L-i}+\mathcal{N}\delta_{n,n(\phi)-1}\xi\quad , \tag{3.90}$$

where $\xi = H\left[\mathcal{N}\right]$ and $H\left[x\right]$ is the discrete Heaviside step function,

$$H\left[x\right]=\begin{cases}0 & \text{, if } x\leq 0\\ \\ 1 & \text{, if } x>0\end{cases}\quad \text{, for } x\in\mathbb{Z}; \tag{3.91}$$

if the number of particles is not larger than the number of sites in $n \in \left[n\left(\phi\right), L-1\right]$ the initial distribution (3.90) simplifies to

$$N_n\left(t=0\right)=\sum_{i=1}^{N}\delta_{n,L-i}\quad \text{, for } N\leq L-n\left(\phi\right); \tag{3.92}$$

Opposite stacking for a pinned system, half the lateral faces are stacked on top of each other at the upper side of the base, until their cumulative outer angle difference surpasses the value of the closing angle , at which point the remaining ones of this half all have an outer angle of $\phi - \Delta\theta$, while the other half of lateral faces all have a zero outer angle; for a suspended system, one half is stacked at the upper side of the base, while the other half is stacked at the lower side, until the cumulative outer angle difference of each stack surpasses the value of the closing angle, at which point the remaining lateral faces of each half all have a respective outer angle of $\phi - \Delta\theta$ for the upper stack, and $-\phi + \Delta\theta$ for the lower one; for both systems, the lateral faces stacked on each

side have an outer angle difference of $\Delta\theta$; if the number of lateral faces is odd, the additional lateral face is always stacked on the upper side;

for N particles, for a pinned system, from the first $\lfloor {}^{(N+1)}/_{2} \rfloor$ particles, the i-th particle from the ensemble of $L - n(\phi)$ particles is placed at site $L - i$, while the remaining $\lfloor {}^{(N+1)}/_{2} \rfloor - (L - n(\phi)) \equiv \mathcal{N}_p$ are all allocated at site $n(\phi) - 1$; the other $\lfloor {}^{N}/_{2} \rfloor$ particles are all allocated at the zeroth site; the initial distribution is then

$$N_n(t=0) = \left\lfloor \frac{N}{2} \right\rfloor \delta_{n,0} + \sum_{i=1}^{\lfloor (N+1)/2 \rfloor - \mathcal{N}_p \xi_p} \delta_{n,L-i} + \mathcal{N}_p \delta_{n,n(\phi)-1} \xi_p \quad , \tag{3.93}$$

where $\xi_p = H[N - 2(L - n(\phi))]$; if the number of particles is not larger than twice the number of sites in $n \in [n(\phi), L-1]$ the initial distribution (3.93) simplifies to

$$N_n(t=0) = \left\lfloor \frac{N}{2} \right\rfloor \delta_{n,0} + \sum_{i=1}^{\lfloor (N+1)/2 \rfloor} \delta_{n,L-i} \quad \text{, for } N \leq 2(L - n(\phi)); \tag{3.94}$$

for a suspended system, from the first $\lfloor {}^{(N+1)}/_{2} \rfloor$ particles, the i-th particle from the ensemble of $n(-\phi)+1$ particles is placed at site $L-i$, while the remaining $\lfloor {}^{(N+1)}/_{2} \rfloor - (n(-\phi)+1) \equiv \mathcal{N}_u$ are all allocated at site $n(\phi) - 1$; from the other $\lfloor {}^{N}/_{2} \rfloor$ particles, the j-th particle (with j starting at $\lfloor {}^{(N+1)}/_{2} \rfloor$) from the ensemble $n(-\phi) + 1$ is placed at site $j - \lfloor {}^{(N+1)}/_{2} \rfloor$, while the remaining $\lfloor {}^{N}/_{2} \rfloor - (n(-\phi)+1) \equiv \mathcal{N}_l$ are all allocated at site $n(-\phi) + 1$; the initial distribution is then

$$\begin{aligned} N_n(t=0) = & \sum_{i=1}^{\lfloor N/2 \rfloor - \mathcal{N}_l \xi_s} \delta_{n,i-1} + \mathcal{N}_l \delta_{n,n(-\phi)+1} \xi_s \\ & + \sum_{i=1}^{\lfloor (N+1)/2 \rfloor - \mathcal{N}_u \xi_s} \delta_{n,L-i} + \mathcal{N}_u \delta_{n,n(\phi)-1} \xi_s \quad , \end{aligned} \tag{3.95}$$

where $\xi_s = H[N - 2(n(-\phi)+1)]$; if the number of particles is not larger than twice the number of sites in $n \in [n(\phi), L-1]$ the initial distribution (3.95) simplifies to

$$N_n(t=0) = \sum_{i=1}^{\lfloor N/2 \rfloor} \delta_{n,i-1} + \sum_{i=1}^{\lfloor (N+1)/2 \rfloor} \delta_{n,L-i} \quad \text{, for } N \leq 2(n(-\phi)+1); \tag{3.96}$$

The initial configurations of a pyramid net with $N = 3$ for these six ICs are schematized in Fig. 3.16.

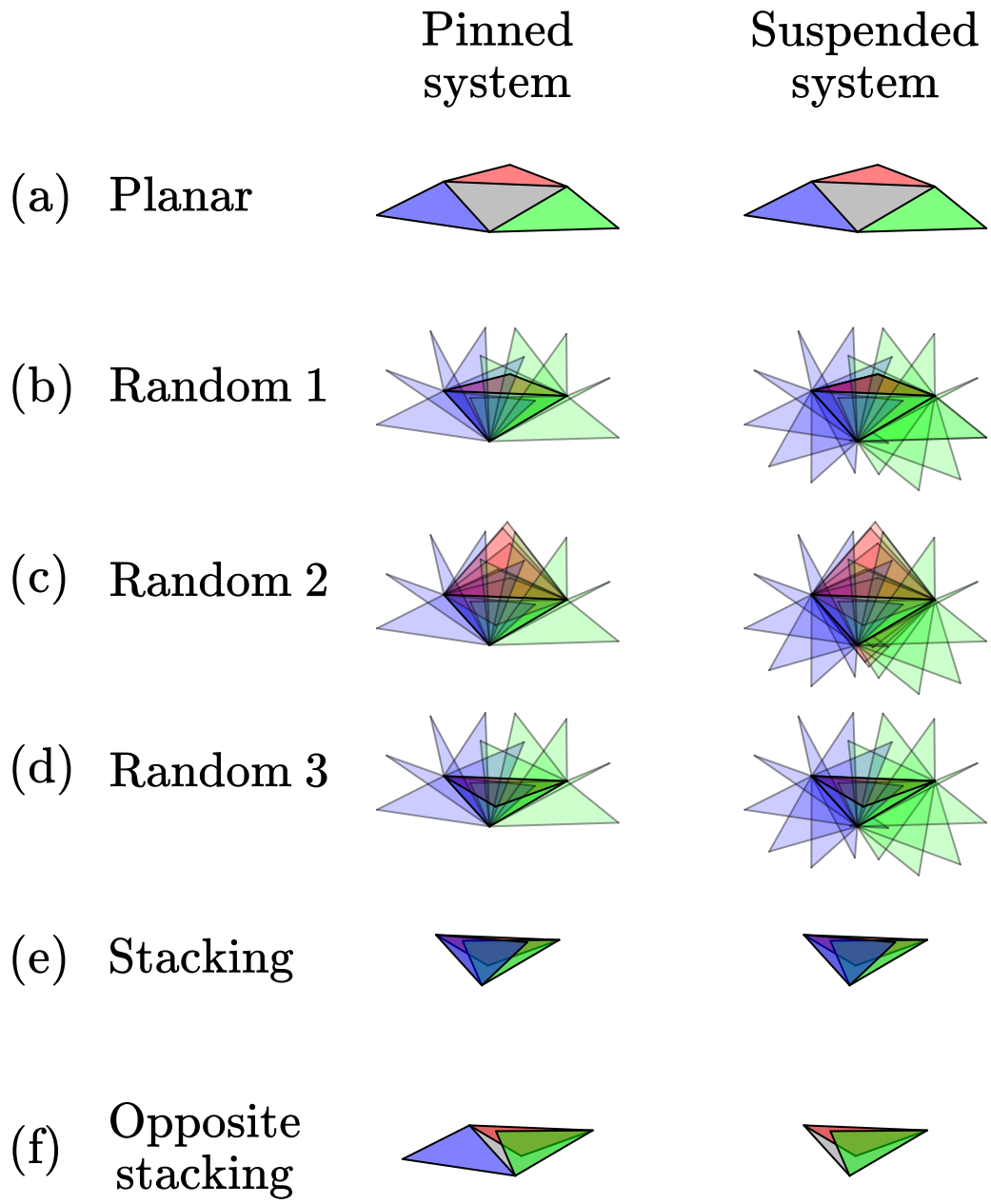

Figure 3.16: Schematic representations of the initial configuration of a pyramid net with $N = 3$ lateral faces and closing angle $\phi = {}^{2\pi}/_{3}$ for the six ICs presented in this section, for both a pinned system (left column) and a suspended one (right column). The overall transparent lateral faces in ICs (b-d) serve to portray the uniform random distribution of their outer angles. The two ICs (a,e) are completely identical between the two types of systems. The third lateral face in IC (f) for a suspended system is not visible since it is touching the base from below.

In Fig. 3.17 we present the average folding times with $N = 2$ for these six different ICs, both for a pinned and suspended system, while in Fig. 3.18 we present the results for $N = 3$. The distributions from which these averages were computed are presented in Figs. E.1-E.4 of appendix E.

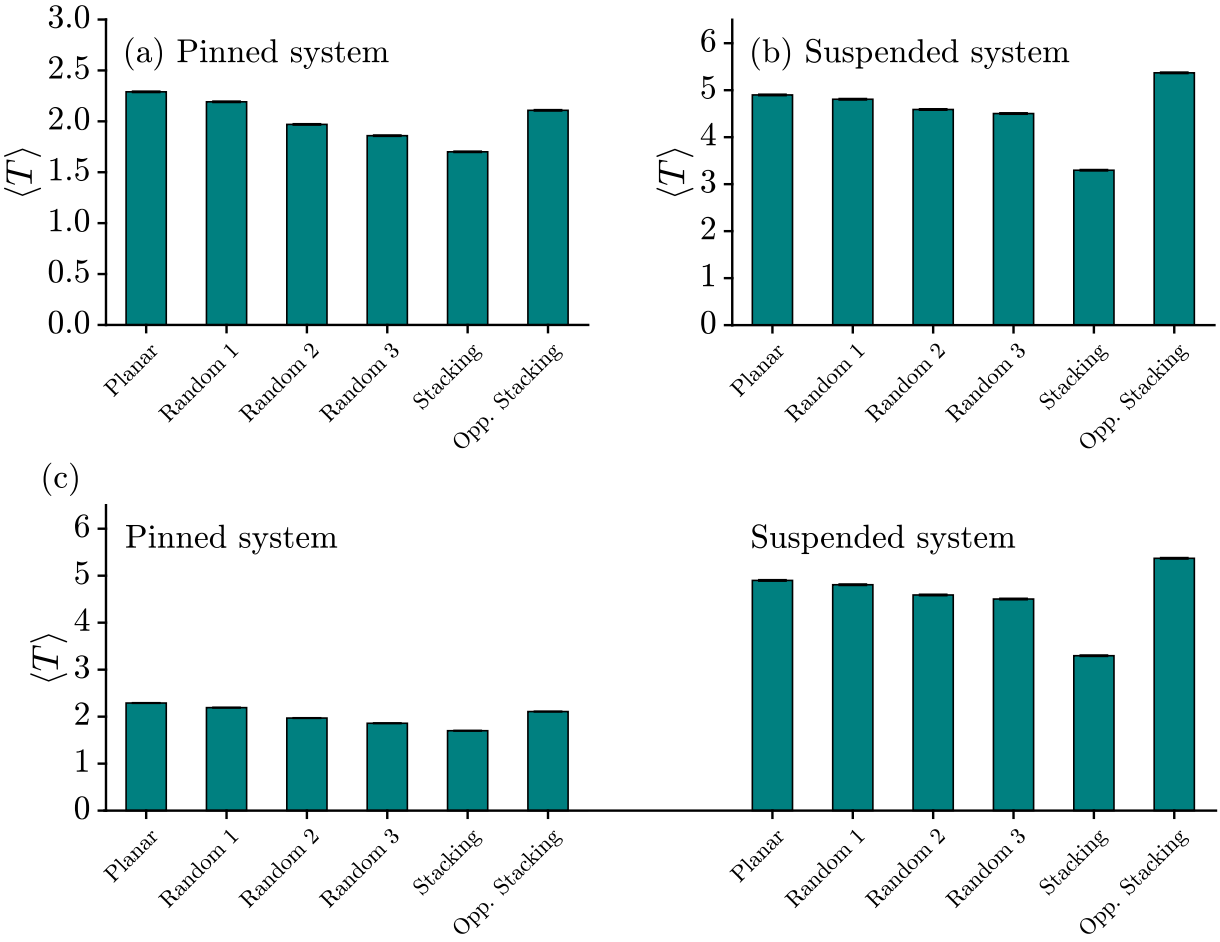

Figure 3.17: Comparison of the average folding times $\langle T \rangle$ for $N = 2$ between six different ICs, for both a pinned (a) and suspended (b) system. The bar plot (c) shows the same results of (a) and (b) but with a common axis scale for a better comparison between the average times of a pinned system and a suspended one. The simulations were performed with SC. Results are averages from $N_S = 10^5$ independent samples. The averages in bar plot (a) and (b) are respectively calculated from the distributions in Figs. E.1 and E.2 of appendix E.

From Fig. 3.17a we see that, for a pinned system with $N = 2$, the planar IC has the slowest average folding time, while the stacking IC has the fastest. The random 2 IC has an average folding time between two other ICs which also have one of the lateral faces starting from an outer angle randomly chosen from a uniform distribution but the other starts from a definite angle. The random 1 IC is slower than the random 2 one, since one of the lateral faces starts at zero outer angle, slowing down the folding process. On the other hand, the random 3 IC has one lateral face starting from an outer angle of π, touching the base from the upper side, and leads to a faster folding when compared to the random 2 IC, speeding up the process.

For a suspended system with $N = 2$, we see in Fig. 3.17b that, while the stacking IC remains the fastest, the slowest average time is now attributed to the opposite stacking IC. This is intuitive since in this type of system, for this IC, the lateral faces start the folding from complete opposite sides, with the farthest

possible angular distance of 2π between them. For the other ICs they seem to retain their order as it was for the pinned system, with slight changes on their relative differences.

Between the pinned and suspended systems, the change on the average folding time is significant. All ICs for a suspended system have an average folding time at least twice slower as those for the pinned system.

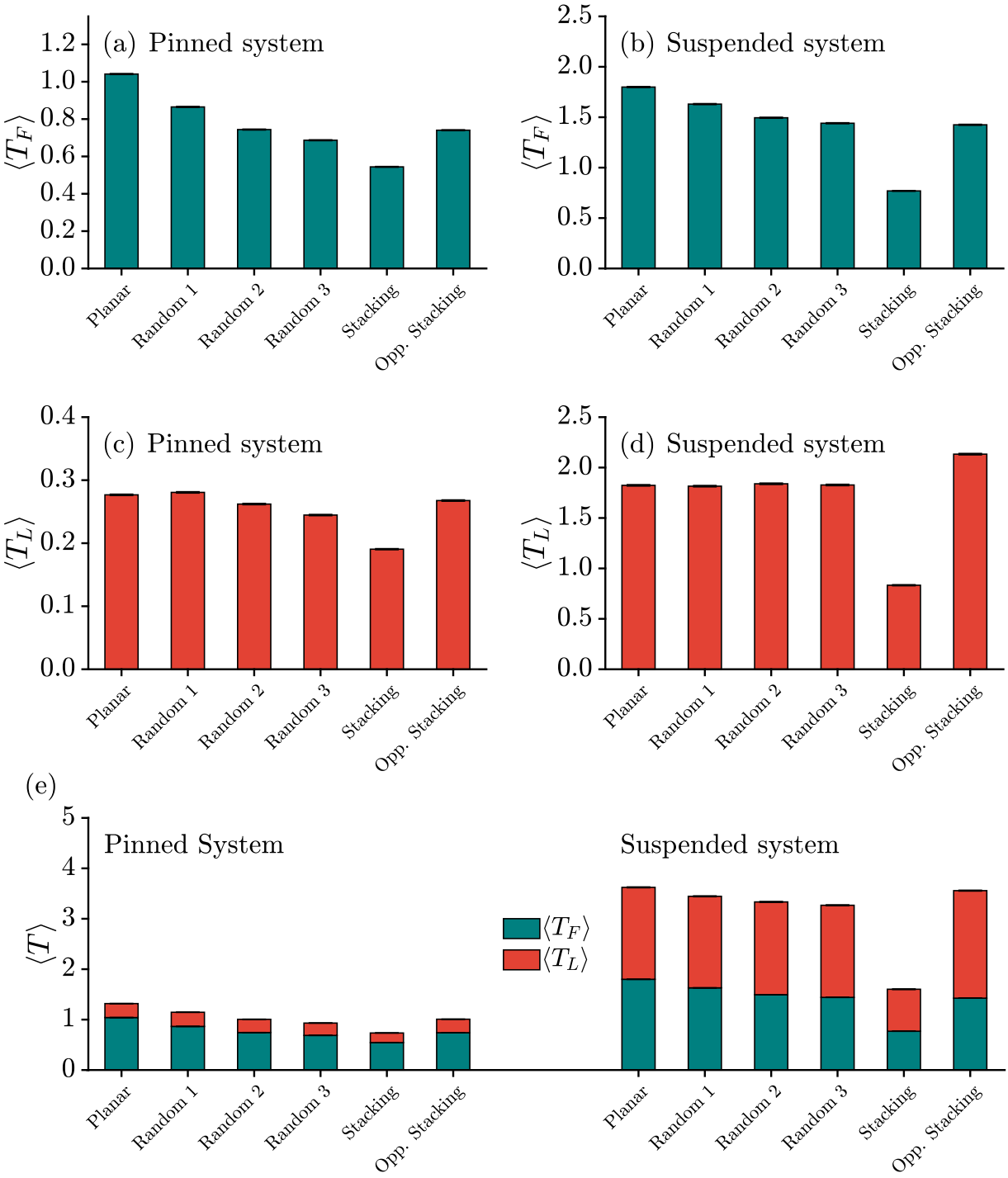

Figure 3.18: Comparison of the average first $\langle T_F \rangle$ (a,b) and last $\langle T_L \rangle$ (c,d) binding times for $N = 3$ between six different ICs, for both a pinned (a,c) and suspended (b,d) system. The bar plot (e) shows the average folding time $\langle T \rangle$ computed from the sum of the numerical results for $\langle T_F \rangle$ and $\langle T_L \rangle$ of plots (a,b,c,d), for the respective pinned or suspended system. These averages are, as supposed, consistent with the expected values of the respective folding time distribution in Figs. E.3 and E.4 of appendix E. The uncertainty in the values of $\langle T \rangle$ were calculated by a sum in quadrature of the uncertainties of the respective values of $\langle T_F \rangle$ and $\langle T_L \rangle$. The simulations were performed with SC. Results are averages from $N_S = 10^5$ independent samples. The averages in (a) and (c) are calculated from the distributions of Fig. E.3 of appendix E, while the ones in (b) and (d) are computed from Fig. E.4 of the same appendix.

As seen from Fig. 3.18a and 3.18b for a pinned and suspended system, respectively, with $N = 3$, the

average first binding times for the six ICs seem to have similar relative behaviour as the ones for the analogous case of $N = 2$, with the fastest average time again belonging to the stacking IC. A notorious exception is the opposite stacking IC for the suspended system, which is now as fast as the random 3 IC. This is due to the fact that now we have two lateral faces starting the folding at the upper side, in contrast with the $N = 2$ case, where the only two lateral faces were 2π apart in the angular space. This proximity likely leads to a higher chance of these two faces binding as soon as their respective outer angle each reach the region of the closing angle. This leaves the planar IC with the slowest average first binding time for both pinned and suspended systems.

In Fig. 3.18c we see that, for a pinned system with $N = 3$, the IC with the slowest average last binding time belongs to either the planar or random 1 IC, while for the suspended system, we see from Fig. 3.18d that the slowest is again the opposite stacking IC. For both types of system, the fastest average last binding time is yet again the one for the stacking IC. It is also noteworthy that, in the suspended system, the average of this time varies little between the planar, random 1, random 2 and random 3 ICs.

In Fig. 3.18e we plot the average folding time, the sum of the average first and last binding times, for the pinned and suspended systems with $N = 3$. A significant difference is observed for the average of folding time between these two systems. Partly it is due to the average first binding time, which becomes slower in the suspended system, but never more than twice the value it has for the pinned one. The major contribution to this difference comes from the average last binding time, where the times become several times slower in the suspended system when compared with the pinned one.

3.5 Conditional probabilities of folding in the upper or lower side

In this section we measure the propensity for the kirigami structure to self-fold without defects as we vary the number of lateral faces N. We recall the definition of a defective self-folding as a final structure which is open, i.e. if lateral faces bind at both sides. For this section, all analyses are performed in a suspended system.

We start by studying the conditional probability $P\left(\text{down} \mid \text{up}\right)$ that the second binding event occurs at the lower side ($\theta = -\phi \pm {}^{\Delta\phi}/_{2}$) provided the first binding occurred at the upper side ($\theta = \phi \pm {}^{\Delta\phi}/_{2}$), and the conditional probability $P\left(\text{up} \mid \text{down}\right)$ that the second binding occurs at the upper side provided that the first occurred at the lower one. For a diffusion without drift, we expect that these two conditional

probabilities are equal if the IC is either planar or random. The unconditional probabilities for each of these two events to occur is then given respectively by $P_{\text{upper}}P\left(\text{down} \mid \text{up}\right)$ and $P_{\text{lower}}P\left(\text{up} \mid \text{down}\right)$, where P_{upper} and P_{lower} are the probabilities for the first binding to occur at the upper and lower side, respectively. For a diffusion without drift, $P_{\text{upper}} = P_{\text{lower}} = 1/2$ as we have seen from the results in Fig. 2.1. Therefore the probability for the final folded structure to already become defective as soon as the second binding occurs is given by $1/2\left(P\left(\text{down} \mid \text{up}\right) + P\left(\text{up} \mid \text{down}\right)\right)$, since we assume that this process is irreversible. The complementary conditional probabilities $P\left(\text{up} \mid \text{up}\right)$ and $P\left(\text{down} \mid \text{down}\right)$ of both first and second binding events occurring in the upper and lower side, respectively, are given by

$$P\left(\text{up} \mid \text{up}\right) = 1 - P\left(\text{down} \mid \text{up}\right) \quad , \tag{3.97}$$

$$P\left(\text{down} \mid \text{down}\right) = 1 - P\left(\text{up} \mid \text{down}\right) \quad , \tag{3.98}$$

and $1/2\left(P\left(\text{up} \mid \text{up}\right) + P\left(\text{down} \mid \text{down}\right)\right)$ is, by an analogous reasoning, the probability for the self-folding structure to still not have any defects after the second binding.

To estimate the values of $P\left(\text{down} \mid \text{up}\right)$ and $P\left(\text{up} \mid \text{down}\right)$ we perform two Bernoulli trials,

$$P\left(\text{down} \mid \text{up}\right) \simeq \frac{B_{\text{down}\mid\text{ up}}}{N_S} \quad \text{, for large } N_S, \tag{3.99}$$

$$P\left(\text{up} \mid \text{down}\right) \simeq \frac{B_{\text{up}\mid\text{ down}}}{N_S} \quad \text{, for large } N_S, \tag{3.100}$$

where $B_{\text{down}\mid\text{ up}}$ ($B_{\text{up}\mid\text{ down}}$) is the number of times the second binding occurred at the lower (upper) side, provided that the first binding occurred at the upper (lower) side, out of a total of N_S trials. We present the numerical results of $P\left(\text{down} \mid \text{up}\right)$ and $P\left(\text{up} \mid \text{down}\right)$ for several values of N in Fig. 3.19 for a planar and random IC.

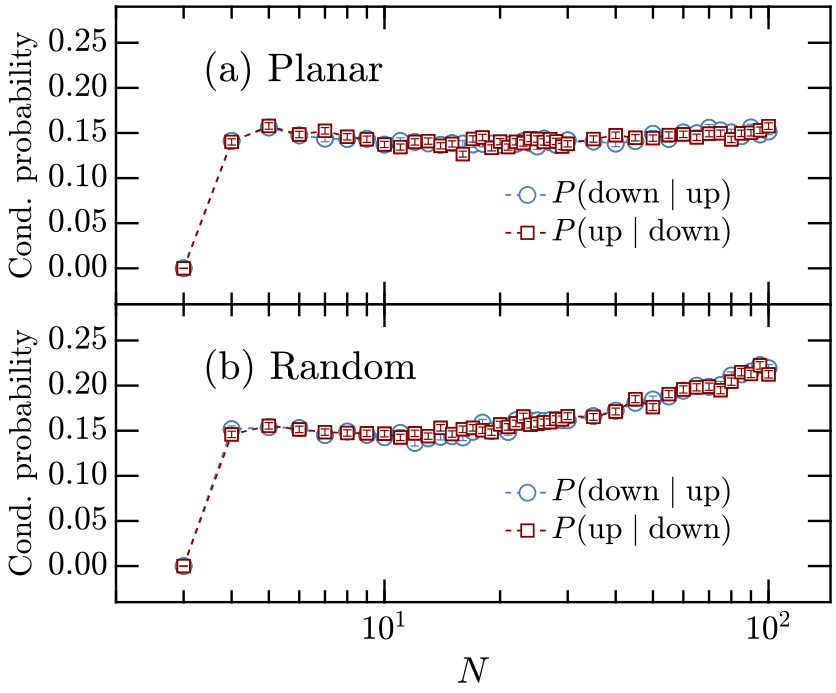

Figure 3.19: Simulation results in a suspended system for the conditional probabilities (3.99) (circle points) and (3.100) (square points) as a function of $N \in [3, 100]$, for a planar (a) and random (b) ICs. The simulations were performed under SC. Results are averages from $N_S = 10^4$ independent trials.

From these results we see that $P\,(\text{down} \mid \text{up}) \simeq P\,(\text{up} \mid \text{down})$, as expected from an unbiased diffusion. Therefore $P\,(\text{down} \mid \text{up})$ and $P\,(\text{up} \mid \text{down})$ directly represent the probability for the final folded structure to already become defective as soon as the second binding occurs. We also observe that this probability remains more or less constant as a function of N for the planar IC, at around 15% for $N > 3$, while for a random IC it slowly increases with N, going from approximately 15% at $N = 3$ to around 20% at $N = 100$.

Next we present in Fig. 3.20 the numerical results of the average first and last binding times, as well as the total total folding time, as a function of N for a suspended system. All samples, with or without defects, were used in the calculation of these time averages.

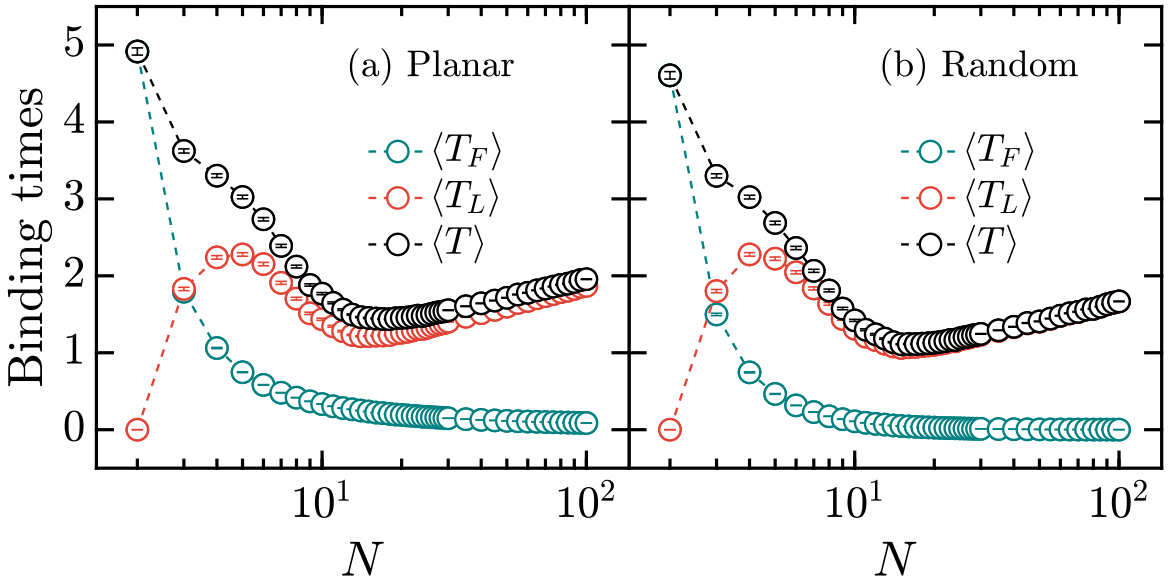

Figure 3.20: Numerical results for the average binding times $\langle T_F \rangle$, $\langle T_L \rangle$ and average folding time $\langle T \rangle$, in Brownian time units, as a function of the number of lateral faces $N \in [2, 100]$ for a planar (a) and random (b) ICs, for a suspended system. In this type of system defects on the folded structure are possible, and both misfolded and defect free samples were considered in the calculation of these times averages. The simulations were performed with SC. Results are averages from $N_S = 10^4$ independent samples.

Results of Fig. 3.20 show that, if the planar structure folds in suspension, the average last binding time is also a non-monotonically function of the lateral faces N, in addition to the total folding time. We believe this happens due to the larger size of the angular domain when compared to the pinned system, and to the additional closing angle at $\theta = -\phi \pm {}^{\Delta\phi}/_{2}$. For low values of N, up to around $N \simeq 5$ for the planar IC and $N = 4$ for the random one, the first binding takes a comparably long time to occur, allowing for the angular probability density to disperse over the larger domain $\theta \in [-\pi, \pi]$ for the planar IC. When the first binding occurs, the remaining lateral faces either bind at the same side of the first two or (if $N > 3$) a pair also happens to bind at the opposite side, in which case the remaining lateral faces can bind one by one at either side. Since N is low it is not very likely for another pairwise binding to occur at the opposite side, only around 15% for $N = 4$, as seen from the results of Fig. 3.19. The majorities of the lateral faces who wandered into the opposite side then need to return to the closing angle where the first binding occurred, leading to long trajectories in the angular space and a tendency for the average of the last binding time $\langle T_L \rangle$ to increase with N in this regime. After this local maximum in $\langle T_L \rangle$ this time average starts to decrease with N, up to around $N \simeq 14$ for a planar IC and $N \simeq 15$ for a random one. In this new regime, the angular probability density at either side of the domain starts to increase, and bindings at both sides are presumably not as unlikely. Therefore, as N increases, there is less chance for lateral faces to wander off far from a closing angle where faces have already bound. Finally, $\langle T_L \rangle$ starts to increase again with N after this regime. Here the number of lateral faces starts to saturate, and the

first two pairs quickly bind at both sides. Assuming each of the first two binding events happen at opposite sides, and that all remaining lateral faces have an outer angle $-\phi < \theta_i\,(t = t_2) < \phi$, this leads to a domain with both boundaries absorbing. At this point increasing N simply increases the number of lateral faces which have to reach one of these absorbing boundaries and, consequently, the average last binding time.

This distinct functional dependence with N for a suspended system leads to a minimum for the folding time shifted to the right, around $N \simeq 18$ for the planar IC and $N \simeq 15$ for the random one, when compared to the analogous results in a pinned system.

Next we analyse the propensity for a given regular pyramid template with N lateral faces to self-fold into the desired structure, free of defects. We start by introducing the folding efficiency η,

$$\eta = \frac{B_{\text{free}}}{N_S} \quad , \tag{3.101}$$

defined as the fraction of samples B_{free} that self-folded without defects out of a total of N_S samples. In Fig. 3.21a we present numerical results for the efficiency η as a function of N, for both planar and random ICs, for a sample size of $N_S = 10^4$.

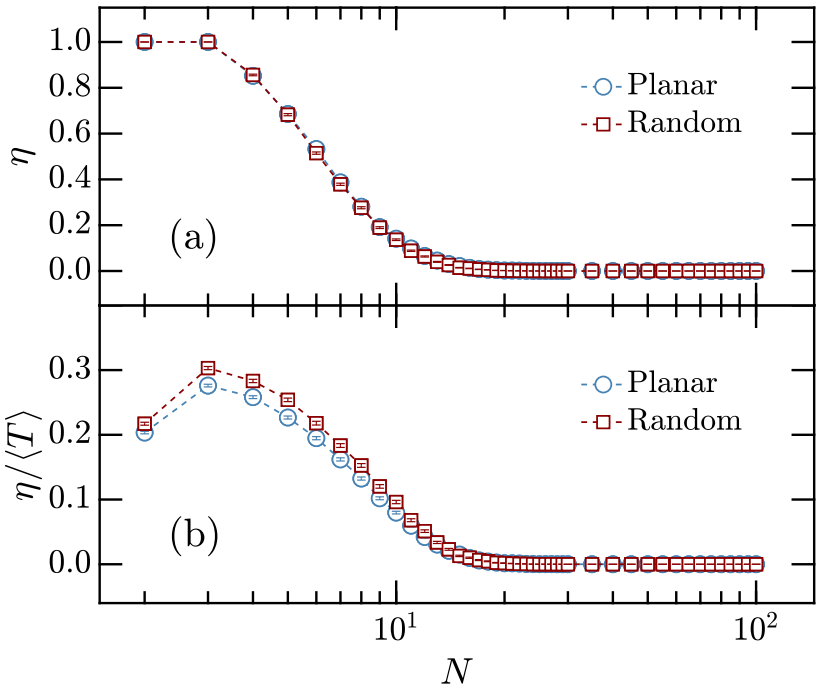

Figure 3.21: (a) Simulation results in a suspended system for the efficiency η, as defined by Eq. (3.101), as a function of $N \in [2, 100]$, for a planar (circle points) and random (square points) ICs, from a sample size of $N_S = 10^4$. (b) The same numerical values of the efficiency η as those in (a) but divided by the corresponding numerical value of the average folding time $\langle T \rangle$, the data set of black points presented in Fig. 3.20 for the respective IC. The simulations were performed under SC.

From Fig. 3.21a we see that the efficiency η does not depend on whether we choose a planar or random IC. For values of $N \gtrsim 7$ the fraction of samples free of defects is already lower than 50%, and for $N \geq 29$ not a single sample self-folded without defects out of a total of 10^4.

In Fig. 3.21b, we divide the numerical values of η of Fig. 3.21a by the average folding time $\langle T \rangle$ of the respective IC of Fig. 3.20, to estimate at which value of N we have an optimal balance between the folding time and efficiency. The uncertainties of these values are propagated from the ones in η and the respective $\langle T \rangle$ in a way analogous to Eq. (3.75). A maximum appears for $N = 3$, indicating that, even though the fastest folding times for suspended systems are achieved, on average, for pyramid templates with around $N \in [15, 18]$ lateral faces (as found in the results of Fig. 3.20) they are undermined by a less favourable efficiency $\eta \lesssim 0.7$, for both planar and random ICs. For $N \lesssim 10$, all results of this ratio for the random IC are larger than the ones for the planar case, since the time $\langle T \rangle$ is faster if the IC is random as opposed to planar (Fig. 3.20). For $N > 10$, the differences of this ratio between the two ICs start to become negligible as the efficiency approaches and eventually reaches a value of zero.

Conclusion

In this work we considered the stochastic folding of templates whose target structure was a regular pyramid with N lateral faces. We assumed that binding is irreversible and, for most of the work, that the base panel of the template is pinned to a substrate, so that misfolding is not possible. Taking advantage of the fact that the rotational motion of these lateral faces around the hinges is consistent with a Brownian process [13], we introduced a lattice model to simulate the angular position of each face, in order to study the time evolution of the folding process using the kMC method.

To validate the model, we compared numerical results with well known analytical predictions from random walks and diffusion theory. We also verified that our lattice model approaches the diffusion limit as we increase the number of sites, and the standard deviation of numerical averages varies accordingly with the sample size. Another assumption made was that lateral faces could interact elastically between each other when their outer angles surpassed the closing angle, the outer angle formed between the lateral faces and the base of the folded structure. However, in this work it was shown that these interactions have negligible effect on the folding time when the templates start the folding process either from a planar configuration or a random one, with its lateral faces uniformly distributed across the angular space. The most noticeable impact, albeit still small, occurs for a low number of lateral faces N, and quickly gets negligible as N increases.

With this model, it is possible to show numerically that the folding at the microscale is an intricate process which depends greatly on the geometry and initial configuration of the net, as well as the binding mechanism. Namely, the results showed that, if the template starts from a planar IC, the average first binding time for the first two lateral faces to bind decays as the inverse of the logarithm of N. If the template starts from a random IC, this time decays as N^{-2} instead. On the other hand, the average last binding time for the last face to bind grows with N for a planar IC. If the IC is random, this growth is slightly slower. These contrasting tendencies lead to a non-monotonic dependence with N of the average total folding time and, therefore, to an optimal number of lateral faces for fast folding, $N = 5$. Between the two ICs, the average

total folding time as a function of N changes only slightly, being faster for the random IC when compared to the planar one, with the optimal value of N unchanged.

Further control of the folding time can be achieved by manipulation of the closing angle. If the template starts from a planar IC, and the number of lateral faces is sufficiently high, most of the lateral faces still have an outer angle with a smaller value than the closing angle by the time the first two bind. In this high N regime, the average folding time grows with ϕ^2. For low values of N this scaling relation fails. For $N = 2$ the average folding time can be approximated by a MFPT to a target in a 2D lattice. For $N = 3$ the average first binding time should, in principle, be consistent with a MFPT to one of a multitude of targets in a 3D lattice, but a practical and elucidative analytical solution remains elusive when the number of targets is too large [41]. For the same N, a semi-analytical solution was found for the average last binding time which highlighted the dependency of this time with the conditional angular distribution of the free lateral face at the moment the first two bind. In turn, we found that this distribution depends non-trivially on the closing angle and the IC.

The type of binding mechanism also affects the folding time. If lateral faces can only bind adjacently through their edges, as opposed to a nonspecific vertex enabled binding, the overall average folding time is considerably longer for all N studied, with the exception of $N = \{2, 3\}$, since the folding process is indistinguishable between a nonspecific or the edge only binding for these two cases. While the shape of the average folding time as a function of N remains similar, with an optimal number of lateral faces still around $N \simeq 5$, the shape of the curve around this minimum is more flattened when compared to the nonspecific binding type. These remarks are valid independently of whether the template starts from a planar or random configuration.

Next, we studied how the average folding time depends on the initial configuration of the net. We additionally compared the results between a pinned and a suspended system, while restraining our analysis to the cases with $N = \{2, 3\}$ so that misfolding is still not possible even if the folding occurs in suspension. From the various ICs investigated, the one where all lateral faces start the folding process stacked on top of each other above the base yielded the fastest folding times on average. The slowest times are associated with either a planar initial configuration or an IC where each half of the total number of lateral faces are stacked on opposite sides of the angular space. If folding occurs in suspension, the average folding time more than doubles if $N = 2$, and more than triples if $N = 3$, when compared to the pinned system, with the particular exception for the stacking IC where this increase is slightly below the pattern for the respective N.

For the last analysis we focused on a system which folds in suspension, where misfolding was possible.

The results demonstrated that the probability for a sample to already be defective due to misfolding at the second binding event is independent of N for a planar IC, with only a slight increase with N if the IC is random. Under these conditions, the average last binding time and the overall average folding time are also a non-monotonic function of N. An optimal number of lateral faces N for fast folding also exist for a suspended system, shifted to higher values of N when compared with the pinned system counterpart, at $N \simeq 18$ and $N \simeq 15$ for the planar and random ICs, respectively. While the folding time gets faster as we approach this minimum, the folding efficiency always decreases as N gets larger, and so a balanced decision needs to be made when choosing the number of lateral faces, so that an optimal trade-off between the folding time and the number of defective samples is obtained according to specific needs.

This work showed how a lattice model can be a computationally efficient way to highlight and analyze the geometrical underpinnings which influence the dynamics of self-folding at the microscale. The possibilities have not been exhausted, however, as further studies can still be conducted. For example, it may also be of interest to study how a possible bias on the rotational motion of the lateral faces influences both the folding time and efficiency.

Existing methods for controlling the folding process rely on external mechanisms. The folding time of templates with cell coated panels can be controlled through the concentration and culturing time of these cells [29]. For templates with graphene-based hinges, the folding process can be manipulated by changing the temperature and pH of the surrounding environment [28]. The findings of this work show how inherent properties of templates can be exploited to expand the available means of controlling the time and efficiency of self-folding at the microscale.

Appendix A

Formulae of error and statistical analysis

In this appendix we present a set of formulae from which we compute the mean and uncertainty of our results (section A.1) and a method to fit a set of data to a linear model (section A.2). The contents of this appendix are based on [43, 44].

A.1 Reporting the results

Suppose we are trying to measure the mean value $\overline{x}$ of some random variable x whose possible outcomes are governed by some distribution $f(x)$ whose analytical form is, in general, unbeknownst to us a priori. This distribution $f(x)$ needs to be normalized

$$\int f(x)\,dx = 1 \quad , \tag{A.1}$$

and is characterized by its so called moments μ_n, with the n-th moment given by

$$\mu_n = \int x^n f(x)\,dx \quad . \tag{A.2}$$

The integration runs through the whole domain where $f(x)$ is defined.

If x is discrete, with possible discrete values x_i, each with an outcome probability of $f(x_i)$, the analogous form of equations (A.1) and (A.2) are

$$\sum_i f(x_i) = 1 \quad , \tag{A.3}$$

and

$$\mu_n = \sum_i x_i^n f(x_i) \quad , \tag{A.4}$$

respectively, where the sum is over all possible outcomes of x.

Lets suppose we now take N_S independent samples, $\{x_1, x_2, ..., x_{N_S}\}$, from our random variable x. We take as our best estimate of the average $\overline{x}$ the sample average

$$\langle x \rangle = \frac{1}{N_S} \sum_{i=1}^{N_S} x_i \quad . \tag{A.5}$$

To quantify our uncertainty $\delta\overline{x}$ in our best estimate of the average $\overline{x}$ we first note that each sample x_i will deviate from our sample average an amount we call a residue $\epsilon_i = x_i - \langle x \rangle$. A natural first try at quantifying the deviation δx of each sample x_i from our sample mean $\langle x \rangle$ might be summing all the residuals ϵ_i and average them by dividing by the number of samples N_S we took, but this method is faulty, since ϵ_i might be either positive or negative and may cancel each other through the summation. An alternative is taking the square of the residues ϵ_i, averaging them, and taking the square root of the result

$$\delta x = \sqrt{\frac{1}{N_S} \sum_{i=1}^{N_S} \epsilon_i^2} = \sqrt{\frac{1}{N_S} \sum_{i=1}^{N_S} (x_i - \langle x \rangle)^2} = \sigma_x \quad , \tag{A.6}$$

which is formally known as the standard deviation σ_x,[1] whose square, σ_x^2, is called the (sample) variance.

Before proceeding into finding an equation for $\delta\overline{x}$, we remark that definition (A.6) is rather inconvenient for computational implementation since it needs to go through all the sample data twice, first to calculate the sample mean $\langle x \rangle$ and then to calculate the residuals ϵ_i. Furthermore, the sample mean is usually not a whole number, so the deviations from the mean are also not whole numbers, and the computation gets error-prone. Instead we reformulate Eq. (A.6) by expanding the square of the residual,

$$\begin{aligned} \sigma_x &= \sqrt{\frac{1}{N_S} \sum_{i=1}^{N_S} (x_i - \langle x \rangle)^2} = \sqrt{\frac{1}{N_S} \sum_{i=1}^{N_S} \left(x_i^2 + \langle x \rangle^2 - 2 \langle x \rangle x_i \right)} = \\ &= \sqrt{\frac{1}{N_S} \sum_{i=1}^{N_S} x_i^2 + \frac{\langle x \rangle^2}{N_S} \sum_{i=1}^{N_S} 1 - \frac{2 \langle x \rangle}{N_S} \sum_{i=1}^{N_S} x_i} = \sqrt{\langle x^2 \rangle + \langle x \rangle^2 - 2 \langle x \rangle^2} = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \quad , \end{aligned} \tag{A.7}$$

[1] The standard deviation as defined in equation (A.6) is more precisely known as the population standard deviation. In our case where we calculate σ_x using a finite number of samples we should be using the sample standard deviation instead, where the factor N_S in the denominator is replaced with $N_S - 1$, with the rationale that there are only $N_S - 1$ independent residuals since their sum, by definition of residuals and the sample mean, must equate to zero, i.e. $\sum_{i=1}^{N_S} \epsilon_i = 0$. However, the difference between the factor N_S and $N_S - 1$ is negligible when N_S is of the order of 10^4 or greater, as is the case in this work, so we adopt definition (A.6).

where we defined the squared sample average

$$\left\langle x^2 \right\rangle = \frac{1}{N_S} \sum_{i=1}^{N_S} x_i^2 \quad . \tag{A.8}$$

If we now treat the outcomes $\{x_1, x_2, ..., x_{N_S}\}$ as independent random variables all with the same distribution, hence the same mean $\langle x \rangle$ (as given by Eq. (A.5) in the limit $N_S \to \infty$) and variance σ_x^2 (as given by the square of Eq. (A.7) in the limit $N_S \to \infty$), and define a new random quantity X as the sum of all the N_S independent random variables

$$X = \sum_{i=1}^{N_S} x_i \quad . \tag{A.9}$$

and compute the variance σ_X^2 of X, using $\langle ... \rangle$ as a notation to denote the operation of averaging whatever random quantity is inside the brackets,

$$\begin{aligned} \sigma_X^2 = \left\langle X^2 \right\rangle - \langle X \rangle^2 = \sum_{i,j=1}^{N_S} \left(\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \right) = \\ = \sum_{i=1}^{N_S} \left(\langle x_i x_i \rangle - \langle x_i \rangle \langle x_i \rangle \right) = N_S \cdot \left(\left\langle x^2 \right\rangle - \langle x \rangle^2 \right) = N_S \sigma_x^2 \quad , \end{aligned} \tag{A.10}$$

where we used the fact that, when $i \neq j$, $\langle x_i x_j \rangle = \langle x_i \rangle \langle x_j \rangle$ since we assume that the data is statistically independent.

The best estimate of the mean value $\overline{x}$ of our original random variable x is just (A.9) divided by N_S, and so the variance $\sigma_{\overline{x}}^2$ of this mean is

$$\sigma_{\overline{x}}^2 = \left\langle \left(\frac{X}{N_S} \right)^2 \right\rangle - \left\langle \frac{X}{N_S} \right\rangle^2 = \frac{1}{N_S^2} \cdot \left(\left\langle X^2 \right\rangle - \langle X \rangle^2 \right) = \frac{N_S \sigma_x^2}{N_S^2} = \frac{\sigma_x^2}{N_S} \quad , \tag{A.11}$$

and by definition of variance, the standard deviation $\sigma_{\overline{x}}$ of the mean value $\overline{x}$ is just the square root of (A.11)

$$\sigma_{\overline{x}} = \frac{\sigma_x}{\sqrt{N_S}} = \sqrt{\frac{\langle x^2 \rangle - \langle x \rangle^2}{N_S}} \quad . \tag{A.12}$$

We take the quantity (A.12) as our uncertainty $\delta \overline{x} = \sigma_{\overline{x}}$ of the best estimate of $\overline{x}$ and report the results of our statistical measurement of $\overline{x}$ as

$$\overline{x} = \langle x \rangle \pm \delta \overline{x} \quad . \tag{A.13}$$

Although the precise meaning of σ_x, and consequently of $\sigma_{\overline{x}}$, is ill-defined unless the distribution $f(x)$ has the form of a normal (or Gaussian) distribution, we argue that as long as the first (μ_1) and second (μ_2) moments of $f(x)$ are convergent, Eq. (A.12) is nevertheless a way to quantify the deviation of $\langle x \rangle$ from the true theoretical mean $\overline{x}$ and access if those two moments are indeed converging by ensuring that the value of $\sigma_{\overline{x}}$ is not too large.

A.2 Linear least square fitting

Suppose we have obtained a set of N_P data points $\{(x_1, y_1), (x_2, y_2), ..., (x_{N_P}, y_{N_P})\}$, where the x_i values are either known exactly or with negligible uncertainty while the y_i values each have an associated uncertainty σ_i. Suppose further we also suspect that this data set should follow a linear relation $f(x_i)$ of the form

$$y_i \simeq f(x_i) = a_0 + a_1 x_i \quad , \tag{A.14}$$

where a_0 and a_1 are the fitting parameters whose values should give the straight line $f(x_i)$ that best fits the data. The question is how can we find the optimal values for these parameters a_0 and a_1?

A possible way to do this goes by the name of least square fitting, in which we minimize a quantity χ^2 (called 'chi-squared') defined as the sum of the squares of the weighted difference between the obtained values y_i and the expected values of the fitting function $f(x_i)$,[2]

$$\chi^2 = \sum_{i=1}^{N_P} \left(\frac{y_i - a_0 - a_1 x_i}{\sigma_i} \right)^2 \quad . \tag{A.15}$$

The squaring of this difference serves to avoid the terms cancelling out due to the possibility of them being either positive or negative, much like how it was for the standard deviation (A.6), while the weighting by the respective uncertainty σ_i favours a fitting which is more tightly bound to y_i values obtained with greater precision rather than with less.

To minimize the χ^2 we differentiate Eq. (A.15) with respect to the parameters a_0 and a_1 and set the results

[2] The '2' in χ^2 is simply a superscript notation and not the square of some quantity χ. This notation is standard.

to zero,

$$\frac{\partial \chi^2}{\partial a_0} = 2\sum_{i=1}^{N_P} \frac{a_0 + a_1 x_i - y_i}{\sigma_i^2} \overset{\text{min}}{=} 0 \quad , \tag{A.16}$$

$$\frac{\partial \chi^2}{\partial a_1} = 2\sum_{i=1}^{N_P} x_i \frac{a_0 + a_1 x_i - y_i}{\sigma_i^2} \overset{\text{min}}{=} 0 \quad , \tag{A.17}$$

which can then be rearranged into a set of linear equations,

$$a_0 \sum_{i=1}^{N_P} \frac{1}{\sigma_i^2} + a_1 \sum_{i=1}^{N_P} \frac{x_i}{\sigma_i^2} = \sum_{i=1}^{N_P} \frac{y_i}{\sigma_i^2} \quad , \tag{A.18}$$

$$a_0 \sum_{i=1}^{N_P} \frac{x_i}{\sigma_i^2} + a_1 \sum_{i=1}^{N_P} \frac{x_i^2}{\sigma_i^2} = \sum_{i=1}^{N_P} \frac{x_i y_i}{\sigma_i^2} \quad . \tag{A.19}$$

To simplify further computations we can introduce the following matrix notation,

$$\boldsymbol{U} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} v_0 \\ v_1 \end{bmatrix} \quad \text{,where } \boldsymbol{U} = \begin{bmatrix} U_{00} & U_{01} \\ U_{10} & U_{11} \end{bmatrix} \text{ is a symmetric matrix,} \tag{A.20}$$

and

$$\begin{cases} U_{00} = \sum_{i=1}^{N_P} \frac{1}{\sigma_i^2} \\ U_{01} = U_{10} = \sum_{i=1}^{N_P} \frac{x_i}{\sigma_i^2} \\ U_{11} = \sum_{i=1}^{N_P} \frac{x_i^2}{\sigma_i^2} \\ v_0 = \sum_{i=1}^{N_P} \frac{y_i}{\sigma_i^2} \\ v_1 = \sum_{i=1}^{N_P} \frac{x_i y_i}{\sigma_i^2} \end{cases} . \tag{A.21}$$

Solving Eq. (A.20) for a_0 and a_1 yields the solution for the fitting parameters as a function of the data,

$$a_0 = \frac{U_{11} v_0 - U_{01} v_1}{\Delta} = \frac{\sum_{i=1}^{N_P} \left(\frac{x_i^2}{\sigma_i^2}\right) \sum_{i=1}^{N_P} \left(\frac{y_i}{\sigma_i^2}\right) - \sum_{i=1}^{N_P} \left(\frac{x_i}{\sigma_i^2}\right) \sum_{i=1}^{N_P} \left(\frac{x_i y_i}{\sigma_i^2}\right)}{\Delta} \quad , \tag{A.22}$$

$$a_1 = \frac{-U_{01} v_0 + U_{00} v_1}{\Delta} = \frac{-\sum_{i=1}^{N_P} \left(\frac{x_i}{\sigma_i^2}\right) \sum_{i=1}^{N_P} \left(\frac{y_i}{\sigma_i^2}\right) + \sum_{i=1}^{N_P} \left(\frac{1}{\sigma_i^2}\right) \sum_{i=1}^{N_P} \left(\frac{x_i y_i}{\sigma_i^2}\right)}{\Delta} \quad , \tag{A.23}$$

where

$$\Delta = U_{00} U_{11} - U_{01}^2 = \sum_{i=1}^{N_P} \left(\frac{1}{\sigma_i^2}\right) \sum_{i=1}^{N_P} \left(\frac{x_i^2}{\sigma_i^2}\right) - \left(\sum_{i=1}^{N_P} \frac{x_i}{\sigma_i^2}\right)^2 \quad . \tag{A.24}$$

To estimate the uncertainties σ_{a_0} and σ_{a_1} of a_0 and a_1, respectively, we propagate the uncertainties σ_i from the corresponding values y_i,

$$\sigma_{a_0} = \sqrt{\sum_{i=1}^{N_P} \left(\sigma_i \frac{\partial a_0}{\partial y_i}\right)^2} = \sqrt{\frac{1}{\Delta^2} \sum_{i=1}^{N_P} \left(\sigma_i \cdot \left[U_{11} \frac{\partial v_0}{\partial y_i} - U_{01} \frac{\partial v_1}{\partial y_i}\right]\right)^2} \quad , \tag{A.25}$$

$$\sigma_{a_1} = \sqrt{\sum_{i=1}^{N_P} \left(\sigma_i \frac{\partial a_1}{\partial y_i}\right)^2} = \sqrt{\frac{1}{\Delta^2} \sum_{i=1}^{N_P} \left(\sigma_i \cdot \left[-U_{01} \frac{\partial v_0}{\partial y_i} + U_{00} \frac{\partial v_1}{\partial y_i}\right]\right)^2} \quad . \tag{A.26}$$

Knowing the results of the following derivatives,

$$\frac{\partial v_0}{\partial y_i} = \sum_{j=1}^{N_P} \frac{1}{\sigma_j^2} \frac{\partial y_j}{\partial y_i} = \sum_{j=1}^{N_P} \frac{\delta_{ji}}{\sigma_j^2} = \frac{1}{\sigma_i^2} \quad , \tag{A.27}$$

$$\frac{\partial v_1}{\partial y_i} = \sum_{j=1}^{N_P} \frac{x_j}{\sigma_j^2} \frac{\partial y_j}{\partial y_i} = \sum_{j=1}^{N_P} \frac{x_j}{\sigma_j^2} \delta_{ji} = \frac{x_i}{\sigma_i^2} \quad , \tag{A.28}$$

where we used $\partial y_j / \partial y_i = \delta_{ji}$, Eqs. (A.25) and (A.26) are then given by

$$\sigma_{a_0} = \sqrt{\frac{1}{\Delta^2} \sum_{i=1}^{N_P} \left(U_{11} \frac{1}{\sigma_i} - U_{01} \frac{x_i}{\sigma_i}\right)^2} = \sqrt{\frac{1}{\Delta^2} \cdot \left(U_{11}^2 \sum_{i=1}^{N_P} \frac{1}{\sigma_i^2} + U_{01}^2 \sum_{i=1}^{N_P} \frac{x_i^2}{\sigma_i^2} - 2U_{11}U_{01} \sum_{i=1}^{N_P} \frac{x_i}{\sigma_i^2}\right)} \quad , \tag{A.29}$$

$$\sigma_{a_1} = \sqrt{\frac{1}{\Delta^2} \sum_{i=1}^{N_P} \left(-U_{01} \frac{1}{\sigma_i} + U_{00} \frac{x_i}{\sigma_i}\right)^2} = \sqrt{\frac{1}{\Delta^2} \cdot \left(U_{01}^2 \sum_{i=1}^{N_P} \frac{1}{\sigma_i^2} + U_{00}^2 \sum_{i=1}^{N_P} \frac{x_i^2}{\sigma_i^2} - 2U_{01}U_{00} \sum_{i=1}^{N_P} \frac{x_i}{\sigma_i^2}\right)} \quad . \tag{A.30}$$

Recalling the definitions (A.21) we can replace the sums in Eqs. (A.29) and (A.30) by the respective element of the matrix $\boldsymbol{U}$. We see that, using the definition (A.24), the term in parenthesis in the rightmost side of Eq. (A.29) gives $U_{11}\Delta$, while the one in Eq. (A.30) gives $U_{00}\Delta$. Thus we have the uncertainty of the fitting parameters also as a function of the data,

$$\sigma_{a_0} = \sqrt{\frac{U_{11}}{\Delta}} = \sqrt{\frac{\sum_{i=1}^{N_P} \frac{x_i^2}{\sigma_i^2}}{\Delta}} \quad , \tag{A.31}$$

$$\sigma_{a_1} = \sqrt{\frac{U_{00}}{\Delta}} = \sqrt{\frac{\sum_{i=1}^{N_P} \frac{1}{\sigma_i^2}}{\Delta}} \quad . \tag{A.32}$$

Appendix B

The normal limit of a Binomial distribution

Consider the probability $P(r,t)$ that a random walker in a one-dimensional lattice, which hops to the right or left with probability p and q, respectively, made r hops to the right (ending up at site n) out of a total of t hops,

$$P(r,t) = \begin{pmatrix} t \\ r \end{pmatrix} p^r q^{t-r} = \frac{t!}{r! \cdot (t-r)!} p^r q^{t-r} \quad . \tag{B.1}$$

The number of hops to the right r, the number of hops to the left l, the total hops t and the site index n are all related through

$$\begin{cases} r = \frac{t+n}{2} \\ l = \frac{t-n}{2} \end{cases} , \tag{B.2}$$

which allow a redefinition of Eq. (B.1) in terms of t and n,

$$P(r,t) = P_n(t) = \frac{t!}{\left(\frac{t+n}{2}\right)! \cdot \left(\frac{t-n}{2}\right)!} p^{\frac{t+n}{2}} q^{\frac{t-n}{2}} \quad . \tag{B.3}$$

Note that $P_n(t) \equiv 0$ if n and t have different parities, since it is impossible to reach a even numbered site after an odd number of hops and vice-versa. Furthermore, $P_n(t) \equiv 0$ if $n > t$ since it is also not possible to reach a site that is farther than the total number of hops performed. This means that the domain where $P_n(t)$ is not strictly zero is $n \in [-t, -t+2, -t+4, ..., t-4, t-2, t]$.

The binomial distribution (B.1) can be approximated to a normal distribution in the limit of large $t \gg n$ using Stirling's formula,

$$x! \simeq x^x e^{-x} \sqrt{2\pi x} \quad \text{, for large } x, \tag{B.4}$$

on the factorials of Eq. (B.1),

$$\begin{cases} t! \simeq t^t e^{-t}\sqrt{2\pi t} \\ r! \simeq r^r e^{-r}\sqrt{2\pi r} \\ (t-r)! \simeq (t-r)^{t-r} e^{-(t-r)}\sqrt{2\pi\cdot(t-r)} \quad , \end{cases} \tag{B.5}$$

so the equation becomes

$$\begin{aligned} P(r,t) &\simeq \frac{t^t e^{-t}\sqrt{2\pi t}}{r^r e^{-r}\sqrt{2\pi r}\cdot(t-r)^{t-r}e^{-(t-r)}\sqrt{2\pi\cdot(t-r)}} p^r q^{t-r} = \\ &= \left(\frac{p}{r}\right)^r \cdot \left(\frac{q}{t-r}\right)^{t-r} \frac{e^{-t}}{e^{-r}e^{-t+r}} t^t \sqrt{\frac{2\pi t}{4\pi^2 r\cdot(t-r)}} = \\ &= \left(\frac{p}{r}\right)^r t^r \cdot \left(\frac{q}{t-r}\right)^{t-r} t^{t-r} \sqrt{\frac{t}{2\pi r\cdot(t-r)}} = \\ &= \left(\frac{tp}{r}\right)^r \cdot \left(\frac{tq}{t-r}\right)^{t-r} \sqrt{\frac{t}{2\pi r\cdot(t-r)}} \quad . \end{aligned} \tag{B.6}$$

Before proceeding we notice that

$$\ln\left[\left(\frac{tp}{r}\right)^r \cdot \left(\frac{tq}{t-r}\right)^{t-r}\right] = r\ln\left[\frac{tp}{r}\right] + (t-r)\ln\left[\frac{tq}{t-r}\right] \quad , \tag{B.7}$$

and defining $\delta \equiv r - tp$, so that

$$\begin{cases} r = \delta + tp \\ tp - r = -\delta \quad \Leftrightarrow \quad t\cdot(1-q) - r = -\delta \quad \Leftrightarrow \quad t - r = tq - \delta \quad , \end{cases} \tag{B.8}$$

we get the following equivalences

$$\begin{cases} \ln\left[\frac{tp}{r}\right] = \ln\left[\frac{tp}{\delta+tp}\right] = -\ln\left[\frac{\delta+tp}{tp}\right] = -\ln\left[1+\frac{\delta}{tp}\right] \\ \ln\left[\frac{tq}{t-r}\right] = \ln\left[\frac{tq}{tq-\delta}\right] = -\ln\left[\frac{tq-\delta}{tq}\right] = -\ln\left[1+\left(-\frac{\delta}{tq}\right)\right] \quad , \end{cases} \tag{B.9}$$

which can be approximated using the expansion of the natural logarithm,

$$\ln[1+x] \simeq x - \frac{1}{2}x^2 \quad \text{, for small } x \in \,]-1,1], \tag{B.10}$$

so equation (B.7) becomes

$$\ln\left[\left(\frac{tp}{r}\right)^{r}\cdot\left(\frac{tq}{t-r}\right)^{t-r}\right]\simeq-(\delta+tp)\cdot\left[\frac{\delta}{tp}-\frac{1}{2}\cdot\left(\frac{\delta}{tp}\right)^{2}\right]-(tq-\delta)\cdot\left[-\frac{\delta}{tq}-\frac{1}{2}\cdot\left(-\frac{\delta}{tq}\right)^{2}\right] \quad . \tag{B.11}$$

Manipulating and simplifying (B.11), ignoring terms of the order of $1/N^2$, we arrive at

$$\ln\left[\left(\frac{tp}{r}\right)^{r}\cdot\left(\frac{tq}{t-r}\right)^{t-r}\right]\simeq-\frac{\delta^2}{2tpq} \quad . \tag{B.12}$$

Exponentiating the previous result we get

$$\left(\frac{tp}{r}\right)^{r}\cdot\left(\frac{tq}{t-r}\right)^{t-r}\simeq e^{-\frac{\delta^2}{2tpq}}=e^{-\frac{(r-tp)^2}{2tpq}} \quad . \tag{B.13}$$

On the other hand, the square root factor in (B.6) can also be approximated, making use of the definition of δ,

$$\begin{aligned}\sqrt{\frac{t}{2\pi r\cdot(t-r)}}=\sqrt{\frac{t}{2\pi\cdot(\delta+tp)\cdot(tq-\delta)}}=\frac{1}{\sqrt{2\pi tpq+2\pi\delta\cdot\left(q-p-\frac{\delta}{t}\right)}}=\\ =\frac{1}{\sqrt{2\pi tpq}}\frac{1}{\sqrt{1+\frac{1}{pq}\frac{\delta}{t}\cdot\left(q-p-\frac{\delta}{t}\right)}} \quad ,\end{aligned} \tag{B.14}$$

where, from the binomial form $1/\sqrt{1+x}=\sum_{k=0}^{\infty}\left(-1/4\right)^{k}\cdot\binom{2k}{k}x^{k}$, we can see that the right term of the last equivalence of (B.14) can be approximated to unity (the 1 comes from the first term of the expansion, when $k=0$, and the remaining terms $k>0$ are all proportional to δ/t), so it simplifies to

$$\sqrt{\frac{t}{2\pi r\cdot(t-r)}}\simeq\frac{1}{\sqrt{2\pi tpq}} \quad . \tag{B.15}$$

Finally going back to Eq. (B.6), using the approximated results (B.13) and (B.15), we can express $P\left(r,t\right)$ as a normal distribution with standard deviation tpq and mean tp,

$$P\left(r,t\right)\simeq\frac{1}{\sqrt{2\pi tpq}}e^{-\frac{(r-tp)^2}{2tpq}} \quad \text{, for large } t \quad . \tag{B.16}$$

Making the change of variable of r to n using relation (B.2) and recalling that $p+q=1$, we can show

that,

$$(r - tp)^2 = {}^1\!/_4 \cdot (n - t \cdot (p - q))^2 \quad , \tag{B.17}$$

and we can express the normal distribution (B.16) as a function of n instead of r, whose mean has now changed into $t \cdot (p - q)$,

$$P(r,t) = P_n(t) \simeq \frac{1}{\sqrt{2\pi tpq}} e^{-\frac{(n - t\cdot(p-q))^2}{8tpq}} \quad \text{, for large } t \gg n \quad . \tag{B.18}$$

As before, the approximation (B.18) is strictly zero if n and t have opposite parities, for the same reasons as Eq. (B.3).

To get from Eq. (B.18) to the continuous space limit, one can introduce the continuous space variable $x = n\Delta\theta$, where $\Delta\theta$ is the lattice spacing. x always varies by factors of $2\Delta\theta$, since n is always either odd or even. For large enough values of t, the probability $P_n(t)$ varies only slightly, and in this regime, a small space interval dx contains ${}^{dx}\!/_{2\Delta\theta}$ possible values of n. Therefore, the probability of finding the random walk at the interval $[x, x + dx]$ at some time t, assuming $P_n(t)$ is constant in this interval, is

$$P(x,t)\,dx = P_n(t)\,\frac{dx}{2\Delta\theta} \quad . \tag{B.19}$$

Appendix C

Solution of the 1D diffusion equation

To solve the diffusion equation

$$\frac{\partial P(\theta,t)}{\partial t} = D\frac{\partial^2 P(\theta,t)}{\partial \theta^2} \quad , \tag{C.1}$$

subject to the IC $\theta(t=0)=a$,

$$P(\theta, t=0) = \delta(\theta - a) \quad , \tag{C.2}$$

and the following boundary conditions (BCs)

$$\left.\frac{\partial P(\theta,t)}{\partial \theta}\right|_{\theta=0} = P(\theta=\phi,t) = 0 \quad , \tag{C.3}$$

corresponding to a reflective boundary at $\theta = 0$ and an absorbing one at $\theta = \phi$, we start by using the method of separation of variables $P(\theta,t) = \Theta(\theta)\,T(t)$,

$$\Theta(\theta)\frac{\partial T(t)}{\partial t} = DT(t)\frac{\partial^2 \Theta(\theta)}{\partial \theta^2} \quad , \tag{C.4}$$

and we divide both sides of Eq. (C.4) by $\Theta(\theta)\,T(t)$,

$$\frac{1}{DT(t)}\frac{\partial T(t)}{\partial t} = \frac{1}{\Theta(\theta)}\frac{\partial^2 \Theta(\theta)}{\partial \theta^2} \quad . \tag{C.5}$$

The only way for Eq. (C.5) to hold true for all values of t and θ is if both sides of the equation are equal to a constant $-k^2$ (the introduction of the minus and the square, instead of just k, is for convenience

reasons),

$$\begin{cases} \frac{1}{DT(t)}\frac{\partial T(t)}{\partial t} = -k^2 & \Leftrightarrow \quad \frac{\partial T(t)}{\partial t} = -k^2 DT\left(t\right) \\ \frac{1}{\Theta(\theta)}\frac{\partial^2 \Theta(\theta)}{\partial \theta^2} = -k^2 & \Leftrightarrow \quad \frac{\partial^2 \Theta(\theta)}{\partial \theta^2} = -k^2 \Theta\left(\theta\right) \end{cases} \quad . \tag{C.6}$$

The general solutions of the ordinary differential equations (C.6) are

$$\begin{cases} T\left(t\right) = e^{-k^2 Dt} \\ \Theta\left(\theta\right) = A\cos\left(k\theta\right) + B\sin\left(k\theta\right) \end{cases} \quad , \tag{C.7}$$

where A and B are arbitrary constants to be defined using the boundary and initial conditions.

A general solution of $P\left(\theta, t\right)$ is then

$$P\left(\theta, t\right) = T\left(t\right)\Theta\left(\theta\right) = e^{-k^2 Dt} \cdot \left(A\cos\left(k\theta\right) + B\sin\left(k\theta\right)\right) \tag{C.8}$$

From the lower reflecting BC at $\theta = 0$ we have

$$\left.\frac{\partial P\left(\theta, t\right)}{\partial \theta}\right|_{\theta=0} = T\left(t\right) \cdot \left[-Ak\sin\left(k\theta\right) + Bk\cos\left(k\theta\right)\right]_{\theta=0} = T\left(t\right) kB \overset{\text{BC}}{=} 0 \quad , \tag{C.9}$$

and since, in general, $T\left(t\right) \neq 0$ and $k \neq 0$ (otherwise we would have a trivial solution), Eq. (C.9) is only true if $B = 0$. From the upper absorbing BC at $\theta = \phi$ we get

$$P\left(\theta = \phi, t\right) = T\left(t\right) A\cos\left(k\phi\right) \overset{\text{BC}}{=} 0 \quad , \tag{C.10}$$

and since, for the same reason as $T\left(t\right)$ and k, $A \neq 0$ in general, Eq. (C.10) is only true if $\cos\left(k\phi\right) = 0$, which imposes that

$$k\phi = \frac{\pi}{2} + \pi m \quad \text{, with } m \in \mathbb{N}_0, \tag{C.11}$$

which, rearranged permits us to define the constant k

$$k \equiv k_m = \frac{\pi}{\phi} \cdot \left(\frac{1}{2} + m\right) \quad . \tag{C.12}$$

Due to the linear properties of the operators present in the diffusion equation, its solution can, in general, be written as an eigeinfunction expansion of all solutions with different m (each term with its corresponding

constant $A \equiv A_m$)

$$P(\theta, t) = \sum_{m=0}^{\infty} A_m e^{-k_m^2 Dt} \cos(k_m \theta) \quad . \tag{C.13}$$

From the initial condition (C.2) we have

$$\sum_{m=0}^{\infty} A_m \cos(k_m \theta) = \delta(\theta - a) \quad . \tag{C.14}$$

To find the constant A_m we can make use of the orthogonality of the $\cos$ function (see appendix D),

$$\int_0^{\phi} d\theta \cos(k_m \theta) \cos(k_n \theta) = \frac{\phi}{2} \delta_{m,n} \quad \text{, where } m,n \in \mathbb{N}_0, \tag{C.15}$$

and multiply Eq. (C.14) in both sides by $\cos(k_n \theta)$ and integrate it in the whole θ domain,

$$\sum_{m=0}^{\infty} A_m \int_0^{\phi} d\theta \cos(k_m \theta) \cos(k_n \theta) = \int_0^{\phi} d\theta \cos(k_n \theta) \delta(\theta - a) \quad , \tag{C.16}$$

which results in

$$\sum_{m=0}^{\infty} A_m \frac{\phi}{2} \delta_{m,n} = \cos(k_n a) \quad . \tag{C.17}$$

The Kronecker delta function in Eq. (C.17) makes all terms zero except for $m = n$, so

$$A_m = \frac{2}{\phi} \cos(k_m a) \quad . \tag{C.18}$$

The solution of the 1D diffusion equation is then found

$$P(\theta, t) = \frac{2}{\phi} \sum_{m=0}^{\infty} e^{-k_m^2 Dt} \cos(k_m a) \cos(k_m \theta) \quad . \tag{C.19}$$

If the BC (C.3) were to be interchanged,

$$P(\theta = 0, t) = \left. \frac{\partial P(\theta, t)}{\partial \theta} \right|_{\theta = \phi} = 0 \quad , \tag{C.20}$$

i.e. if we had instead an absorbing boundary at $\theta = 0$ and a reflecting one at $\theta = \phi$, the general solution (C.8) would still be valid but the constants A and B would have to be recalculated. At $\theta = 0$ we would have

$$P(\theta = 0, t) = T(t) A \overset{\text{BC}}{=} 0 \quad , \tag{C.21}$$

and thus $A = 0$ (since, again, $T(t) \neq 0$ in general, otherwise the solution would be a trivial one). At

$\theta = \phi$ we would get

$$\left.\frac{\partial P(\theta,t)}{\partial \theta}\right|_{\theta=\phi} = T(t)\cdot\left[Bk\cos(k\theta)\right]_{\theta=\phi} = T(t)\,kB\cos(k\phi) \overset{\text{BC}}{=} 0 \quad , \tag{C.22}$$

and since $k \neq 0$ and $B \neq 0$ in general for the usual reason, Eq. (C.22) is only true if $\cos(k\phi) = 0$, therefore leading to the same imposition on the value of k as before,

$$k \equiv k_m = \frac{\pi}{\phi}\cdot\left(\frac{1}{2}+m\right) \quad \text{, with } m \in \mathbb{N}_0. \tag{C.23}$$

The eigeinfunction expansion solution would then be (each term with its corresponding constant $B \equiv B_m$)

$$P(\theta,t) = \sum_{m=0}^{\infty} B_m e^{-k_m^2 Dt}\sin(k_m\theta) \quad \text{, for swapped BCs.} \tag{C.24}$$

To determine the constants B_m we can again make use of the orthogonality of the trigonometric functions, this time using the $\sin$ function (see appendix D),

$$\int_0^{\phi} d\theta \sin(k_m\theta)\sin(k_n\theta) = \frac{\phi}{2}\delta_{m,n} \quad \text{, where } m,n \in \mathbb{N}_0, \tag{C.25}$$

and multiply both sides of Eq. (C.24) at $t = 0$ by $\sin(k_n\theta)$ and integrating it in θ over the whole domain. For the same initial condition as before, Eq. (C.2), this would give

$$\sum_{m=0}^{\infty} B_m \frac{\phi}{2}\delta_{m,n} = \sin(k_n a) \quad \Leftrightarrow \quad B_m = \frac{2}{\phi}\sin(k_m a) \quad . \tag{C.26}$$

The solution of the diffusion equation for interchanged BCs with the same initial condition would then be given by

$$P(\theta,t) = \frac{2}{\phi}\sum_{m=0}^{\infty} e^{-k_m^2 Dt}\sin(k_m a)\sin(k_m\theta) \quad \text{, for swapped BCs.} \tag{C.27}$$

Appendix D

Calculation of the trigonometric integral

To evaluate the integral

$$I_{m,n} = \int_0^{\lambda} dx \cos\left[\frac{\pi}{\lambda}\cdot\left(\frac{1}{2}+m\right)x\right]\cos\left[\frac{\pi}{\lambda}\cdot\left(\frac{1}{2}+n\right)x\right] \quad , \tag{D.1}$$

where $n, m \in \mathbb{N}_0$ and $\lambda \in \mathbb{R}$, we start by considering the following trigonometric identity

$$\cos(a)\cos(b) = \frac{1}{2}\cdot[\cos(a-b)+\cos(a+b)] \quad \text{, where } a, b \in \mathbb{R}, \tag{D.2}$$

and use it in the $\cos$ functions of the integral

$$I_{m,n} = \frac{1}{2}\int_0^{\lambda} dx \cdot \left[\cos\left(\frac{\pi}{\lambda}\cdot(m-n)\,x\right)+\cos\left(\frac{\pi}{\lambda}\cdot(1+m+n)\,x\right)\right] \quad . \tag{D.3}$$

Performing the integration, we have

$$I_{m,n} = \frac{\lambda}{2}\cdot\left[\frac{\sin[\pi\cdot(m-n)]}{\pi\cdot(m-n)}+\frac{\sin[\pi\cdot(1+m+n)]}{\pi\cdot(1+m+n)}\right] \quad , \tag{D.4}$$

and knowing $\sin[\pi\cdot(1+m+n)] = -\sin[\pi\cdot(m+n)]$,

$$I_{m,n} = \frac{\lambda}{2}\cdot\left[\frac{\sin[\pi\cdot(m-n)]}{\pi\cdot(m-n)}-\frac{\sin[\pi\cdot(m+n)]}{\pi\cdot(1+m+n)}\right] \quad . \tag{D.5}$$

Notice that for $m \neq n$ both $\sin$ functions give zero, so $I_{m,n} = 0$ for that case. On the other hand, if we have $m = n$, $\sin[\pi\cdot(m+n)]$ is always zero, but the first term is an indetermination, and we can use

L'Hôpital's rule to evaluate the limit $m - n \equiv \Delta \to 0$,

$$\lim_{\Delta\to 0} \frac{\sin(\pi\Delta)}{\pi\Delta} = \lim_{\Delta\to 0} \frac{\frac{d}{d\Delta}[\sin(\pi\Delta)]}{\frac{d}{d\Delta}[\pi\Delta]} = \lim_{\Delta\to 0} \cos(\pi\Delta) = \cos(0) = 1 \quad , \tag{D.6}$$

and so we get $I_{m,n} = \lambda/2$ for $m = n$.

In conclusion, the integral $I_{m,n}$ results in

$$I_{m,n} = \frac{\lambda}{2}\delta_{m,n} \quad . \tag{D.7}$$

If the $\cos$ functions in the integral (D.1) were to be replaced by $\sin$ functions with the same respective arguments the result would be identical. This can be easily verified by considering the following trigonometric identity

$$\sin(a)\sin(b) = \frac{1}{2}[\cos(a-b) - \cos(a+b)] \quad \text{, where } a, b \in \mathbb{R}, \tag{D.8}$$

where the only difference compared to the right hand side of the identity (D.2) is a change of sign in the second trigonometric function. Thus, calculations would go unchanged, leading again to Eq. (D.5) save for a plus sign in the second term, instead of a minus one. Furthermore we have seen that this second term is always zero no matter the values of m and n, and so this sign change would be irrelevant. Therefore we also have that

$$\int_0^\lambda dx \sin\left[\frac{\pi}{\lambda}\cdot\left(\frac{1}{2}+m\right)x\right]\sin\left[\frac{\pi}{\lambda}\cdot\left(\frac{1}{2}+n\right)x\right] = I_{m,n} = \frac{\lambda}{2}\delta_{m,n} \quad . \tag{D.9}$$

Appendix E

Folding time distributions

In this appendix we group in table E.1 the explicit average values of the binding and folding times of the results presented in the figures of section 3.4. In Figs. E.1-E.4 we show the histograms with the obtained distributions for each IC and type of system, from which the aforementioned averages were respectively computed.

Table E.1: Average values obtained for the first binding $\langle T_F \rangle$, last binding $\langle T_L \rangle$ and folding times $\langle T \rangle$ for both pinned and suspended systems, with $N = 2$ (a) and $N = 3$ (b) lateral faces for each of the six ICs defined in section 3.4. The first column specifies the type of system for which the numerical value was obtained, while the second one specifies which time average the numerical value is referring to. The values reported in sub-table (a) and (b) are the same as those plotted in the Figs. 3.17 and 3.18, respectively. All these time averages were computed from the expectation value of the respective distributions of Figs. E.1, E.2, E.3 and E.4, with the exception of the average total folding time $\langle T \rangle$ for $N = 3$ (last row of sub-table b), which was instead calculated by summing the corresponding numerical average values of the first and last binding times. The expected values of the folding time distributions in Figs. E.3 and E.4 (black histograms) are, as supposed, consistent with this sum. All values are in units of Brownian time.

(a) $N = 2$

		Planar	Random 1	Random 2	Random 3	Stacking	Opposite stacking
Pinned	$\langle T \rangle$	2.290±0.006	2.192±0.006	1.970±0.006	1.859±0.006	1.701±0.006	2.108±0.006
Suspended	$\langle T \rangle$	4.90±0.01	4.81±0.01	4.59±0.01	4.50±0.01	3.30±0.01	5.37±0.01

(b) $N = 3$

		Planar	Random 1	Random 2	Random 3	Stacking	Opposite stacking
Pinned	$\langle T_F \rangle$	1.041±0.003	0.865±0.002	0.744±0.002	0.687±0.002	0.544±0.002	0.740±0.002
	$\langle T_L \rangle$	0.277±0.001	0.281±0.001	0.262±0.001	0.245±0.001	(1905±9)×10^{-4}	0.268±0.001
	$\langle T \rangle$	1.318±0.003	1.146±0.002	1.006±0.002	0.932±0.002	0.735±0.002	1.008±0.002
Suspended	$\langle T_F \rangle$	1.799±0.005	1.630±0.005	1.495±0.005	1.440±0.005	0.769±0.004	1.425±0.004
	$\langle T_L \rangle$	1.824±0.007	1.816±0.007	1.839±0.007	1.827±0.007	0.833±0.005	2.133±0.007
	$\langle T \rangle$	3.623±0.009	3.446±0.009	3.334±0.009	3.267±0.009	1.602±0.006	3.558±0.008

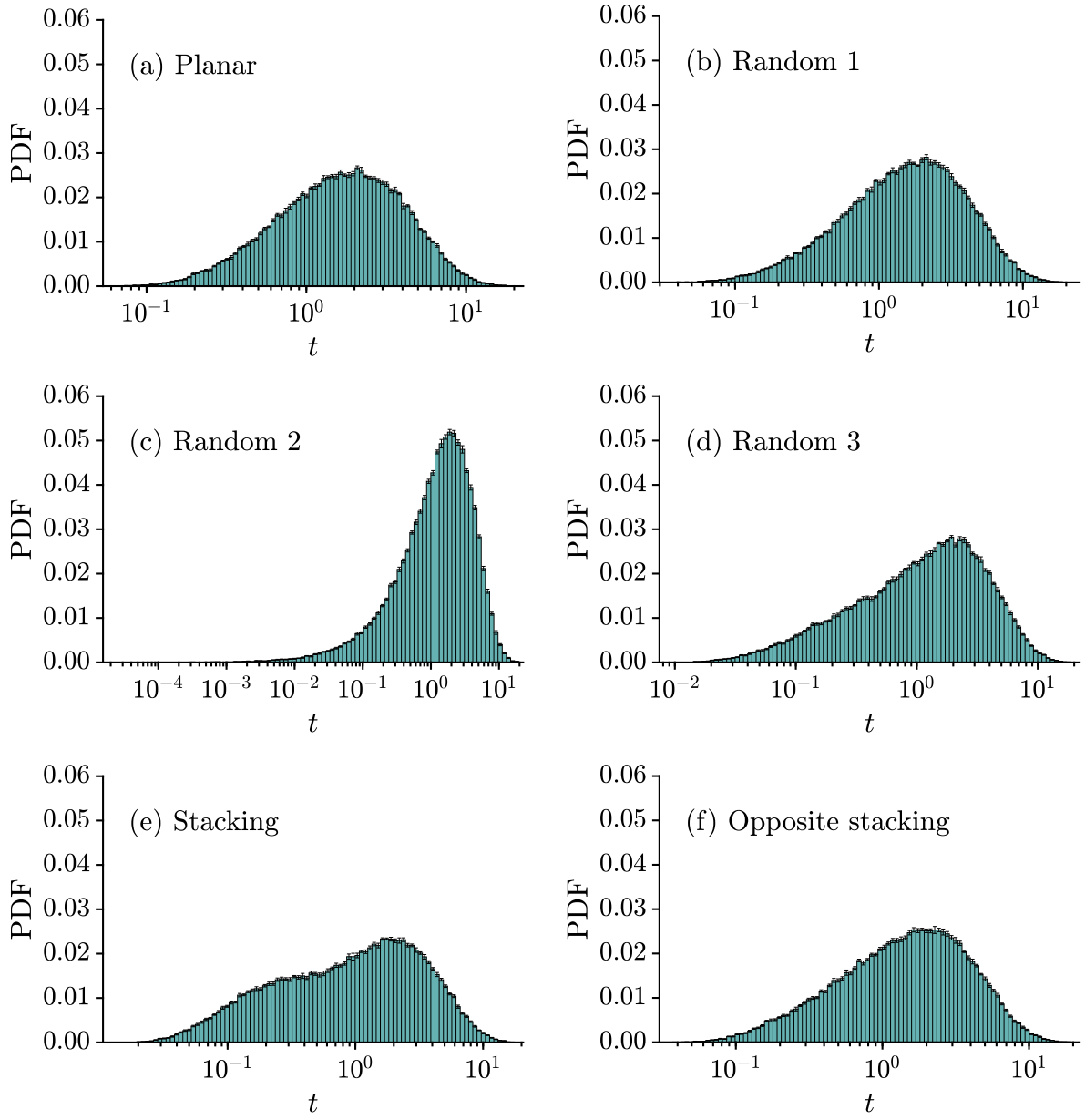

Figure E.1: Numerically obtained distributions of the folding time, in Brownian time units, for the six ICs defined in section 3.4 for $N = 2$ lateral faces, for a pinned system. Each histogram was made from $N_S = 10^5$ independent samples, distributed with logarithmic binning across 100 bins. The simulations were performed under SC for all six ICs. The averages of these distributions are listed in the second row of Table E.1a and plotted in Fig. 3.17a of section 3.4.

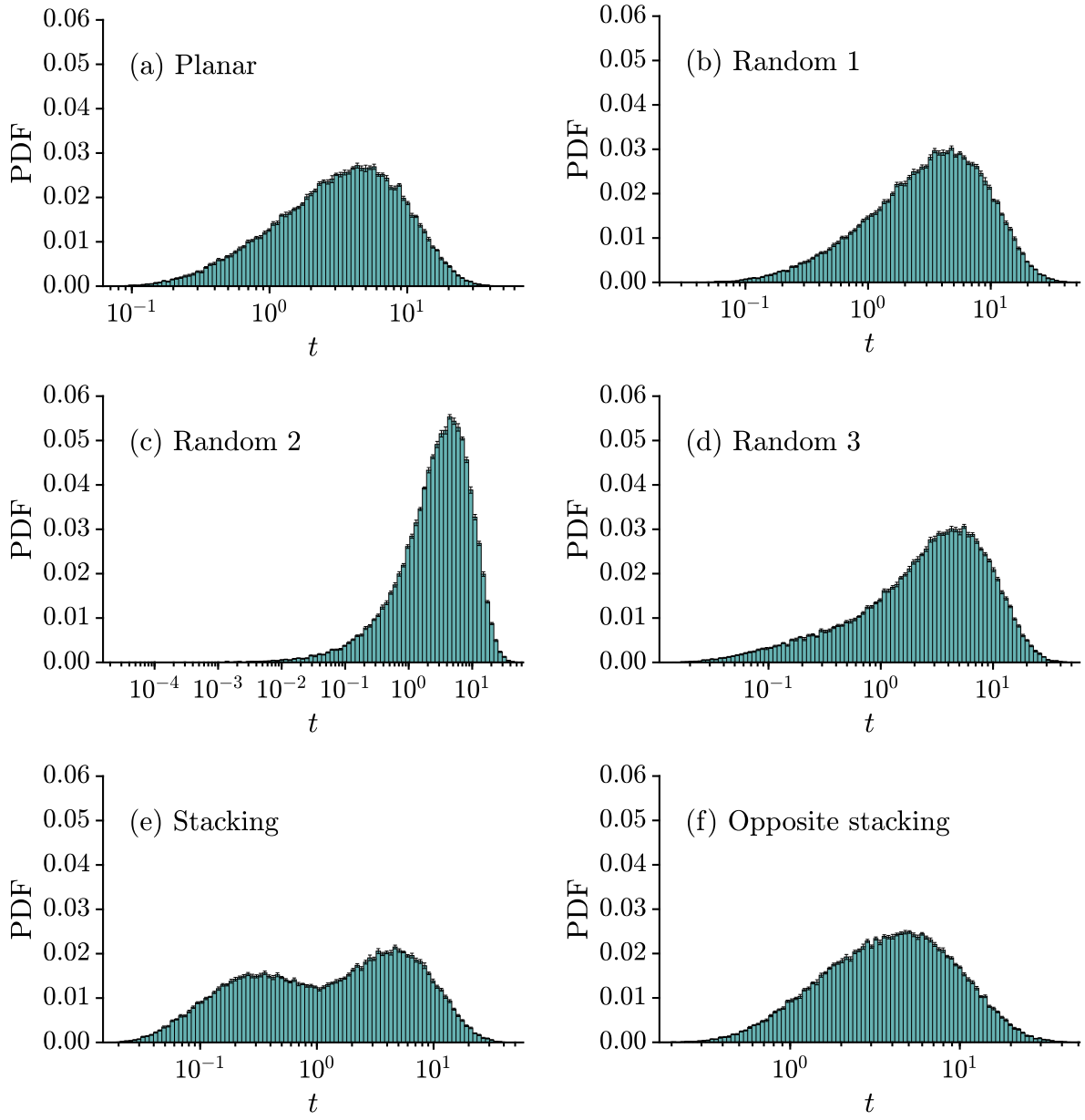

Figure E.2: Numerically obtained distributions of the folding time, in Brownian time units, for the six ICs defined in section 3.4 for $N = 2$ lateral faces, for a suspended system. Each histogram was made from $N_S = 10^5$ independent samples, distributed with logarithmic binning across 100 bins. The simulations were performed under SC for all six ICs. The averages of these distributions are listed in the last row of Table E.1a and plotted in Fig. 3.17b of section 3.4.

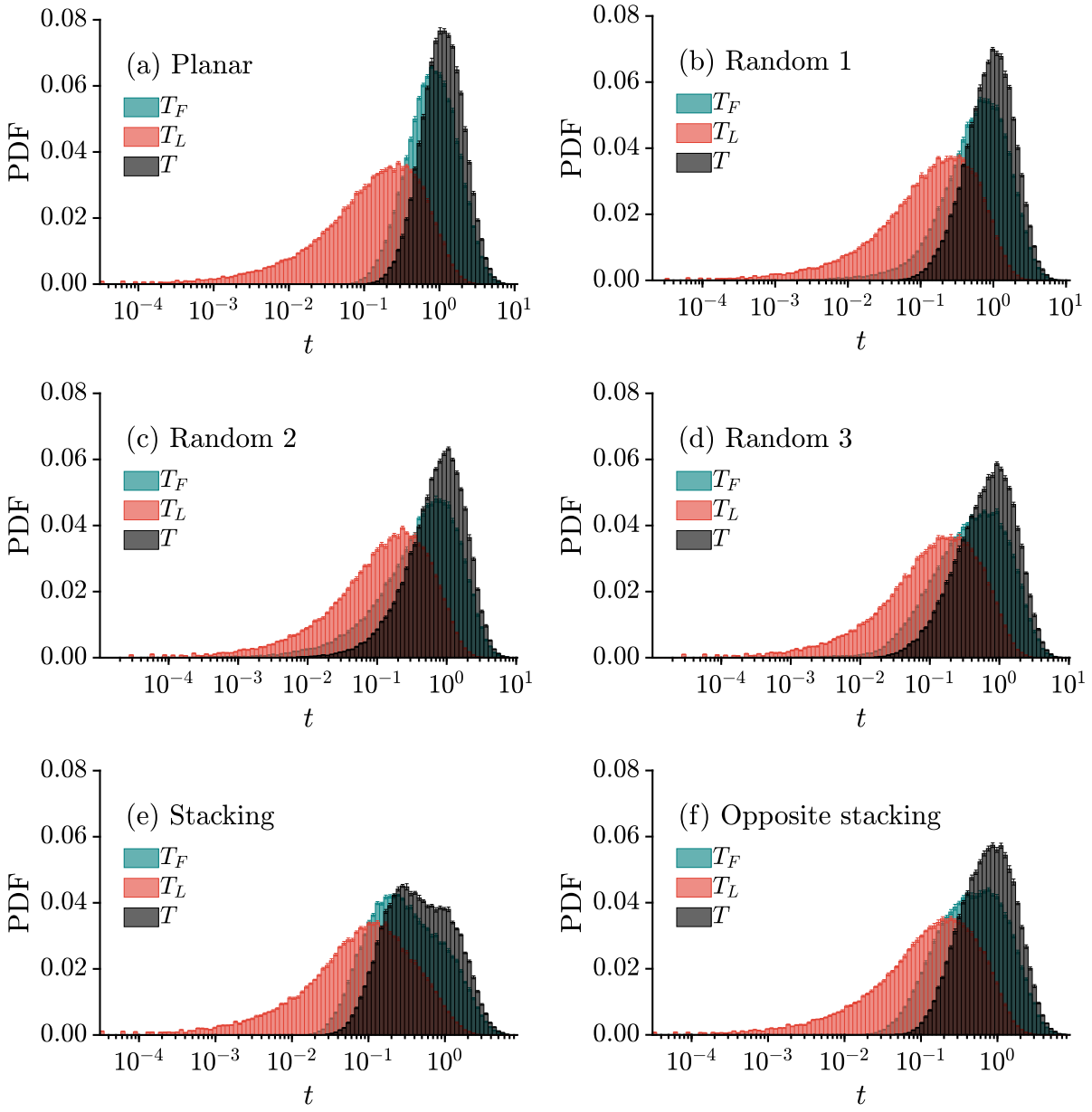

Figure E.3: Numerically obtained distributions of the first binding (T_F), last binding (T_L) and folding time (T), in Brownian time units, for the six ICs defined in section 3.4 for $N = 3$ lateral faces, for a pinned system. Each histogram was made from $N_S = 10^5$ independent samples, distributed with logarithmic binning across 100 bins. The simulations were performed under SC for all six ICs. The averages of these distributions are listed in the second to fourth rows of Table E.1b and plotted in Figs. 3.18a, 3.18c and 3.18e (left bar plot) of section 3.4.

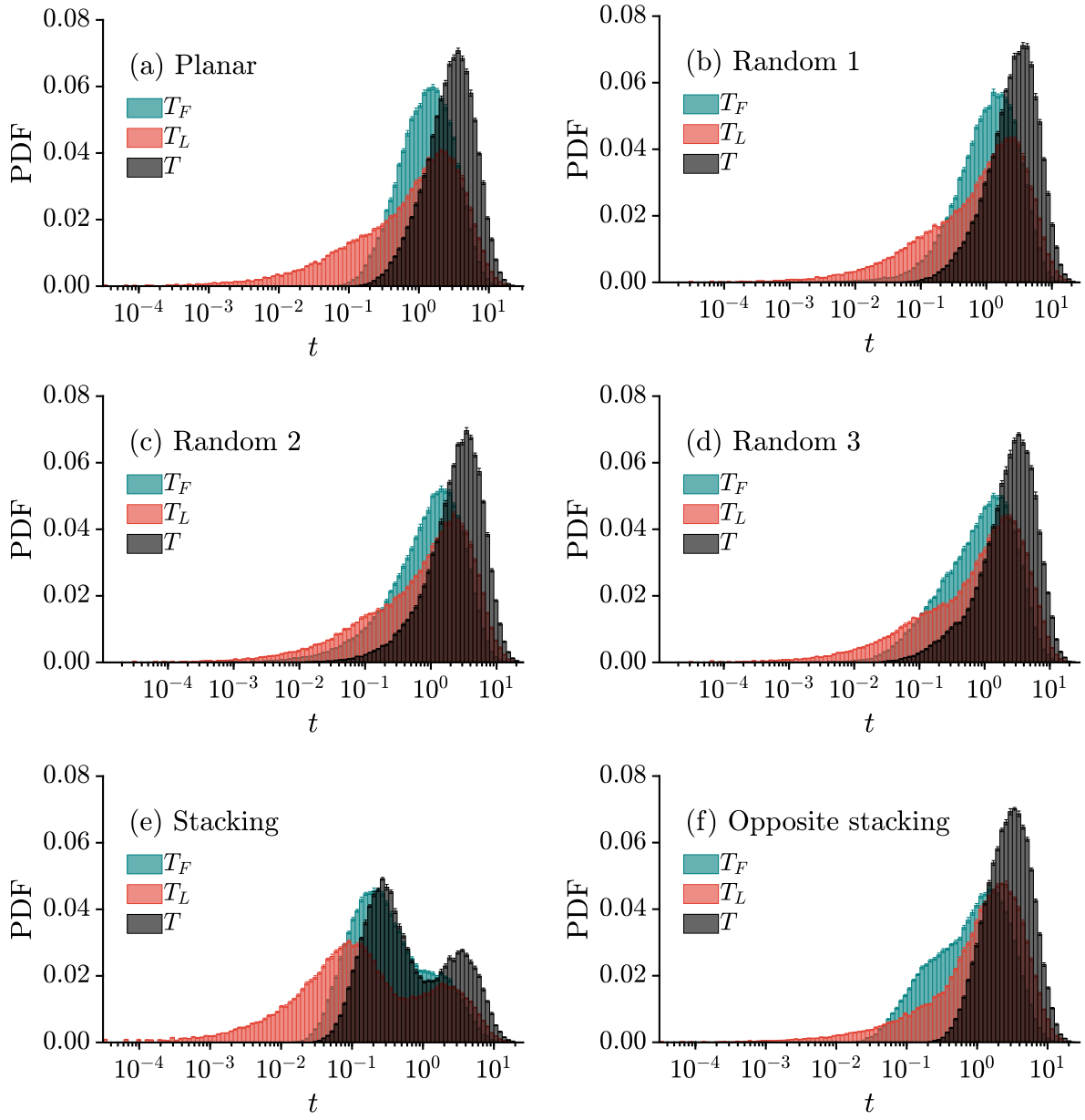

Figure E.4: Numerically obtained distributions of the first binding (T_F), last binding (T_L) and folding time (T), in Brownian time units, for the six ICs defined in section 3.4 for $N = 3$ lateral faces, for a suspended system. Each histogram was made from $N_S = 10^5$ independent samples, distributed with logarithmic binning across 100 bins. The simulations were performed under SC for all six ICs. The averages of these distributions are listed in the fifth to last rows of Table E.1b and plotted in Figs. 3.18b, 3.18d and 3.18e (right bar plot) of section 3.4.

Bibliography

[1] Erik D. Demaine and Joseph O'rourke. Geometric folding algorithms: Linkages, origami, polyhedra, volume 9780521857. 2007.

[2] Graham P. Collins. Kirigami and technology cut a fine figure, together. Proceedings of the National Academy of Sciences of the United States of America, 113(2):240–241, 2016.

[3] Albrecht Dürer. The Painter's Manual. Abaris, New York, 1st editio edition, 1977.

[4] Xin Ning, Xueju Wang, Yi Zhang, Xinge Yu, Dongwhi Choi, Ning Zheng, Dong Sung Kim, Yonggang Huang, Yihui Zhang, and John A. Rogers. Assembly of Advanced Materials into 3D Functional Structures by Methods Inspired by Origami and Kirigami: A Review. Advanced Materials Interfaces, 5(13):1–13, 2018.

[5] Yihui Zhang, Fan Zhang, Zheng Yan, Qiang Ma, Xiuling Li, Yonggang Huang, and John A. Rogers. Printing, folding and assembly methods for forming 3D mesostructures in advanced materials. Nature Reviews Materials, 2(4), 2017.

[6] Shanshan Chen, Jianfeng Chen, Xiangdong Zhang, Zhi Yuan Li, and Jiafang Li. Kirigami/origami: unfolding the new regime of advanced 3D microfabrication/nanofabrication with 'folding'. Light: Science and Applications, 9(1), 2020.

[7] Mohammad Humood, Yan Shi, Mengdi Han, Joseph Lefebvre, Zheng Yan, Matt Pharr, Yihui Zhang, Yonggang Huang, John A. Rogers, and Andreas A. Polycarpou. Fabrication and Deformation of 3D Multilayered Kirigami Microstructures. Small, 14(11):1–9, 2018.

[8] William M. Jacobs and Daan Frenkel. Self-Assembly of Structures with Addressable Complexity. Journal of the American Chemical Society, 138(8):2457–2467, 2016.

[9] Terry C. Shyu, Pablo F. Damasceno, Paul M. Dodd, Aaron Lamoureux, Lizhi Xu, Matthew Shlian, Max Shtein, Sharon C. Glotzer, and Nicholas A. Kotov. A kirigami approach to engineering elasticity in nanocomposites through patterned defects. Nature Materials, 14(8):785–789, 2015.

[10] Daniel M. Sussman, Yigil Cho, Toen Castle, Xingting Gong, Euiyeon Jung, Shu Yang, and Randall D. Kamien. Algorithmic lattice kirigami: A route to pluripotent materials. Proceedings of the National Academy of Sciences of the United States of America, 112(24):7449–7453, 2015.

[11] Sheng Xu, Zheng Yan, Kyung In Jang, Wen Huang, Haoran Fu, Jeonghyun Kim, Zijun Wei, Matthew Flavin, Joselle McCracken, Renhan Wang, Adina Badea, Yuhao Liu, Dongqing Xiao, Guoyan Zhou, Jungwoo Lee, Ha Uk Chung, Huanyu Cheng, Wen Ren, Anthony Banks, Xiuling Li, Ungyu Paik, Ralph G. Nuzzo, Yonggang Huang, Yihui Zhang, and John A. Rogers. Assembly of micro/nanomaterials into complex, three-dimensional architectures by compressive buckling. Science, 347(6218):154–159, 2015.

[12] Yihui Zhang, Zheng Yan, Kewang Nan, Dongqing Xiao, Yuhao Liu, Haiwen Luan, Haoran Fu, Xizhu Wang, Qinglin Yang, Jiechen Wang, Wen Ren, Hongzhi Si, Fei Liu, Lihen Yang, Hejun Li, Juntong Wang, Xuelin Guo, Hongying Luo, Liang Wang, Yonggang Huang, and John A. Rogers. A mechanically driven form of Kirigami as a route to 3D mesostructures in micro/nanomembranes. Proceedings of the National Academy of Sciences of the United States of America, 112(38):11757–11764, 2015.

[13] H. P. M. Melo, C. S. Dias, and N. A. M. Araujo. Optimal Number of Faces for Fast Self-Folding Kirigami. Communications Physics, (2020):1–5, 2020.

[14] Aaron Lamoureux, Kyusang Lee, Matthew Shlian, Stephen R. Forrest, and Max Shtein. Dynamic kirigami structures for integrated solar tracking. Nature Communications, 6:1–6, 2015.

[15] Jian Sun, Fabrizio Scarpa, Yanju Liu, and Jinsong Leng. Morphing thickness in airfoils using pneumatic flexible tubes and Kirigami honeycomb. Journal of Intelligent Material Systems and Structures, 27(6):755–763, 2016.

[16] Sina Sareh and Jonathan Rossiter. Kirigami artificial muscles with complex biologically inspired morphologies. Smart Materials and Structures, 22(1), 2013.

[17] Zeming Song, Xu Wang, Cheng Lv, Yonghao An, Mengbing Liang, Teng Ma, David He, Ying Jie Zheng, Shi Qing Huang, Hongyu Yu, and Hanqing Jiang. Kirigami-based stretchable lithium-ion batteries. Scientific Reports, 5:1–9, 2015.

[18] Shivendra Pandey, Margaret Ewing, Andrew Kunas, Nghi Nguyen, David H. Gracias, and Govind Menon. Algorithmic design of self-folding polyhedra. Proceedings of the National Academy of Sciences of the United States of America, 108(50):19885–19890, 2011.

[19] N. A.M. Araújo, R. A. Da Costa, S. N. Dorogovtsev, and J. F.F. Mendes. Finding the Optimal Nets for Self-Folding Kirigami. Physical Review Letters, 120(18):188001, 2018.

[20] Rohan Fernandes and David H. Gracias. Self-folding polymeric containers for encapsulation and delivery of drugs. Advanced Drug Delivery Reviews, 64(14):1579–1589, 2012.

[21] Jongmin Shim, Claude Perdigou, Elizabeth R. Chen, Katia Bertoldi, and Pedro M. Reis. Buckling-induced encapsulation of structured elastic shells under pressure. Proceedings of the National Academy of Sciences of the United States of America, 109(16):5978–5983, 2012.

[22] M. Filippousi, T. Altantzis, G. Stefanou, M. Betsiou, D. N. Bikiaris, M. Angelakeris, E. Pavlidou, D. Zamboulis, and G. Van Tendeloo. Polyhedral iron oxide core-shell nanoparticles in a biodegradable polymeric matrix: Preparation, characterization and application in magnetic particle hyperthermia and drug delivery. RSC Advances, 3(46):24367–24377, 2013.

[23] S. Felton, M. Tolley, E. Demaine, D. Rus, and R. Wood. A method for building self-folding machines. Science, 345(6197):644–646, 2014.

[24] John Rogers, Yonggang Huang, Oliver G. Schmidt, and David H. Gracias. Origami MEMS and NEMS. MRS Bulletin, 41(2):123–129, 2016.

[25] Melina K. Blees, Arthur W. Barnard, Peter A. Rose, Samantha P. Roberts, Kathryn L. McGill, Pinshane Y. Huang, Alexander R. Ruyack, Joshua W. Kevek, Bryce Kobrin, David A. Muller, and Paul L. McEuen. Graphene kirigami. Nature, 524(7564):204–207, 2015.

[26] Paul M. Dodd, Pablo F. Damasceno, and Sharon C. Glotzer. Universal folding pathways of polyhedron nets. Proceedings of the National Academy of Sciences of the United States of America, 115(29):E6690–E6696, 2018.

[27] Anum Azam, Timothy G. Leong, Aasiyeh M. Zarafshar, and David H. Gracias. Compactness determines the success of cube and octahedron self-assembly. PLoS ONE, 4(2):2–7, 2009.

[28] Marc Z. Miskin, Kyle J. Dorsey, Baris Bircan, Yimo Han, David A. Muller, Paul L. McEuen, and Itai Cohen. Graphene-based bimorphs for micron-sized, tautonomous origami machines. Proceedings of the National Academy of Sciences of the United States of America, 115(3):466–470, 2018.

[29] Kaori Kuribayashi-Shigetomi, Hiroaki Onoe, and Shoji Takeuchi. Cell Origami: Self-Folding of Three-Dimensional Cell-Laden Microstructures Driven by Cell Traction Force. PLoS ONE, 7(12):1–8, 2012.

[30] Mie Andersen, Chiara Panosetti, and Karsten Reuter. A practical guide to surface kinetic Monte Carlo simulations. Frontiers in Chemistry, 7(APR), 2019.

[31] Pavel L. Krapivsky, Sidney Redner, and Eli Ben-Naim. A Kinetic View of Statistical Physics. Cambridge University Press, Cambridge, 2010.

[32] Dani Darmawan. Radiation Effects in Solids, volume 235 of NATO Science Series. Springer Netherlands, Dordrecht, 2007.

[33] Michail Stamatakis. Kinetic modelling of heterogeneous catalytic systems. Journal of Physics Condensed Matter, 27(1), 2015.

[34] Makoto Matsumoto and Takuji Nishimura. Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator. ACM Transactions on Modeling and Computer Simulation, 8(1):3–30, 1998.

[35] Sidney Redner. A Guide to First-Passage Processes. Number 1. Cambridge University Press, aug 2001.

[36] Abramowitz Milton and Stegun Irene. Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables. United States Department of Commerce; National Bureau of Standards (NBS); Dover Publications, Inc., New York, 10th print edition, 1972.

[37] Eric W. Weisstein. CRC Concise Encyclopedia of Mathematics. Chapman and Hall/CRC, dec 2002.

[38] Barry C. Arnold, N. Balakrishnan, and H. N. Nagaraja. A First Course in Order Statistics. Society for Industrial and Applied Mathematics, jan 2008.

[39] K. Basnayake, Z. Schuss, and D. Holcman. Asymptotic Formulas for Extreme Statistics of Escape Times in 1, 2 and 3-Dimensions. Journal of Nonlinear Science, 29(2):461–499, 2019.

[40] Sergei Nechaev, Kirill Polovnikov, Senya Shlosman, Alexander Valov, and Alexander Vladimirov. Anomalous one-dimensional fluctuations of a simple two-dimensional random walk in a large-deviation regime. Physical Review E, 99(1):1–10, 2019.

[41] Luca Giuggioli. Exact Spatiotemporal Dynamics of Confined Lattice Random Walks in Arbitrary Dimensions: A Century after Smoluchowski and Pólya. Physical Review X, 10(2):021045, may 2020.

[42] Elliott W. Montroll and George H. Weiss. Random walks on lattices. II. Journal of Mathematical Physics, 6(2):167–181, 1965.

[43] John R. Taylor. An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. University Science Books, Sausalito, California, 2nd edition, 1997.

[44] Peter Young. Everything you wanted to know about Data Analysis and Fitting but were afraid to ask. 2012.