A Few Facets of Random Walks & A Hint of Stochastic Modelling

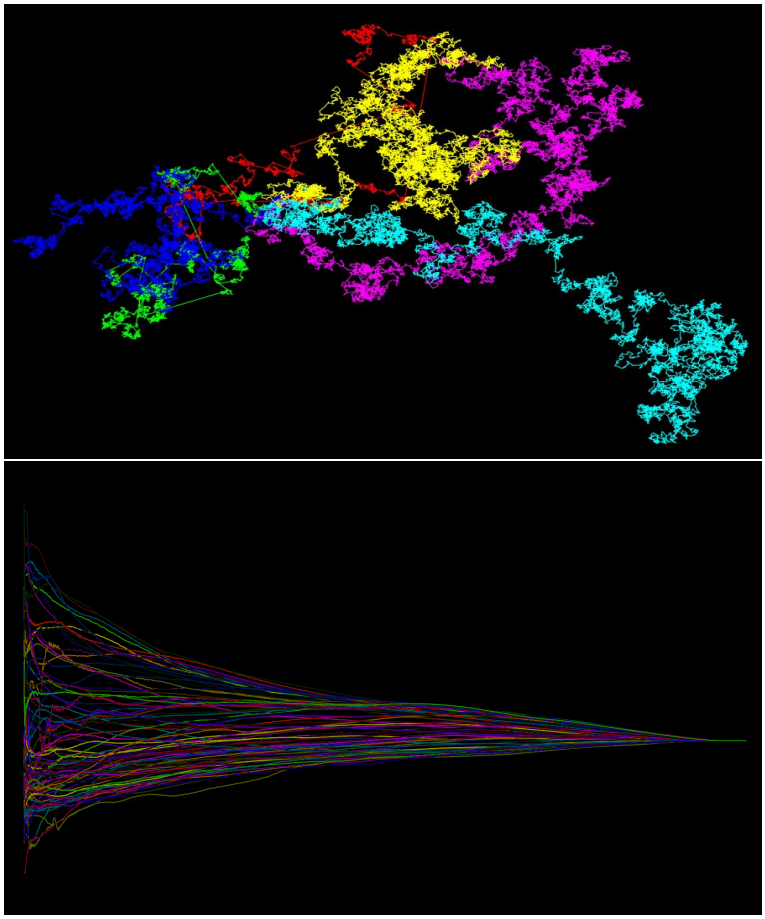
Julien Randon-Furling

A Few Facets of Random Walks

&

A Hint of Stochastic Modelling

Julien Randon-Furling
SAMM, Université Paris 1 Panthéon Sorbonne



Habilitation **– 7 December 2018**

Jury

Eva Löcherbach (SAMM, Univ. Paris-1) – University Referee

Marc Barthélémy (IPhT, CEA Saclay) – External Referee

Pierre Vallois (Institut Elie Cartan, Univ. Lorraine) – External Referee

and

Nilanjana Datta (StatsLab, Univ. Cambridge)

Jean-Pierre Nadal (LPS ENS / CAMS EHESS)

Fabrice Rossi (SAMM, Univ. Paris-1)

Ellen Saada (MAP5, CNRS / Univ. Paris 5)

Gernimo Uribe Bravo (UNAM, Mexico)

CONTENTS

Aux temps d'atteinte...

INTRODUCTION

Since I finished my PhD in 2009, my work has evolved in more than one direction, more or less fruitfully, as is the case for many a young scientist, I suppose. The following report is an attempt to give an overview of what constitutes perhaps the two main stems of my research: some theoretical aspects of certain stochastic processes on the one hand, and studies of certain types of complex systems on the other hand. The two are not unrelated, though. The reader will encounter random walks and a general probability-theoretic flavour in both cases.

As I was reflecting upon what could potentially appear, if not original, at least maybe a little peculiar or specific in the type of work I have been pursuing, some words have sprung back to my mind. Those were words spoken by the late Marc Yor when he invited theoretical physicist Satya N. Majumdar and his then PhD student (the author of these lines – only a bit younger then!) to discuss some properties of Brownian motion. We had computed explicitly certain densities, and Marc was convinced (rightly, of course) that these could be linked to a set of more general theorems sometimes called *agreement formulae*. We spent half a day exchanging ideas in the most lively and agreeable manner. I have to admit I did not understand much of the extremely abstract concepts with which Marc was playing so easily... But I was deeply impressed, and in particular by his profound goodwill at partaking in discussions with theoretical physicists. Far from despising us for not being fluent in *filtrations* or for not using standard probability-theory notations, he tirelessly engaged in exchanging with us. At some point, he said: "I have long thought how great and useful it would be to have a sort of *bilingual* dictionary of Brownian motion, with double entries allowing physicists and mathematicians to share their knowledge more easily". Together with Stavros Vakeroudis, then a PhD student with Marc, we even had a brief discussion on how to start this as an online, collaborative platform. But then of course, our PhD theses took their toll on our spare time...

Yet, the idea of regularly wandering on both sides of the divide between probabilists and statistical physicists has borne with, and even grown on, me. Therefore the reader will find in this report both $\mathbb{E}$'s and $\langle \dots \rangle$'s, expectations and ensemble averages, transition probabilities and propagators... May s/he forgive me for this constant to-and-fro – let it be perceived as a healthy gymnastic for the mind.

Apart from somewhat exotic vocabulary and notations, treading a line between probability theory and statistical physics has led me, I hope, to introduce at times certain types of reasoning that perhaps are not so common in the physics literature on stochastic processes. Path transformations would be an example. The use of reciprocal processes is another one. Conversely, keeping a physicist's eye leads one to ask questions that tend to be less prominent in the probability-theory community. Obtaining explicit formulae for densities and moments, examining specific configurations are key questions for physicists. Also, using mathematical tools accessible to an audience wider than the stochastic processes community may be just as important, depending on the applications. One should bear in mind that most physicists (including

statistical physicists) do not always remember what a σ-algebra is – but all of them decidedly have a taste for mathematical reasoning and know how to check for flaws and fallacies. Hence, part of my work on stochastic processes is guided by a form of "minimalism" as far as mathematical structures are concerned.

Among common grounds for mathematicians and physicists, when it comes to stochastic processes, is a strong interest in what the latter call "universal" properties, and the former "central limit theorems" or "distribution-free" results. For instance, results that are valid for all random walks, independently of the jump distributions, as long as they are symmetric – think of Erik Sparre Andersen's theorems. Some of the results I have been able to establish on Brownian motion in the years since my PhD exhibit some sort of universality, and may be extended to more general Lévy processes. This is particularly the case when the proofs rely on path transformations, as I will try to show in the following pages.

The first stem of my work is partly rooted in the problems that had remained open at the end of my PhD, and partly in new questions that have come up from other sides. Six papers are published in international peer-reviewed journals:

- J. Randon-Furling: *Convex hull of n planar Brownian paths: an exact formula for the average number of edges*, Journal of Physics A: Mathematical & Theoretical 46, 015004 (2012). This paper was distinguished as a "Highlight of 2013" by the journal.

- J. Randon-Furling: *Universality and time-scale invariance for the shape of planar Lévy processes*, Physical Review E 89, 052112 (2014).

- J. Randon-Furling: *From Markovian to non-Markovian persistence exponents*, Europhysics Letters (EPL) 109 40015 (2015). This paper was distinguished as an "Editor's choice" by the journal.

- E. Ben-Naim, P.L. Krapivsky, J. Randon-Furling: *Maxima of two random walks: universal statistics of lead changes*, Journal of Physics A: Mathematical & Theoretical 49, 205003 (2016).

- J. Randon-Furling, F. Wespi: *Facets on the convex hull of d-dimensional Brownian and Lévy motion*, Physical Review E 95, 032129 (2017).

- J. Randon-Furling, S. Redner: *Residence time near an absorbing set*, Journal of Statistical Mechanics 103205 (2018).

Three of these papers are directly related to the study of Brownian or Lévy convex hulls, thus pertaining to a wider field called *stochastic geometry* – I have been a member of the CNRS research group GeoSto (GdR 3477) since 2011.
The other three papers deal with first-passage properties, hitting times, and local times for various types of processes: discrete-time random walks, continuous-time Lévy processes, non-Markovian processes (in particular, supremum processes).
I have been lucky enough to convince a few students of the beauty there is in the properties of stochastic

processes, and my research has benefited greatly from exchanging with them and supervising their work: Apollinaire Barme (M1 student) worked on edge and facet probabilities, Nicolas Jouvin (M1 student) worked on symmetric stable processes, Emma Krebs (M1 student) and Morgane Goibert (M1 student) worked on hitting times and crossing times of Lévy processes, Matthieu Vert (M2 student) worked on the asymptotic distributions of non-homogeneous processes, Sophie Rossi (2nd year undergraduate student) worked on correlations in spatial processes, and Antoine Lucquiaud (PhD student) works on several of the same topics.

The second main stem of my work in the recent years is rooted in an interest I have developed for interdisciplinary modelling and in particular for the field that has become known as *complex systems*. These generally are systems with many-body and/or nontrivial interactions. Cities, for instance, may be regarded as an epitome of complex systems: individual and institutional agents interact on multiple levels of numerous networks (both physical and virtual), leading to nontrivial collective behaviour and nontrivial patterns at many scales. Most of my interdisciplinary research so far has focused on cities, and more specifically on segregation phenomena. I started with a variant of Thomas Schelling's celebrated model, but I then felt the need to move towards more data-based approaches (my belonging to a research team that counts many statisticians probably played a role there!).
This part of my work has been much enriched by the supervision of four Master (M2) theses (those of Rachneet Kaur, Thomas Lavoisier, Antoine Lucquiaud, Enrica Racca), one Master (M1) report (that of Luis Pinto Castaneda) and one undergraduate internship report (that of Ulysse Reverre). Under my supervision and the co-supervision of Paris-1 colleague Prof. Fabrice Rossi, Antoine Lucquiaud is now investigating further, in his PhD thesis, the theoretical aspects of a multi-agent model that we have introduced as a variant of the Schelling model. His work is also part of some of the research projects jointly organized with Paris-1 colleague Madalina Olteanu and UCLA Professor of Geography William Clark. Another important aspect in this part of my work has been the joint organizing, with the *Centre d'analyse et de mathématique sociales* (CAMS - EHESS/CNRS) of a series of seminars and workshops, notably a biennial interdisciplinary workshop (*Interactions*) bringing together social scientists and mathematical scientists, as well as an ongoing seminar series *in memoriam* Thomas Schelling.
Papers in international peer-reviewed journals and communications in international peer-reviewed workshops or conferences include:

- A. Hazan, J. Randon-Furling: *A Schelling model with switching agents: decreasing segregation via random allocation and social mobility*, The European Physical Journal B 86, 421. (2013)

- M. Olteanu, J. Randon-Furling: *Analyzing spatial dissimilarities via effective-time series*, International Time-Series Conference - ITISE, Granada, Spain. (2017)

- M. Cottrell, A. Hazan, M. Olteanu, J. Randon-Furling: *Multidimensional urban segregation - Toward a neural network measure*, WSOM+17, Nancy (2017). Also to appear in a special issue of Neural

Network and Applications (2019).

- J. Randon-Furling, M. Olteanu, A. Lucquiaud: *From urban segregation to spatial structure detection*, Environment & Planning B: Urban Analytics and City Science, doi.org/10.1177/2399808318797129 (2018)

- W. Clark, M. Olteanu, J. Randon-Furling: *A new method for analyzing ethnic mixing: studies from Southern California*, European Network for Housing Research Conference, Uppsala, Sweden (2018).

- J. Randon-Furling, M. Olteanu, W. Clark: *Converging to the city: a myriad trajectories*, International Conference on Complex Systems, Thessaloniki, Greece (2018).

I have been and I am involved in other interdisciplinary collaborations, which have led or will lead to publications and communications, but I only mention in this introduction the topics upon which I will enlarge in the following pages.

The first part of this thesis will be devoted to studies on stochastic processes and the second part to studies on segregation phenomena. Ongoing and future projects will be presented in the conclusion. A full list of publications and communications as well as a short CV are appended after the bibliography.

I. RANDOM WALKS, BROWNIAN MOTION, LÉVY PROCESSES

A. Edge and Facet Probabilities for Lévy-Brownian Convex Hulls

Stochastic processes are common in the physical and natural sciences (227), in particular as models of actual motion (26; 52; 90; 230), interfaces (120) or long proteins and other polymers (80; 84; 89; 233). Spatial characteristics of stochastic processes thus come to reflect physically meaningful variables: home ranges of animals (238), heights of surfaces (157), shapes of molecules (72; 117),...

One way to describe the spatial extent of a stochastic process's path is to examine its convex hull. That is, the smallest convex set enclosing all points visited by the process. In two dimensions, exact expressions for the values of geometric quantities such as the mean perimeter or the mean area of the convex hull have been found for Brownian motion (37; 68; 91; 222) as well as for other types of stochastic motions (136; 155; 198), for multi-walker systems (191) and for confined configurations (62; 63). In some cases, equivalent results exist in higher dimensional space (79; 92; 136; 137; 171; 187).

Results (108; 117; 202) pertaining to the shape of such random convex hulls appear scarcer though, despite their importance in biophysics: a protein's or an antibody's ability to play its part depends crucially on its shape (4; 237). Convex hulls methods are central in the description of long-molecule configurations (122; 150; 167; 220; 231), leading to questions on the structure of these hulls, such as on their number of edges (in 2D) or facets (in 3D) and on the distribution of distances between the molecule and its convex hull's edges or facets (11; 70).

Among the questions left open in my PhD thesis (187) was indeed that of counting the average number of edges appearing on the boundary of the convex hull of a Brownian path in the plane (see Fig. 1). Support functions and integral formulae had allowed us to compute the average perimeter and average enclosed area for the planar Brownian convex hull (158; 191). But the average number of edges on the boundary of the convex hull appeared not to be amenable to the same technique. Starting from another angle, I introduced the idea of "edge probabilities" and showed how these could be computed from the transition probabilities for Brownian excursions and meanders (61), provided one paid careful attention to normalization. Calculations actually led me to the discovery of a path transformation that allowed to go beyond the original planar Brownian path setting. I present below the main results obtained and the main ideas leading to them. This work was published as it unfolded, in a sequence of three papers (188; 189; 193). It was the subject of a plenary talk I was invited to give at the 2018 Stochastic Geometry Days – the annual conference of the CNRS research group GeoSto (GdR 3477).

1. Edges on planar Brownian convex hulls

Consider a planar Brownian motion $B(t) = (b_1(t), b_2(t))$, that is, a random variable in the plane corresponding to the position of a Brownian particle in some canonical reference frame. The coordinates b_1 and

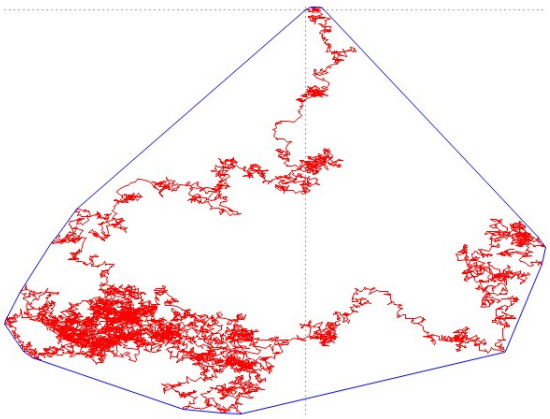


FIG. 1 A planar Brownian path. The dashed lines are coordinate axes in a canonical reference frame. We have represented the boundary of the convex hull: that is, the boundary of the smallest convex set enclosing all points of the path.

b_2 are thus two independent copies of standard, one-dimensional Brownian motion. We call $\Gamma(T)$ the path described in the plane by the Brownian particle up to time T,

$$\Gamma(T) = \{(b_1(s), b_2(s))\,,\ 0 \leq s \leq T\}\,. \tag{1}$$

$\mathcal{C}(T)$ is the convex hull of $\Gamma(T)$, i.e. the smallest convex set in the plane containing all points of the path at time T, and we write $\partial\mathcal{C}(T)$ for the boundary of $\mathcal{C}(T)$. Let us recall that $\partial\mathcal{C}(T)$ is (almost surely) a countable union of straight line segments (77; 91; 93).

Strictly speaking the number of edges is countably infinite, so we use a regularization strategy to analyze its behaviour and derive an explicit formula. The (simple) idea is to compute an "edge probability". Namely, picking two points on the path, $B(\tau), B(\tau+s) \in \Gamma(T)$, what is the probability that the line segment $[B(\tau), B(\tau+s)]$ is an edge on the boundary of the convex hull? To be more precise, let us write

$$A_{\tau,\tau+s} = \{[B(\tau), B(\tau+s)] \subset \partial\mathcal{C}(T)\}\,. \tag{2}$$

The edge probability is thus simply the probability of the event $A_{\tau,\tau+s}$. This probability may be written as the expectation of the indicator function of the event:

$$P\left(A_{\tau,\tau+s}\right) = \mathbb{E}\left[\mathbb{I}_{A_{\tau,\tau+s}}\right]. \tag{3}$$

So that, somewhat naively, the average number $\mathcal{N}(T)$ of edges on $\partial\mathcal{C}(T)$ may be written using the following, formal summation:

$$\text{“}\quad \mathcal{N}(T) = \mathbb{E}\left[\int\limits_{\substack{\Delta t \leq s \leq T \\ 0 \leq \tau \leq T-s}} \mathbb{I}_{A_{\tau,\tau+s}}\right] = \int\limits_{\substack{\Delta t \leq s \leq T \\ 0 \leq \tau \leq T-s}} P\left(A_{\tau,s}\right) \quad \text{”} \tag{4}$$

with Δt a regularization parameter (akin to a time separation below which we do not consider line segments). To make this rigorous, we need to compute a bivariate density $\rho(\tau, s)$ associated with $P\left(A_{\tau,\tau+s}\right)$, and we

shall see later on that this is indeed a *bona fide* joint density[1]. $\mathcal{N}(\mathcal{T})$ will then take the following form:

$$\mathcal{N}(T) = \int\limits_{\substack{\Delta t \le s \le T \\ 0 \le \tau \le T-s}} \rho\,(\tau, s)\; d\tau\, ds. \tag{5}$$

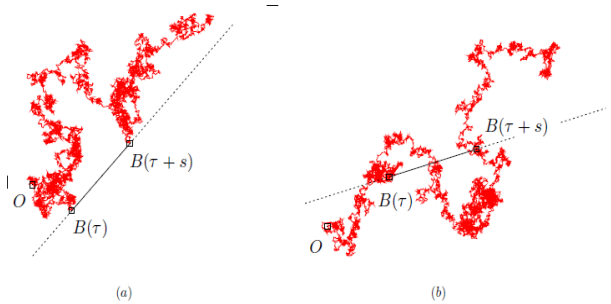


FIG. 2 Two possible configurations of the Brownian path $\Gamma(T)$, with respect to the line joining $B(\tau)$ and $B(\tau + s)$.

The key step is thus to compute the edge probability $P\,(A_{\tau,\tau+s})$. This may be done by shifting to a reference frame adapted to the straight line segment under consideration, as we explain below.

Given times $s \in [\Delta t, T]$ and $\tau \in [0, T - s]$, $\Gamma(T)$ may be found either:

- to lie on one side only of the line through $B(\tau)$ and $B(\tau + s)$, as in Fig.2 (a)
- or, to cross this line, as in Fig.2 (b).

The adapted reference frame we consider is one where the line through $B(\tau)$ and $B(\tau + s)$ is taken as the x-axis, as shown in Fig. 3. The "edge" event $A_{\tau,\tau+s}$ thus translates into constraints for the (new) y coordinate of the motion, *viz.*

C1. $y(t)$ starts at $y_0 > 0$ (say) and hits 0 for the first time at $t = \tau$;

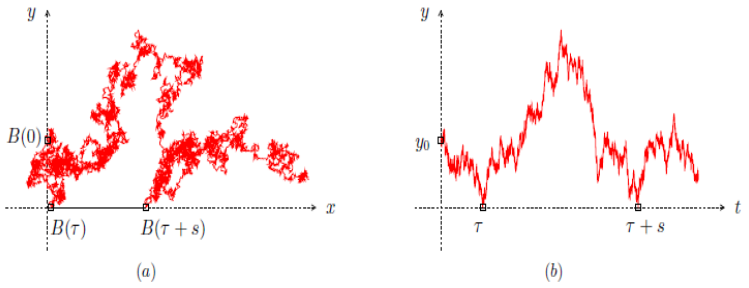


FIG. 3 The LHS configuration of Fig. 2 shown in an adapted reference frame (a), and the (new) y-coordinate plotted with respect to time (b).

[1] We shall see that $\rho(\tau, s)$ may be viewed as the joint density of the time of the global minimum of a Brownian motion with duration $T - s$, and the sojourn time in $\mathbb{R}_+$ of a Brownian bridge with duration s.

C2. $y(t)$ hits 0 again at $t = \tau + s$ but remains strictly positive between $t = \tau$ and $t = \tau + s$;

C3. $y(t)$ does not hit 0 after time $t = \tau + s$, it remains strictly positive up to time $t = T$.

Thanks to the Markovian nature of Brownian motion, we can separate the path into three parts and treat them independently. The constraints C1, C2, C3 mean we are then dealing with subpaths that are Brownian meanders and Brownian excursion[2]. We refer the reader to (188) for the full details of the calculation, in particular that of the normalization constants. Let us simply mention here the expression thus obtained for the edge probability:

$$P\left(A_{\tau,\tau+s}\right) = \frac{2\,d\tau\,ds}{\pi s\sqrt{\tau[(T-s)-\tau]}}, \tag{6}$$

that is,

$$\rho\left(\tau,\tau+s\right) = \frac{2}{\pi s\sqrt{\tau[(T-s)-\tau]}}. \tag{7}$$

The most interesting thing in this expression is the probabilistic interpretation that may be given to its factors. Indeed, $\frac{1}{\pi\sqrt{\tau[(T-s)-\tau]}} \times \frac{1}{s}$ is nothing but the product of the uniform density on an interval of length s (that is, $1/s$), and an arcsine density on the interval $[0, T-s]$ (that is, $1/\pi\sqrt{\tau[(T-s)-\tau]}$). This may be understood by "cutting and stitching" the y-coordinate path in the adapted reference frame shown in Fig. 3 (b). Treating the middle part (between τ and $\tau + s$) separately, one has a Brownian excursion or, seen otherwise, a Brownian bridge with a sojourn time on the positive side equal to its full duration, s. The corresponding sojourn time density is well-known to be uniform (151) and this leads to the $\frac{1}{s}$-factor. As for the two end parts of the path, they can be concatenated to produce a Brownian path of duration $T-s$ that hits its minimum at τ. The corresponding density is the well-known arcsine one, derived by Paul Lévy (152) – this gives the $\frac{1}{\pi\sqrt{\tau[(T-s)-\tau]}}$-factor. The overall factor of 2 comes from the fact that there are two sides to a straight line.

Although, in the Brownian case, the edge probability can be computed by "brute force" (one is, after all, simply talking about Gaussian integrals), the path transformation approach is much more appealing – and not only to those of us who dislike long calculations! Indeed a path transformation corresponds to a bijection between sets of paths, and chances are that it generalizes from Brownian motion to more general Markov processes. Path transformations are commonplace in the study of Brownian motion and Lévy processes (1; 29; 30; 31; 32; 36; 53; 54; 55; 59; 156; 228), and we shall make use of a few other ones in the following sections.

For now, let us remark that the "cutting and stitching" described above, together with the arcsine law for the time of the minimum of a Brownian motion and the uniform law for the sojourn time of a Brownian bridge, leads immediately to Eqs. (6) and (7). From which, upon integration of $\rho\left(\tau,\tau+s\right)$ over τ, one finds

$$\int_0^{T-s} \rho\left(\tau,\tau+s\right)\, d\tau = \frac{2}{s}, \tag{8}$$

[2] Let us recall that: (i) a Brownian bridge is a Brownian motion that starts at 0 and is constrained to come back to 0, (ii) a Brownian excursion is a Brownian bridge further constrained to always keep the same sign (equivalently, it is a Brownian motion constrained to always stay positive and come back to 0), (iii) a Brownian meander is a Brownian motion that starts at 0 and is constrained to stay positive (but it does not have to come back to 0). See Section I.B.

because $\rho(\tau, \tau+s)$ depends on τ only through the arcsine density on $[0, T-s]$, so that the integral over $\tau \in [0, T-s]$ yields 1. Then, to compute the average number of edges $\mathcal{N}(T)$ from Eqs. (5) and (8), all that is left is to integrate over $s \in [\Delta t, T]$, leading to

$$\mathcal{N}(T) = 2 \ln\left(\frac{T}{\Delta t}\right) \tag{9}$$

This result generalizes the discrete-time (random walk) result (20), through the idea of computing edge probabilities. This new idea was subsequently put to fruitful use also in the discrete-time setting (132), and in higher dimension (193).

But before presenting higher-dimensional results, let us mention some other two-dimensional results derived using edge probabilities.

2. Multiple planar Brownian paths

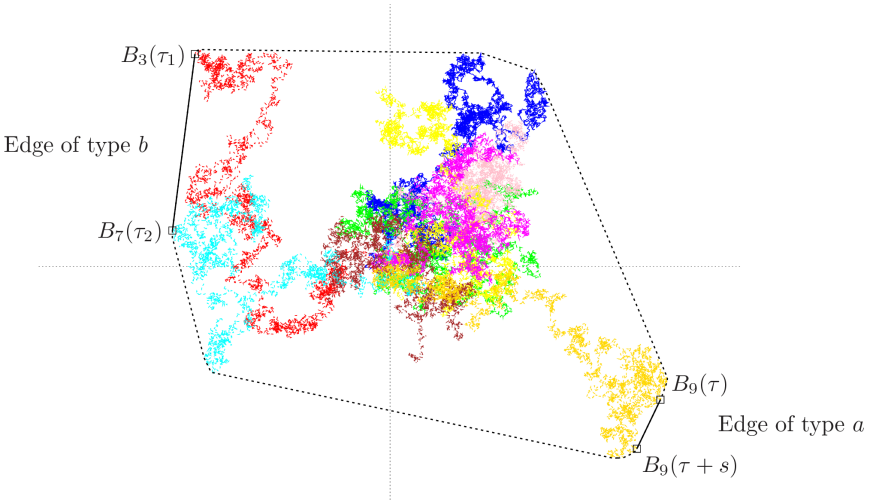


FIG. 4 Two types of edges on the convex hull of $n = 9$ independent planar Brownian paths (with same origin and same duration): edges joining points on the same path (a), edges joining points on two different paths (b).

Indeed, long but straightforward computation of Gaussian integrals allows one to obtain the average number, $\mathcal{N}_n(T)$, of edges on the boundary of the convex hull of $n \geq 1$ independent Brownian paths. One has to consider the two types of edges that can exist: edges joining two points on the same Brownian path, and edges joining two points on two different paths (Fig. 4). In the relevant adapted reference frame, this induces two types of configurations for the y-coordinates, as shown in Fig 5. Computing the associated densities leads to a closed expression for $\mathcal{N}_n(T)$ (this was done in the second part of (188) and we skip details here):

$$\mathcal{N}_n(T) = \alpha_n(T, \Delta t) + \beta_n(T, \Delta t) \tag{10}$$

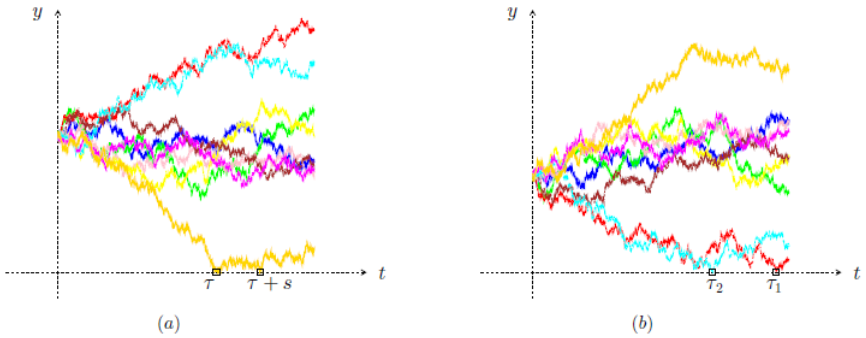


FIG. 5 One-dimensional (y-coordinate) configurations corresponding to the two types of edges on the convex hull of $n = 9$ independent planar Brownian paths.

with

$$\alpha_n(T, \Delta t) = \int_{\frac{\Delta t}{T}}^{1} \int_0^{1-v} \int_0^{\infty} h_n^{(a)}(u, v, z)\, du\, dz\, dv, \tag{11}$$

$$\beta_n(T, \Delta t) = \int_{\frac{\Delta t}{T}}^{1} \int_{\max(0,v-1)}^{\min(1,v)} \int_0^{\infty} h_n^{(b)}(u, v, z)\, du\, dz\, dv, \tag{12}$$

where

$$h_n^{(a)}(u, v, z) = 4n\, \frac{u \exp\left(-\frac{u^2}{z}\right) [\mathrm{erf}(u)]^{n-1}}{\pi v z \left[z(1-v-z)\right]^{\frac{1}{2}}} \tag{13}$$

$$h_n^{(b)}(u, v, z) = 4n(n-1)\, \frac{v^{\frac{1}{2}} u^2 \exp\left(-\frac{vu^2}{z(v-z)}\right) [\mathrm{erf}(u)]^{n-2}}{\pi z(v-z)\, \left[\pi z(1-z)(v-z)(1-v+z)\right]^{\frac{1}{2}}} \tag{14}$$

(We write $\mathrm{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-x^2}\, dx$ for the error function.)

These expressions, though maybe not very much pleasing to the eye, are an example of exact analytical results that may be obtained using edge probabilities in a context where the underlying probability densities are tractable (typically, Gaussian densities).

3. Edges on planar Lévy convex hulls

Now, let us come back to single path contexts. As mentioned above, the path transformation that leads directly to the edge probability may be used to compute the average number of edges on more general Lévy convex hulls, *eg* convex hulls of symmetric α-stable Lévy processes other than simple Brownian motion. Recall Eq.6:

$$P\left(A_{\tau,\tau+s}\right) = \mathbb{E}\left[\mathbb{I}_{A_{\tau,\tau+s}}\right] = \frac{2\, d\tau\, ds}{\pi s \sqrt{\tau[(T-s)-\tau]}}.$$

Recall also that the $\frac{1}{\pi\sqrt{\tau[(T-s)-\tau]}}$- factor corresponds to the two end parts of the path: once stitched together, these amount to a process of duration $T - s$ that attains its minimum at time τ. Last but not least, recall

that the edge probability is integrated over $\tau \in [0, T-s]$. So that, in fact, one does not even need to know the specific form of the underlying marginal density (the arcsine density in the Brownian case): it is going to be integrated out and yield 1 in any case. Thus the only important factor in the edge probability is $\frac{2}{s}$. Where did this factor come from? From the middle part of the process, the one that corresponds to an excursion away from the edge. Such an excursion may also be viewed as a bridge with a sojourn time in the positive (or negative) side equal to its full duration. As it happens, the sojourn time density is uniform not only for Brownian bridges, but also for bridges of a large class of cyclically exchangeable processes (104; 139), including for instance most symmetric α-stable processes[3]. These so-called uniform laws are the continuous-time equivalent of combinatorial identities for random walks (20; 134; 210; 211; 212; 213; 214). They emerge from the profound symmetry of cyclically exchangeable processes and are intimately linked to Ballot theorems (135) (we shall encounter Ballot theorems (6; 16; 34; 224; 235) also in the second part of this thesis, see Section II.B.5.)

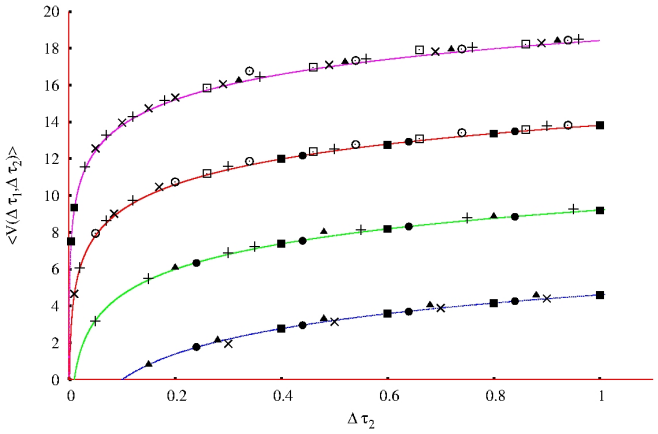


FIG. 6 Computer simulation results (points) compared with the exact formula from Eq. (15) (solid lines) giving the average number of s-edges with $s \in [\Delta\tau_1, \Delta\tau_2]$, for a range of values of $\Delta\tau_2$ and $\Delta\tau_1$ ($\Delta\tau_1 = 10^{-1}$ on the bottom curve, 10^{-2} on the second curve, 10^{-3} on the third one and 10^{-4} on the top curve). Symbols correspond to computer simulations with 10^4 paths for: Brownian motion with diffusion constant set to 1/2 (filled squares) or 50 (filled circles), Cauchy-Lorentz process (filled triangles), symmetric α-stable processes with $\alpha = 0.75$ (squares) and $\alpha = 1.5$ (circles), compound Poisson processes with drift and with normal (pluses) or Cauchy (crosses) jump variables.

Therefore, one obtains the same logarithmic form as in Eq. 9 for the general symmetric Lévy case. This form may be further generalized as to express the average number of edges joining points with a time-separation in $[\Delta t_1, \Delta t_2]$

$$\mathcal{N}_{[\Delta t_1, \Delta t_2]}(T) = 2 \int_{\Delta\tau_1}^{\Delta\tau_2} \frac{1}{s}\, ds = 2 \ln\left(\frac{\Delta\tau_2}{\Delta\tau_1}\right) \tag{15}$$

This formula shows the remarkable "universal" behaviour of the average number of edges on the convex hull of symmetric processes, as well as a time-scale invariance: there are as many edges joining points

[3] The Fourier transform of the process needs to be integrable.

separated by $\delta t \in [\Delta t_1, \Delta t_2]$ as edges joining points separated by $\delta t \in [\lambda \Delta t_1, \lambda \Delta t_2]$ (for $\frac{\Delta t}{\Delta t_1} \leq \lambda \leq \frac{T}{\Delta t_2}$). Numerical simulations illustrate the validity of this general formula for various symmetric Lévy processes, as can be seen in Fig. 6.

Let us now show that the universality observed for the average number of edges in two dimensions extends to the average number of facets in $d \geq 3$ dimensions (Fig.7).

4. Convex hulls of higher-dimensional Lévy processes

Moving to higher dimensions, edges become facets and the same idea of computing "facet probabilities" leads to an exact, integral formula for the average number of facets on the convex hull of a vast class of higher-dimensional Lévy processes. I published this extension with Florian Wespi (193), a student of Prof. Ilya Molchanov (171) at the University of Bern, Switzerland. In particular, we showed that in dimension $d = 3$, the integral formula for the average number of facets admits a closed form,

$$\mathcal{N}^{(3)}(T) = 2\left[\ln\left(T/\Delta t\right)\right]^2 + 4\left\{\ln\left(T/\Delta t\right)\ln\left(1-\Delta t/T\right) - \mathrm{Li}_2\left(1-\Delta t/T\right) + \pi^2/12\right\},$$

where Li_2 is the dilogarithm function, T is the total duration of the process and Δt the usual cut-off parameter representing an arbitrary minimal time separation between two extreme points on the process's path. We also established asymptotics in the general d-dimensional case,

$$\mathcal{N}^{(d)}(T) \sim 2\left[\ln\left(T/\Delta t\right)\right]^{d-1}, \tag{16}$$

which nicely generalizes the two-dimensional case.

For the sake of clarity, let us detail here first the 3-dimensional discrete-time case.

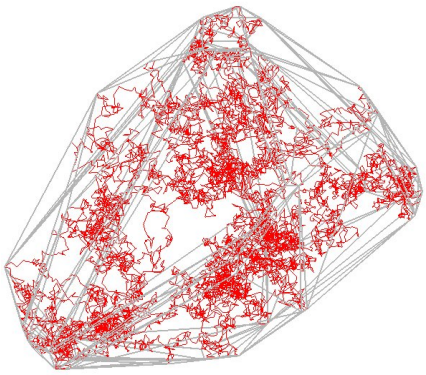


FIG. 7 A sample realization of 3-dimensional Brownian motion and the facets on the boundary of its convex hull.

Discrete-time case

In three-dimensional space, the convex hull of the successive positions of a random walker is almost surely made up of triangular facets. The idea is therefore to compute the average number of such facets by computing the probability for a triangle defined by three points on the walker's path to form a facet of the convex hull (see Fig. 7).

Let S_n denote the position of the walker after n steps ($S_0 = 0$, $n \geq 3$). Given three integers $n_1 < n_2 < n_3$, the triangle joining S_{n_1}, S_{n_2}, and S_{n_3} will be a facet of the walk's convex hull if the walker always stays on the same side of the plane defined by S_{n_1}, S_{n_2}, S_{n_3}. Let us choose a reference frame where the xy-plane coincides with the one defined by S_{n_1}, S_{n_2}, S_{n_3}. Then the three-dimensional random walk always stays on the same side of the xy-plane if its z-coordinate is a one-dimensional random walk that: (i) starts at some value – say a positive one – and hits 0 for the first time at time n_1; (ii) performs a positive excursion away from 0 between n_1 and n_2 (i.e. it hits 0 at these times but remains positive in between); (iii) performs another positive excursion away from 0 between n_2 and n_3; and finally (iv) leaves from 0 at time n_3 and stays positive thereafter.

Relabeling $i = n_1$, setting $k_1 = n_2 - n_1$ and $k_2 = n_3 - n_2$, one may formally write the following formula for the mean number $\mathcal{N}_N^{(3)}$ of facets

$$2 \sum_{k_1=1}^{N-1} \sum_{k_2=1}^{N-k_1} \sum_{i=0}^{N-(k_1+k_2)} m_i^{(0)} e_{i+k_1}^{(i)} e_{i+k_1+k_2}^{(i+k_1)} m_N^{(i+k_1+k_2)}, \tag{17}$$

with $m_i^{(0)}$, $e_{i+k_1}^{(i)}$, $e_{i+k_1+k_2}^{(i+k_1)}$, $m_N^{(i+k_1+k_2)}$ standing for the probabilities of each of the four parts (i)-(iv) described above. The prefactor 2 accounts for the fact that there are two sides to every facet.

Since we are considering only random walks with independent and identically distributed jumps, we have $e_{i+k_1}^{(i)} = e_{k_1}^{(0)} \equiv e_{k_1}$ and $e_{i+k_1+k_2}^{(i+k_1)} = e_{k_2}^{(0)} \equiv e_{k_2}$ for all i, k_1, and k_2. Therefore, Eq. (17) becomes

$$\mathcal{N}_N^{(3)} = 2 \sum_{k_1=1}^{N-1} \sum_{k_2=1}^{N-k_1} e_{k_1} e_{k_2} \sum_{i=0}^{N-(k_1+k_2)} m_i^{(0)} m_N^{(i+k_1+k_2)}. \tag{18}$$

Markovian independence allows one to stitch the parts of the z-random walk described above as separate pieces (because the cuts occur at special times, namely stopping times). In particular, stitching together the first and fourth parts produces a sub-walk of duration $N-(k_1+k_2)$. This walk starts at some positive value and hits its minimum (0 by construction) at time $i = n_1$. The probability associated with this sub-walk is just the probability for a random walk with $N-(k_1+k_2)$ steps to hit its minimum at step i, just as before in the two-dimensional continuous-time case, when this corresponded to the arcsine law (Section I.A.1). Thus interpreting the product $m_i^{(0)} m_N^{(i+k_1+k_2)}$ in Eq. (18), one finds that

$$\sum_{i=0}^{N-(k_1+k_2)} m_i^{(0)} m_N^{(i+k_1+k_2)} = \sum_{i=0}^{N-(k_1+k_2)} \text{Prob}\,(\text{Min. is hit at } i) = 1.$$

Substituting into Eq. (18) yields

$$\mathcal{N}_N^{(3)} = 2 \sum_{k_1=1}^{N-1} \sum_{k_2=1}^{N-k_1} e_{k_1} e_{k_2}. \tag{19}$$

There remains to compute the excursion probabilities e_{k_1}, e_{k_2}. As before, let us notice that the excursion probability for a random walk is the probability that the walk, pinned at 0 at times 0 and N, visits only the positive half-space. This is simply given by the probability that a bridge (that is, a random walk pinned at 0 at times 0 and N) attains its minimum at the initial step. The discrete-time equivalent of the uniform law for bridges, Baxter's combinatorial lemma (20), therefore leads to $e_k = 1/k$. (Note that using generalizations of this combinatorial lemma (17; 133), one could treat the three-dimensional case here without resorting to a projection on one of the coordinates.)
Finally, we obtain

$$\mathcal{N}_N^{(3)} = 2 \sum_{k_1=1}^{N-1} \sum_{k_2=1}^{N-k_1} \frac{1}{k_1 k_2} = 2 \sum_{\substack{k_1+k_2 \leq N \\ k_1, k_2 \geq 1}} \frac{1}{k_1 k_2}. \tag{20}$$

In higher dimension, the same reasoning is readily carried out. Counting facets that can be defined by d points on the path of a random walk with $N \geq d$ steps, one obtains a $(d-1)$-fold sum

$$\mathcal{N}_N^{(d)} = 2 \sum_{\substack{k_1+\cdots+k_{d-1} \leq N \\ k_1, \ldots, k_{d-1} \geq 1}} \frac{1}{k_1 \cdot k_2 \cdots k_{d-1}}. \tag{21}$$

While Florian Wespi and I were writing our paper, a pre-print (132) came to our attention: it contained our Eq. (21). Several other results pertaining to the discrete-time case are to be found in (132), but none on the continuous-time case, which we now examine.

Continuous-time case

Given the universality observed in both the discrete-time case and the two-dimensional continuous time case, one expects a simple continuous version of Eq.(21) to hold in the d-dimensional continuous-time case.

This may be proven using the same strategy as in the previous section. In dimension d, a facet will be a $(d-1)$-dimensional object lying in a hyperplane and defined by d points on the process's path. One thus splits the path into $d+1$ independent parts, and then computes the probability density (pdf) associated with each part before multiplying them together and integrating. When $d = 3$, the four parts and their associated pdfs are:

- from time 0 up to time τ, the path lies on one side only (say side +) of a (hyper)plane $\mathcal{H}$, and it hits $\mathcal{H}$ at time τ; write $p_\tau^{(0)}$ for the corresponding pdf;

- from time τ up to time $\tau + k_1$, the path is an excursion away from $\mathcal{H}$ (in side +) and is pinned on $\mathcal{H}$ at times τ and $\tau + k_1$; write $f(\tau, \tau + k_1)$ for the pdf;
- from time $\tau + k_1$ up to time $\tau + k_1 + k_2$, the path is an excursion away from $\mathcal{H}$ (in side +) and is pinned on $\mathcal{H}$ at times $\tau + k_1$ and $\tau + k_1 + k_2$; write $f(\tau + k_1, \tau + k_1 + k_2)$ for the pdf;
- from time $\tau+k_1+k_2$ up to time T, the path lies on side + of $\mathcal{H}$ and it is pinned on $\mathcal{H}$ at time $\tau+k_1+k_2$; write $p_T^{(\tau+k_1+k_2)}$ for the pdf.

Formally, one obtains the following integral

$$\mathcal{N}^{(3)}(T) = 2\int_{\Delta t}^{T-\Delta t}\int_{\Delta t}^{T-k_1}\int_{0}^{T-(k_1+k_2)} p_\tau^{(0)} f(\tau, \tau + k_1) f(\tau + k_1, \tau + k_1 + k_2) p_T^{(\tau+k_1+k_2)}\, d\tau\, dk_2\, dk_1. \quad (22)$$

Since Lévy processes have stationary and independent increments, we have, just as in the discrete-time case, for all τ, k_1 and k_2: $f(\tau, \tau + k_1) = f(0, k_1) \equiv f(k_1)$ and $f(\tau + k_1, \tau + k_1 + k_2) = f(0, k_2) \equiv f(k_2)$. Also, the product $p_\tau^{(0)} p_T^{(\tau+k_1+k_2)}$ can be interpreted as giving the pdf of the time τ at which a similar process of duration $T - (k_1 + k_2)$ attains its minimum. Thus,

$$\int_{\tau=0}^{T-(k_1+k_2)} p_\tau^{(0)} p_T^{(\tau+k_1+k_2)}\, d\tau = 1, \quad (23)$$

and the triple integral in Eq. (22) reduces to

$$\mathcal{N}^{(3)}(T) = 2\int_{\Delta t}^{T-\Delta t}\int_{\Delta t}^{T-k_1} f(k_1) f(k_2)\, dk_2\, dk_1. \quad (24)$$

Lastly, $f(k)$ may be viewed as the pdf that the sojourn time of a process on a given side of a (hyper)plane to which it is pinned (at both endpoints) is equal to the full duration k. This means, $f(k)$ is the pdf of the sojourn time of a Lévy bridge. Direct higher dimensional results (39; 83) or the same reasoning as before (i.e. using an adapted frame and reducing the problem to a one-dimensional one) shows that the sojourn-time distribution is uniform for a vast class of Lévy processes (104; 135; 139). Therefore $f(k) = 1/k$, and we obtain

$$\mathcal{N}^{(3)}(T) = 2 \int\limits_{\substack{k_1+k_2\le T \\ k_1,k_2\ge\Delta t}} \frac{dk_1\, dk_2}{k_1 k_2}. \quad (25)$$

In higher dimension, Eq. (22) becomes a d-fold integral, accounting for facets defined by d points and characterised by $d-1$ excursions:

$$\mathcal{N}^{(d)}(T) = 2\int_{\Delta t}^{T-(d-2)\Delta t}\int_{\Delta t}^{T-(d-3)\Delta t-k_1} \cdots \int_{\Delta t}^{T-(k_1+\cdots+k_{d-2})} \frac{dk_{d-1}\cdots dk_2\, dk_1}{k_1 k_2 \cdots k_{d-1}},$$

that is,

$$\mathcal{N}^{(d)}(T) = 2 \int\limits_{\substack{k_1+\cdots+k_{d-1}\le T \\ k_1,\ldots,k_{d-1}\ge\Delta t}} \frac{dk_1\cdots dk_{d-2}\, dk_{d-1}}{k_1 k_2 \cdots k_{d-1}}. \quad (26)$$

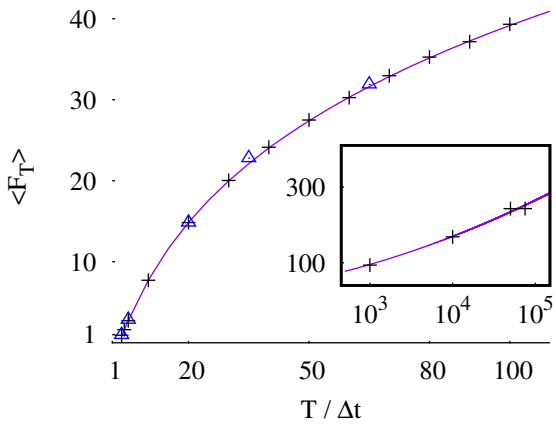


FIG. 8 Numerical simulations for the average number of triangular facets on the convex hulls of some 3-dimensional Lévy processes (with 10^4 and 10^3 sample paths realizations): standard Brownian motion (pluses), and a stable process with stability index 1.5 (triangles). The solid line corresponds to the exact formula Eq. (29). The inset shows the asymptotics for large values of $T/\Delta t$, from Eq (30), with numerical simulations of 3-dimensional Brownian motion.

In the three-dimensional case, it is possible to obtain a closed form for this integral formula, as detailed in the next paragraph.

Closed formula and asymptotics

Starting from Eq. (25), we change variables to $y_1 = k_1$, $y_2 = k_1 + k_2$. We find

$$\mathcal{N}^{(3)}(T) = 2\int_{\Delta t}^{T-\Delta t} \frac{dy_1}{y_1} \int_{y_1+\Delta t}^{T} \frac{dy_2}{y_2 - y_1}, \tag{27}$$

in which the second integral is easily computed. This leads to:

$$\mathcal{N}^{(3)}(T) = 2\int_{\Delta t}^{T-\Delta t} \frac{dy_1}{y_1} \left[\ln\left(\frac{T}{\Delta t}\right) + \ln\left(1 - \frac{y_1}{T}\right)\right]. \tag{28}$$

Setting $z = y_1/T$ allows one to compute the remaining integral (see also (166; 239)) and yields the announced result,

$$\mathcal{N}^{(3)}(T) = 2\left[\ln\left(T/\Delta t\right)\right]^2 + 4\left\{\ln\left(T/\Delta t\right)\ln\left(1 - \Delta t/T\right) - \mathrm{Li}_2\left(1 - \Delta t/T\right) + \pi^2/12\right\}, \tag{29}$$

where Li_2 is the dilogarithm function. As $\mathrm{Li}_2\left(1^-\right)$ is finite, the first term on the right hand side of Eq.(29) is also clearly the leading term when $T/\Delta t$ becomes large. Thus,

$$\mathcal{N}^{(3)}(T) \sim 2\left[\ln\left(T/\Delta t\right)\right]^2. \tag{30}$$

This asymptotical behaviour may readily be directly extracted from Eq. (28), focusing on the $\ln\left(T/\Delta t\right)$ term in the integrand. This transposes straightforwardly, by simple iteration, to the $(d-1)$-dimensional integral in Eq (26), leading to:

$$\mathcal{N}^{(d)}(T) \sim 2\left[\ln\left(T/\Delta t\right)\right]^{d-1}. \tag{31}$$

This result generalizes in a very simple manner the two-dimensional one as well as the discrete-time one. The exact formula (Eq. (29)) and the asymptotical behaviour (Eq. (30)) are illustrated in numerical simulations, see Fig. 8.

Once again, the exact formula established here for the average number of $(d-1)$-dimensional facets on the convex hull of a general d-dimensional Lévy process shows how universal certain aspects of the shape of spatial random processes are. Indeed this formula is valid not only for d-dimensional Brownian motion but for a vast class of Lévy processes, including all stable processes. This is reminiscent of the universality observed in the Sparre Andersen theorem (211; 212), to which the formula established here is linked via the arcsine and the uniform laws.

B. More on Excursions, Meanders, and Path Transformations

As mentioned and illustrated in the previous section, path transformations may be very useful tools in the study of stochastic processes. We give here another example, and a third one will be given in the next section. As excursions and meanders will be involved again here, first let us take time to recall some definitions.

1. Excursions and Meanders

In his seminal work on Brownian excursions (61), following (87; 125; 151; 152), K.L. Chung surveys variants of Brownian motion conditioned to stay positive. For a standard Brownian motion B_t, the excursion straddling t and the meandering process ending at t may be defined from the last zero before t and the first zero after t, namely:

$$\gamma(t) = \sup\{s \leq t, B_s = 0\}$$

and

$$\beta(t) = \inf\{s \geq t, B_s = 0\}.$$

Then the absolute value of the process between $\gamma(t)$ and $\beta(t)$, conditional on these, is the excursion straddling t. Thus an excursion starts from 0 and returns to 0 but without hitting 0 between its initial and final points (see Fig. 9). If one performs conditioning simply on $\gamma(t)$ and focuses on the pre-t part, one obtains a meandering process: a process starting at 0 and never hitting 0 again on the time interval $[\gamma(t), t]$.

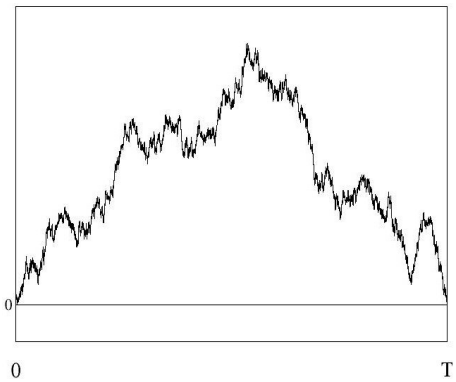


FIG. 9 Brownian excursion: Brownian motion conditioned to stay positive and to return to 0.

Excursions and meanders may similarly be defined for general Lévy processes. In the case of Brownian motion, the propagator[4] for the meander is known to be (61)

$$g(x;t) = \frac{x}{t}\, e^{-\frac{x^2}{2t}}. \tag{32}$$

[4] I use here more of a physicist's vocabulary. The propagator is the same as the transition density of the stochastic process. For instance the propagator of standard Brownian motion is $f(x,y;t) = \frac{e^{-(y-x)^2/2t}}{\sqrt{2\pi t}}$.

Recalling that the first passage-time density for a Brownian motion, that is the probability density associated with the first zero of a Brownian motion starting at $x > 0$ is

$$\rho(x,t) = \frac{x}{\sqrt{2\pi t^3}}\, e^{-\frac{x^2}{2t}}, \tag{33}$$

and that the standard Brownian propagator from 0 to 0 in time t is

$$f(0,0;t) = \frac{e^{-(0-0)^2/2t}}{\sqrt{2\pi t}} = \frac{1}{\sqrt{2\pi t}}, \tag{34}$$

one notices that:

$$g(x;t) = \frac{\rho(x,t)}{f(0,0;t)}. \tag{35}$$

This is in fact not a mere coincidence. A simple path transformation accounts for this observation and allows one to establish the same result for all symmetric α-stable processes, as we shall see below.

2. Path transformations

A well known example of a path transformation is the so-called method of images, which according to Feller is due to Lord Kelvin (97). This transformation allows one to compute the propagator from $x > 0$ to $y > 0$ for a Brownian motion constrained to stay positive:

$$q(x,y;t) = \frac{1}{\sqrt{2\pi t}} \left(e^{-\frac{(y-x)^2}{2t}} - e^{-\frac{(y+x)^2}{2t}} \right). \tag{36}$$

Indeed, the method of images is equivalent to transforming paths that go from $x > 0$ to $y > 0$ *and* hit 0 into paths that go from $-x$ to y. The bijection thus established enables one to identify the right factor to subtract from the standard Brownian density in order to obtain the constrained propagator (Eq. 36).

Another example of path transformation is Vervaat's construction of a Brownian excursion from a Brownian bridge (228).

Let us describe now the path transformation we wish to use here. We start with a Brownian motion B of duration $t > 0$. The probability density function (pdf) for such a motion to be actually a Brownian bridge (that is, $B_0 = B_t = 0$) is given by the Brownian propagator from 0 to 0 in time t:

$$f(0,0;t) = \frac{e^{-(0-0)^2/2t}}{\sqrt{2\pi t}} = \frac{1}{\sqrt{2\pi t}}. \tag{37}$$

We constrain B to be a Brownian bridge and we let τ_m be the time when B attains its minimum on the interval $[0,t]$ – note that this minimum is non-positive. Consider now the pre-τ_m and the post-τ_m parts. For $s \in [0, \tau_m]$, define:

$$W_s^+ = B_{\tau_m - s} - B_{\tau_m}, \tag{38}$$

and for $s \in [\tau_m, t]$:

$$W_s^+ = B_s - 2B_{\tau_m}. \tag{39}$$

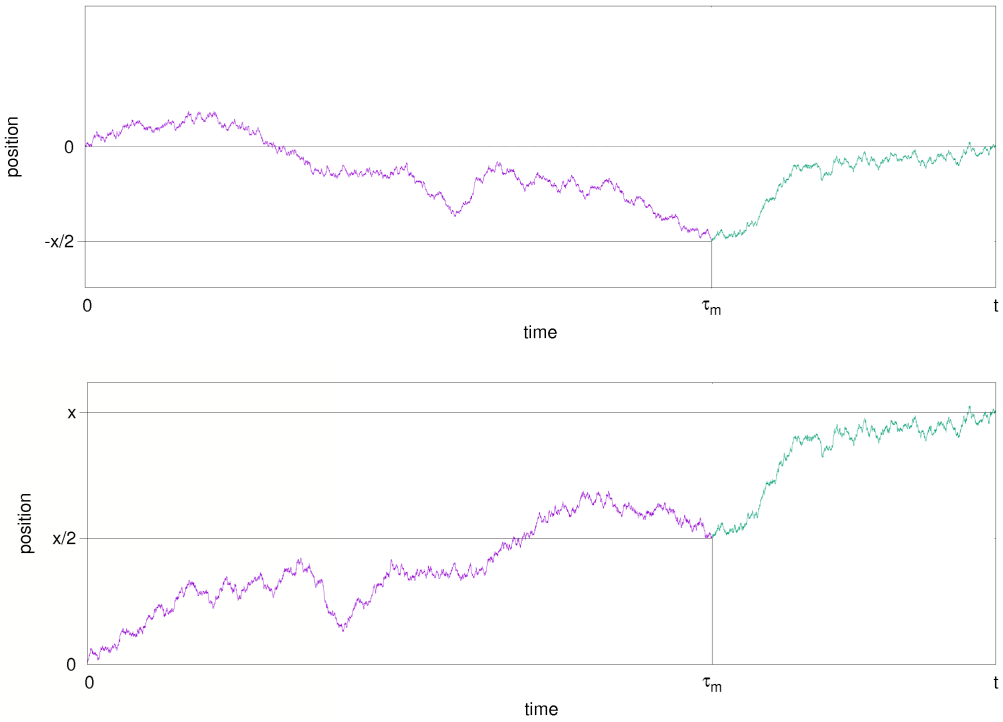


FIG. 10 Path transformation leading from a Brownian bridge with minimum $-x/2$ (top) to a Brownian meander with final value $x > 0$ (bottom). The pre-minimum part is time-reversed and shifted upward by $x/2$ while the post-minimum part is shifted upxard by x.

That is, the pre-τ_m part is time-reversed and the post-τ_m part is shifted to be appended to the new path, as shown in Fig. 10. Recall that $B(0) = 0$ and note that $B(\tau_m) \leq 0$.
This path transformation produces a Brownian meander W^+ of duration t, with final point $x > 0$, where $-x/2$ is the minimum of the initial Brownian bridge.

Having reached it independently, I found out later on that, although this transformation does not exist in the physics literature, it does in the probability literature (29; 32).

3. Propagator for the Brownian meander

The path transformation defined in Eq. 38 and 39 means that the propagator for a Brownian meander with terminal point $x > 0$ is the same as the pdf for the minimum of a Brownian bridge evaluated at $-x/2$. (Both meander and bridge having same duration t.) The latter is well known (85) to be $\frac{x}{t}\, e^{-\frac{x^2}{2t}}$ and thus one obtains directly the propagator of the meander, as in Eq. 32.
One may also, thus, arrive directly at Eq 35 and understand why this equality is not a mere coincidence. Indeed, since the path transformation is a bijection, it implies that the total probability density associated with Brownian meanders of duration t (whatever their endpoint) is the same as that associated with Brownian bridges of duration t (whatever the value of their minimum). In other terms, the partition function for Brownian meanders of duration t is $f(0, 0; t) = \frac{1}{\sqrt{2\pi t}}$. And the numerator to be set against this partition

function if one wishes to compute the propagator of a meander ending at $x > 0$ is simply the pdf associated with a standard Brownian path starting from x and first hitting 0 at time t – namely the first-passage time density $\rho(x,t)$. Hence Eq. 35:

$$g(x,t) = \frac{\rho(x,t)}{f(0,0;t)}.$$

Incidentally, this also explains why

$$\int_0^\infty \rho(x,t)\,dx = f(0,0;t); \tag{40}$$

an equality which, though easily checked in the Brownian case, could have been mistaken as purely coincidental, whereas it reflects the bijection between meanders and bridges.

4. Generalization to other Lévy processes

The path transformation introduced above relies only on properties of Brownian motion that are actually valid also for symmetric stable processes, most importantly Markovian and time-reversal properties. The same result therefore holds – more precisely, one needs to be careful with the jumps that appear in stable processes, and it is better in this case to append the time-reversed pre-mminimum part after the post-minimum one (30; 54; 57; 58). This leads to new equalities on the spatial integrals of first-passage-time densities for symmetric stable processes.

For instance in the case of a Cauchy process (i.e. a symmetric α-stable process with $\alpha = 1$), one finds that:

$$\int_0^\infty \rho(x,t)\,dx = \frac{1}{\pi t}. \tag{41}$$

And also, more generally, for a symmetric α-stable process:

$$\int_0^\infty \rho(x,t)\,dx = f(0,0;1)t^{-\frac{1}{\alpha}}. \tag{42}$$

Thus, thanks to a path transformation, one obtains identities relating probability density functions that, but in a few cases, are not known explicitly. We will see this at play again in the next section, where I detail results obtained on persistence exponents and lead changes.

C. Persistence Exponents and Lead Changes

In this section, we continue our presentation of a number of theoretical results pertaining to Brownian motion and Lévy processes. We put the emphasis on the use of techniques such as path transformations and reciprocal processes. Let us start with a few remarks on persistence exponents and survival probabilities.

1. Persistence and Survival

In many applications of stochastic modeling, a key question regards the time that a random process will "survive" before reaching (or crossing, if the process has jumps) a certain point or an absorbing boundary, away from which it started. Examples abound, from spin dynamics (13) to financial markets (163), reaction kinetics (25), biological systems (115) or climate science (145) – to name but a few (197).

Such survival times are usually described by the exponent governing the power-law decay of the long-time asymptotics for the first passage or first crossing time (FPT) density (13). They are linked with the time at which the maximum is achieved (56; 139) and with excursions (fluctuations) away from the average (14). In a handful of cases, the probability density functions (pdf) associated with the FPT at 0 can be computed exactly. For instance, for Brownian motion starting at $x_0 > 0$ with diffusion constant D, the image method yields the well-known density $\rho_D(x_0, \tau) = x_0(4\pi D\tau^3)^{-1/2} \exp(-x_0^2/4D\tau)$, already recalled in the previous section. The *survival* probability is defined as the tail of the distribution, that is the complementary cumulative distribution function. In the Brownian case, its leading term in the large-τ asymptotics will decay as $\tau^{-1/2}$. Hence a *persistence* exponent $\theta = 1/2$.

As a consequence of E. Sparre Andersen's celebrated results (211; 212), this value of $1/2$ expresses a universal behaviour that extends to the more general case of *symmetric* continuous-time Markov processes (60; 143; 197; 241), even when the method of images breaks down (60). In the case of skewed processes, however, the situation is more complex and offers a larger variety of exponents (144). De Mulatier *et al* (174) showed that the persistence exponent associated with the survival probability $P_{LF}^{(x_0)}$ for asymmetric Lévy flights[5] with stability index $\alpha \neq 1$ and asymmetry parameter $\beta \in [-1, 1]$ can be computed exactly: the result is a family of exponents $\theta(\alpha, \beta)$, which continuously deviates from the Sparre Andersen value of $1/2$ (recovered for $\beta = 0$).

In the case of non Markov processes, persistence exponents have been studied especially in connection with non-equilibrium dynamics and turbulence (121; 162; 168; 176; 181; 183). Particular interest has been devoted to dynamical phenomena involving extreme-value processes (21; 47; 82; 159; 160; 161; 208). Ben-Naim and Krapivsky (22) computed the persistence exponent associated with the order statistics of two Brownian maxima (see Fig. 11). They showed that if $m_1 - m_2 = a > 0$ initially, the probability $P_{BM}^{(a)}(t)$ that the maxima have remained ordered in the same way up to a time τ, decays, when τ is large, like $\tau^{-\delta}$, where δ

[5] In the physics literature, Lévy flights are α-stable processes (or discrete-time random walks with heavy-tailed jump distributions, that converge to α-stable processes). See also the next subsection, where we recall definitions.

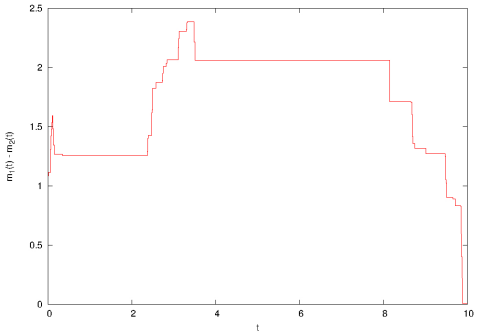


FIG. 11 Sample path, until it first crosses 0, of the non Markov process $m_1(t) - m_2(t)$, where m_1 and m_2 are the running maxima of two independent Brownian walkers.

is linked to the diffusion constants of the Brownian particles as follows: $\delta(D_1, D_2) = (1/\pi)\arctan(\sqrt{D_2/D_1})$. In particular, when $D_1 = D_2$, $\delta = 1/4$.

In (190), I established a formula that relates the survival probability of certain Lévy flights (therefore Markov processes) studied by de Mulatier *et al.*, with the survival probability examined by Ben-Naim and Krapivsky in a non Markovian context. This formula allows one to show how the persistence exponents in both cases are directly related via:

$$\delta(D_1, D_2) = \theta(\alpha, \beta)/2, \tag{43}$$

with $\alpha = 1/2$ and $\beta = \left(\sqrt{D_1} - \sqrt{D_2}\right) / \left(\sqrt{D_1} + \sqrt{D_2}\right)$. The result also reveals how the 1/4-exponent obtained in (22) when $D_1 = D_2$, can be derived directly from the Sparre-Andersen universal behaviour of an underlying, hidden Markov process governing the order statistics of the two Brownian maxima.

A key point on which I based my reasoning is the fact, noted by P. Lévy (152), that the path of the running-maximum process, $M(t)$, of a Brownian motion, is increasing and therefore admits a reciprocal function[6]. This function corresponds to a process, $T(x)$, which is also increasing. But contrary to $M(t)$, $T(x)$ is Markovian. Let us first recall some details about the T process and about general Lévy-Brownian processes.

2. Hitting Times of New Maxima

Consider a Brownian motion on the real line, with diffusion constant D and with position at time t given by $B(t)$. From B, the maximum (or supremum) process, that corresponds to the running maximum of the Brownian motion at time t, is defined as $M(t) = \max\{B(\tau), 0 \leq \tau \leq t\}$. One of the difficulties when working with Brownian maxima is the fact that M is not a Markov process. In some cases, this difficulty can be circumvented by working with $M(t) - B(t)$ which is Markovian (and remarkably happens to be simply

[6] One just needs to take care of the "plateaux" in the supremum function, that lead to discontinuities in the reciprocal function.

a reflected Brownian motion (152)). In other cases, it can prove useful to rely on a perhaps lesser-known process, associated with the times when new maxima are achieved: $T(x) = \inf\{t \geq 0, B(t) = x\}$.

Lévy showed that T is Markovian and α-stable with $\alpha = 1/2$, since the characteristic function (which is just the Fourier transform of the law) of the random variable $T(x)$ has the form (152):

$$\mathbb{E}\left[e^{iuT(x)}\right] = e^{-(1-i\ \mathrm{sign}(u))x\sqrt{\frac{|u|}{2D}}}, \tag{44}$$

where $\mathbb{E}[\dots]$ denotes probabilistic expectation, and $\mathrm{sign}(u)$ is the sign of u. When the diffusion constant D is $1/2$, this process is called the standard stable subordinator of index $1/2$. Its propagator can be expressed easily (173) from the first-passage time density of Brownian motion, and one finds a transition kernel from s to $t > s$ in a "time" a given by:

$$\rho_D(a, t-s) = \frac{a}{\sqrt{4\pi D(t-s)^3}} e^{-\frac{a^2}{4D(t-s)}}. \tag{45}$$

(Recall that here a is the time parameter for the T process, while s and t correspond to positions. Also, note that this is a densitiy in t and therefore will be multiplied by dt upon integration so that, dimension-wise, there is nothing wrong with the denominator.)

Stable subordinators are an example of α-stable processes. These processes play an important role in line with the generalized central limit theorem (97): so-called Lévy flights, also known as Lévy-Brownian motions (88), have emerged as powerful tools to model non Gaussian phenomena (128; 180; 182; 199; 240). They are continuous time Markov processes that can be described in Langevin formalism as generalized Brownian motion, where the driving term, instead of a simple Gaussian white noise is a general Lévy noise ζ, that induces independent increments identically distributed according to an α-stable law (88): $\dot{x}(t) = \zeta(t)$. Such stable laws are specified by four parameters (7; 33): their stability index $\alpha \in]0, 2]$, a scale parameter σ, a location parameter μ, and a skewness parameter β. For $\alpha \neq 1$, the characteristic function of their position X at time t admits the following Lévy-Khintchine general form (143):

$$\mathbb{E}\left[e^{iuX(t)}\right] = e^{i\mu u - \sigma^\alpha t|u|^\alpha\left(1-i\beta\tan\left(\frac{\pi}{2}\alpha\right)\mathrm{sign}(u)\right)}. \tag{46}$$

Standard Brownian motion corresponds to $\alpha = 2$, the Cauchy-Lorentz process is $\alpha = 1, \beta = 0$, and the so-called Lévy-Smirnov process, with $\alpha = 1/2, \beta = 1$, is the same as the stable subordinator mentioned above.

3. Time-lag Process

Let now B_1 and B_2 be two independent Brownian walkers. Write m_1 and m_2 for their running maxima and write T_1 and T_2 for the associated hitting-time processes as defined in the previous subsection. We set $Z(x) = T_2(x) - T_1(x)$ and call Z the *time-lag process*: indeed, if $Z(x) = \tau$, it means that, compared to B_1, B_2 first hit x with a *delay* τ.

The important point here is that if m_1 and m_2 switch order, so will T_1 and T_2, and therefore Z will change signs. One can therefore compute the probability that m_1 and m_2 remain in the same order up to

time t by examining the probability that Z remains positive on the interval $[0, m_1(t)]$ (of course, one will need to pay attention to the fact that the right-hand side of this interval is itself a random variable).

Let us first determine the characteristic function of $Z(x)$, in terms of the diffusion constants D_1 and D_2. Using the independence of T_1 and T_2, one has:

$$\mathbb{E}\left[e^{iuZ(x)}\right] = \mathbb{E}\left[e^{iuT_2(x)}\right]\mathbb{E}\left[e^{-iuT_1(x)}\right]. \tag{47}$$

Substituting formula (44) for the characteristic functions on the right-hand side yields:

$$\mathbb{E}\left[e^{iuZ(x)}\right] = \exp\left\{-\left(1 - i\frac{\sqrt{D_1}-\sqrt{D_2}}{\sqrt{D_1}+\sqrt{D_2}}\ \mathrm{sign}(u)\right)\ \left(\frac{\sqrt{D_1}+\sqrt{D_2}}{\sqrt{2D_1D_2}}\right) x\sqrt{\mid u \mid}\right\}. \tag{48}$$

From this characteristic function and (46), one can readily identify an α-stable process with $\alpha = 1/2$, and asymmetry coefficient $\beta = \left(\sqrt{D_1} - \sqrt{D_2}\right) / \left(\sqrt{D_1} + \sqrt{D_2}\right)$. One immediately notices that when the diffusion constants are the same, β will be zero and therefore Z will simply be a symmetric Lévy flight.

We refer the reader to (190) for detailed formulae, let me simply mention here two main results:

- the survival probability, that is, here, the probability that the initial order (say, $m_1 > m_2$) is preserved up to time t, decreases asymptotically as $t^{-\theta(\beta)/2}$, where $\theta(\beta) = 1/2 - 2/\pi \arctan(\beta)$ (this is the persistence exponent for a 1/2-stable Lévy process with asymmetry parameter β (174)).

- when the two Brownian motions have the same diffusion constant, $\beta = 0$ and $\theta(\beta) = 1/2$. So that the "anomalous" 1/4-exponent computed in (22) emerges in fact from the underlying *standard* Sparre Andersen behaviour of the time-lag process.

Using hitting-time processes, I was thus able to establish a bijection between persistence exponents of non-Markov processes (maxima of Brownian motions) and persistence exponents of Markov processes (time-lag processes). After the publication of these results (190), a collaboration was initiated with Eli Ben-Naim and Paul Krapivsky to study not only persistence but also lead changes in the order of Brownian maxima.

4. Average number of lead changes

The survival probability of the initial order of two Brownian maxima may be viewed as the probability that there is not any lead change taking place between the two maxima, up to time t. Now we ask: what is the probability that there are n lead changes taking place? Or, if we cannot compute this distribution, can we at least compute its first moment, the average number of lead changes taking place up to a time t? In (23), we were able to compute this average number and also, thanks to a path transformation argument, to generalize our result to Lévy processes other than simple Brownian motion. I give below the main lines of the reasoning; all details can be found in (23).

As before, we denote by m_1 and m_2 the maxima of two independent, standard Brownian motion. If $m_1(t) = m > 0$, the probability that m_2 becomes the leader in the interval $[t, t + \Delta t]$ is given by the

probability that the first passage time of x_2 at level m is in $[t, t+\Delta t]$. Let us write $\rho(m,t)$ for the first-passage density, and $Q(m,t)$ for the density of the maximum. Then, integrating over m leads to the time derivative of the average number $n(t)$ of lead changes[7]

$$\frac{dn(t)}{dt} = 2\int_0^\infty \rho(m,t)\, Q(m,t)\, dm. \tag{50}$$

(The multiplicative factor 2 on the right-hand side takes into account that $m_2(t)$ could have been the leader.)

In the case of Brownian motion, the density of the maximum is of course known, as well as the first-passage density $\rho(m,t)$ (152; 197) (as recalled in the previous sections).

Now we stand in fact in a situation similar to that of Section I.A.3: the integral in Eq. 50 can be computed, in the Brownian case, because all densities are well known and tractable. However we wish to find a more general way of evaluating the integral in Eq. 50, without reference to the explicit form of Q or ρ, so that the result will hold for more general processes – not just for Brownian motion.

By construction, the integrand in Eq 50 corresponds to the product of probability densities associated with two independent paths: (i) a path first hitting m at time t, and (ii) a path having on $[0,t]$ a maximum equal to m attained at some time τ_*. We show in Figure 12 a transformation that puts in one-to-one correspondence pairs of paths (i) and (ii) with paths of duration $2t$ attaining their maximum at time t . More precisely with "one half" only of these, since the transformation produces only paths with non-positive final value. (We omit here the technical details of the path transformation.) When integrating over m the product $\rho(m,t)\, Q(m,t)$, we therefore obtain one half of the probability density associated with Brownian paths of duration $2t$ attaining their maximum at time t. This density is easily derived from the arcsine law (151; 152), which we already encountered in Section I.A.3.

Thus,

$$\int_0^\infty \rho(m,t)\, Q(m,t)\; dm = \frac{1}{2}\,\frac{1}{\pi\sqrt{t(2t-t)}} = \frac{1}{2\pi t}. \tag{51}$$

Combining Eq. 51 with Eq. 50, one obtains a very simple expression for the average number number of lead changes up to time t:

$$n(t) \sim \frac{1}{\pi\, \ln t}. \tag{52}$$

As was the case with path transformations in the previous sections, one main advantage is that the result may be generalized to processes other than Brownian motion.

---

[7] Note that, similarly to the average number of facets on convex hulls of Lévy-Brownian paths (see Section I.A.1), $n(t)$ may be seen formally as the expectation of a sum of indicator functions:

$$\text{“}\quad n(t) = \mathbb{E}\left[\int\limits_{\Delta t \le s \le t} \mathbb{I}_{A_{s,s+ds}}\right] \quad \text{”} \tag{49}$$

where $\mathbb{I}_{A_{s,s+ds}}$ is the indicator function of the event "a lead change occurs in the time interval $[s, s+ds]$". The time derivative of $n(t)$ is therefore indeed the same as the (infinitesimal) probability that a lead change occurs in $[t, t+dt]$.
(The starting point of the sum, Δt, is a cut-off on the initial instants, when infinitely many lead changes may occur – it plays the same role as the cut-off parameter when counting edges or facets on the convex hull of Lévy processes. We omit Δt here, or equivalently set it equal to 1. That is, implicitly, we focus on the large-t behaviour.)

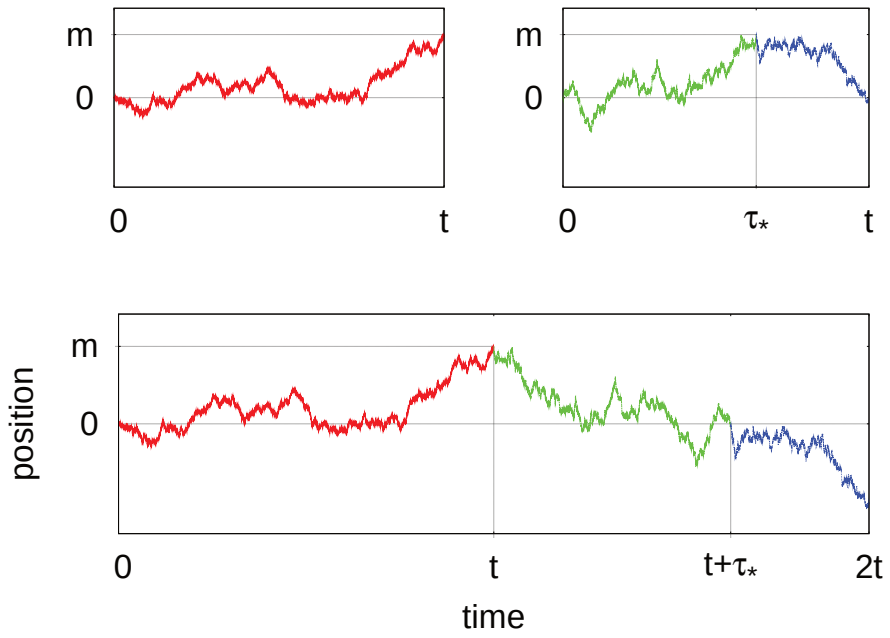


FIG. 12 Path transformation leading directly to Eq. 51: split the second Brownian path (top right) into its pre-maximum part and its post-maximum part (these two parts are independent by property of the time at which a Markov process attains its maximum); reverse time in the pre-maximum part (in green) and shift downward by $-m$ the post-maximum part (in blue); concatenating this with the other, independent Brownian path (top left), one obtains a new Brownian path of double duration (bottom). Note that its final value will be non-positive, and its maximum, also equal to m, will be attained at time t.

For symmetric α-stable processes, we denote by $Q_\alpha(m,t)$ the probability density of the maximum process and by $\rho_\alpha(m,t)$ the density of the first passage time at level m. Following the same reasoning as above, we write for the average rate at which lead changes take place:

$$\frac{dn}{dt} = 2\int_0^\infty \rho_\alpha(m,t)\, Q_\alpha(m,t)\, dm. \tag{53}$$

And we find that

$$2\int_0^\infty \rho_\alpha(m,t)\, Q_\alpha(m,t)\, dm = \frac{1}{\pi t}, \tag{54}$$

since the path transformation as well as the arcsine law are both valid for a wide class of Lévy processes (incl. symmetric α-stable processes). The key points are: cyclic exchangeability of the jumps (which is guaranteed when these are independent and identically distributed), a continuous and symmetric jump distribution, and a certain type of regularity for the supremum, see (33; 139). Fig. 13 illustrates the "universality" of the average number of lead changes.

5. Distribution of the number of lead changes

Let us now study the probability $f_n(t)$ to have exactly n lead changes until time t. We saw before that for two independent Brownian motion, the probability to have no lead change scales as $f_0 \sim t^{-1/4}$ (22; 190).

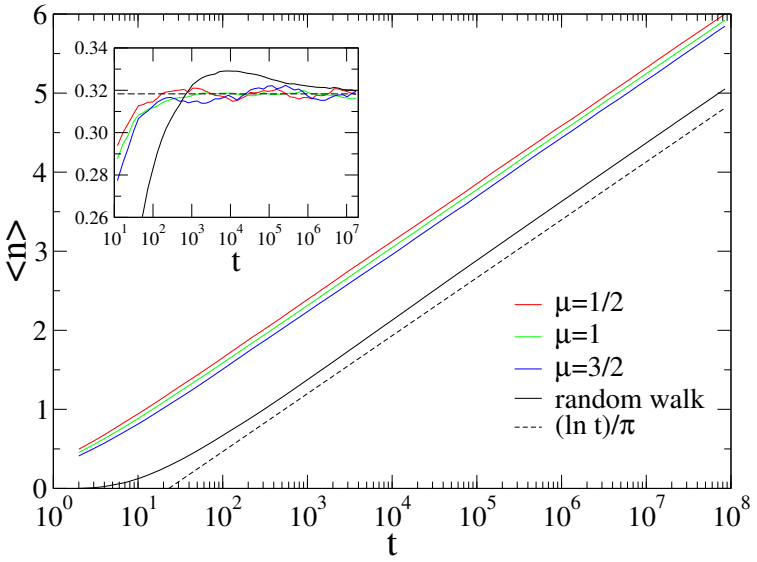


FIG. 13 The average number of lead changes $n(t)$ versus time t. Shown (top to bottom) are simulation results for two independent Lévy stable processes with stability index $\mu = 1/2, \mu = 1, \mu = 3/2$ and for standard (Brownian) random walks ($\mu > 2$). Inset: Plot of $d\ln n/dt$ illustrating that the amplitude A is universal and equal to $1/\pi$.

Since $f_n(t)$ can be interpreted as the probability that the $(n+1)$-th lead change takes place after time t, the probability density for the $(n+1)$-th lead change to occur at time t is given by $-df_n/dt$. In particular, from f_0, one can write down the conditional probability density g_s of the time of the next lead change, given that it takes place after time s:

$$g_s(\tau) = \frac{1}{f_0(s)}\left[-\frac{df_0(\tau)}{d\tau}\right] \sim s^{\frac{1}{4}}\,\tau^{-\frac{5}{4}}. \tag{55}$$

Now, $f_1(t)$ is the probability that a first lead change occurs at some time $s \in [0,t]$ and that the next lead change after s occurs at some time $\tau \geq t$. The probability density associated with such a configuration is thus simply proportional to

$$\left[-\frac{df_0}{ds}\right] \times \int_t^\infty g_s(\tau)\,d\tau \quad \sim \quad s^{-\frac{5}{4}} \times s^{\frac{1}{4}}\,t^{-\frac{1}{4}} \quad \sim \quad s^{-1}\,t^{-\frac{1}{4}}.$$

Integrating over s (we could again introduce a cut-off Δs, but this is akin to a discretization of time, so we might as well work in the large-t limit and set $s \in [1, t-1]$), one finds

$$f_1(t) \quad \sim \quad \int_1^{t-1} s^{-1}t^{-\frac{1}{4}}\,ds \quad \sim \quad t^{-\frac{1}{4}}\,\ln t.$$

Hence, there is a logarithmic enhancement of the probability to have one lead change compared with having no lead change. The above argument can be generalized to arbitrary n (by induction) to yield

$$f_n \quad \sim \quad t^{-\frac{1}{4}}\,[\ln t]^n\,. \tag{56}$$

Thus the logarithmic enhancement observed in f_1 is raised to the power n in f_n.

We probed the behavior of the cumulative distribution $F_n(t) = \sum_{0\leq k\leq n} f_k(t)$ using numerical simulations. The quantity $F_n(t)$ is the probability that the number of lead changes in the time interval $[0,t]$ does

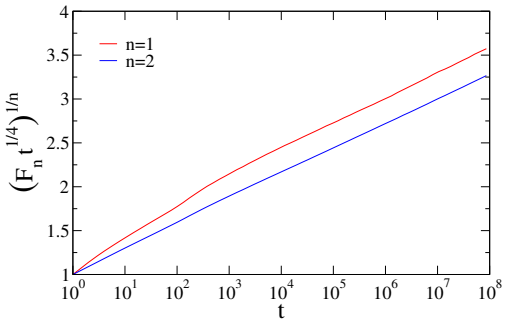


FIG. 14 The quantities $(F_n t^{\frac{1}{4}})^{\frac{1}{n}}$ versus time t for $n = 1$ and $n = 2$.

not exceed n. The dominant contribution to F_n is provided by f_n and hence

$$F_n(t) \quad \sim \quad t^{-\frac{1}{4}} \left[\ln t\right]^n . \tag{57}$$

Eq. (57) is confirmed numerically for $n = 1$ and $n = 2$ in Fig. 14.

The distribution of the number of lead changes in the case of Lévy symmetric stable processes can be established in a similar manner. Namely, one realizes that the derivation of (56) is fully transferable: one must only use the proper persistence exponent. Thus

$$f_n(t) \sim t^{-\beta(\alpha)} \left[\ln t\right]^n , \tag{58}$$

where $\beta(\alpha)$ is the persistence exponent characterizing the survival probability of the initial order of the maxima, for two independent symmetric stable processes with the same stability index α.

D. Residence Times

In the previous section, we have seen another example of path transformation at work, leading to explicit results for more general processes than Brownian motion. We now turn to questions pertaining to what mathematicians call *local times* and physicists *residence times*. This is a recent work (192), in collaboration with Sidney Redner (Santa Fe Institute, New Mexico). We retrieve some existing results and establish new ones using first-passage densities and generating functions.

Suppose that a diffusing particle in one dimension starts at $x_0 > 0$ and is absorbed, or equivalently, dies, when $x = 0$ is reached.

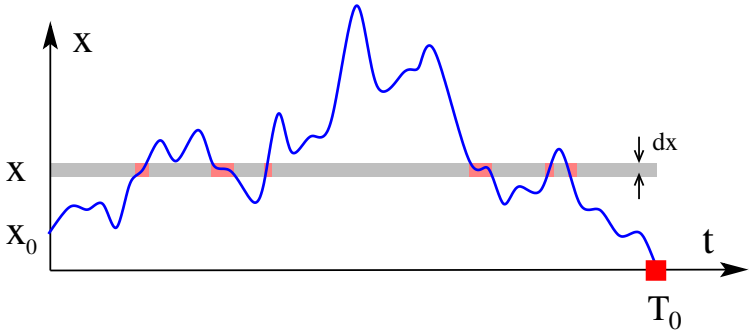


FIG. 15 Schematic trajectory of a diffusing particle in space time that starts at x_0 and is absorbed at time T_0 (square). The time $T(x)$ that the particle spends in $[x, x + dx]$ is indicated by the colored segments.

We are concerned here with the basic question (see Fig. 15): How much time does the particle spend in the range $[x, x + dx]$ before being absorbed? We term this latter quantity as the residence time. The properties of the residence time have been addressed in the mathematics literature by local-time theorems (138; 152; 173; 194) that specify the time that a Brownian particle spends in the region $[x, x + dx]$ before being absorbed when the origin is reached. When $x < x_0$, the distribution of this residence time was shown to be related to the distribution of the radial distance of a two-dimensional Brownian motion (138; 194). If the particle wanders in a finite domain with reflection at the domain boundary and absorption at a given point (or points) within the domain, the residence time at each site is related to the first-passage time to the absorbing set (27; 28; 71). This general formalism allows one to compute both the average residence time and the distribution of residence times at a given location.

While the consequences of local-time theorems are profound, the mathematical literature is sometimes presented in a style that is not readily accessible to the community of physicists who study random walks, and some of the results derived in Refs. (27; 28; 71) are extremely general in their formulation. In this work, we investigate residence-time phenomena for both continuum diffusion and the discrete random walk by using simple ideas and approaches from first-passage processes. We focus on cases where the particle is eventually absorbed at a specified boundary (*e.g.*, one specific side of an interval) and/or starts close to this boundary. Some of these situations have been treated previously in Ref. (3), by a more formal approach than that presented here.

1. Residence Time for Diffusion

Isotropic diffusion on the semi-infinite line

Consider a particle with diffusion coefficient D that starts at x_0 and is absorbed when $x = 0$ is reached. For such a particle, the image method gives the probability density for the particle to be at position $x > 0$ as (97; 197)

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}}\,\left[e^{-(x-x_0)^2/4Dt} - e^{-(x+x_0)^2/4Dt}\right]. \tag{59}$$

The average time $T(x)$ that the particle, which starts at x_0, spends in the range $[x, x+dx]$ before being absorbed at $x = 0$ (Fig. 15) is simply the integral of the probability density over all time, times dx (see Refs. (98; 99; 100; 172; 215) for related approaches). Performing this integral, with $P(x,t)$ from Eq. (59), the residence time $T(x)$ is given by

$$T(x) = dx\int_0^\infty dt\, P(x,t) = \begin{cases} \dfrac{x}{D}\,dx & x < x_0\,, \\[2ex] \dfrac{x_0}{D}\,dx & x > x_0\,. \end{cases} \tag{60}$$

To illustrate this result, we present simulations for a nearest-neighbor random walk that starts at: (a) $x_0 = 1$ and (b) $x_0 = 10$ in Fig. 16. As a function of the number of walks in the ensemble, $T(x)$ slowly converges to the asymptotic time-independent value in Eq. (60). A curious feature of this residence-time data is that it becomes erratic for large x, as shown in Figs. 16 (c) and (d). We can understand the origin of these large fluctuations by the following rough argument: for a diffusing particle that starts at x_0, the probability $S(t)$ that it survives until time t is $S(t) = \mathrm{erf}\big(x_0/\sqrt{4Dt}\big) \simeq x_0/\sqrt{4Dt}$ for $t \to \infty$ (197). For $\mathcal{M}$ random walks, we estimate the longest lived of them by the extreme-statistics criterion $S(t_{\max}) \simeq 1/\mathcal{M}$ (109; 146), which states that one out of $\mathcal{M}$ walks survives until at least time $t_{\max}$. This criterion gives $t_{\max} \simeq (\mathcal{M}x_0)^2/4D$. Correspondingly, the maximal range reached by an ensemble of $\mathcal{M}$ random walks is, roughly, $x_{\max} \sim \sqrt{Dt_{\max}} \sim \mathcal{M}x_0$. We now use this estimate to determine the large-x fluctuations in Figs. 16(c) and (d). To obtain an accuracy of, say, 10%, in $N(x)$, the number of times that the lattice site at x is visited by a random walk, we need roughly 100 walks to reach this value of x. Since $x_{\max}$ scales linearly in the number of realizations, roughly $m = 100$ walks will reach a distance $x = \mathcal{M}x_0/m$. For example, for 25000 walks starting at $x_0 = 1$, roughly 100 of them will reach $x = 250$. Thus up to $x \approx 250$, the variation in $N(x)$ should be smaller than 10%, and beyond this point fluctuations should become progressively more pronounced. This estimate is consistent with the data of Figs. 16 (c) and (d).

The approach given here can be readily extended to any situation where the spatial probability distribution can be computed explicitly. We now investigate three such cases: (a) biased diffusion, (b) diffusion constrained to remain in the interval $[0, L]$, and (c) diffusion exterior to an absorbing sphere in general spatial dimension d.

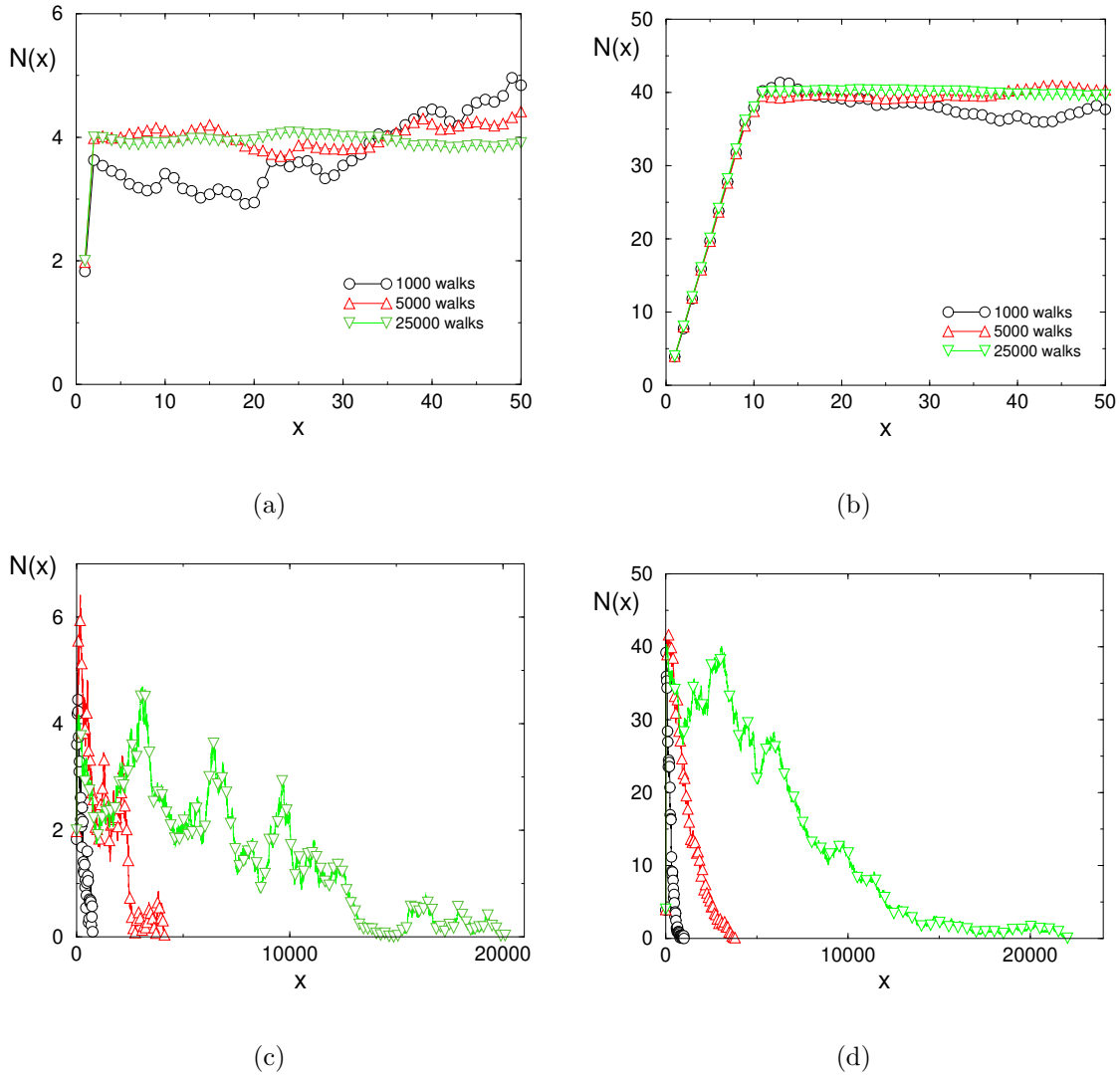


FIG. 16 Simulation results for $N(x)$, the average number of times that a random walk visits x when it starts at: (a) $x_0 = 1$, and (b) $x_0 = 10$. Here $N(x)$ is the discrete analog of the residence time $T(x)$. (c) & (d): The same data as in (a) and (b) over the full range of x.

Biased diffusion on the semi-infinite line

Suppose that a diffusing particle also experiences a constant bias velocity $-v$ that systematically pushes the particle towards the origin, so that the average time for the particle to reach the origin is finite. For a diffusing particle that starts at x_0, its probability density can be obtained by the image method (78; 197), and is given by

$$P(x,t) = \frac{1}{\sqrt{4\pi D t}} \left[e^{-(x-x_0+vt)^2/4Dt} - e^{-vx_0/D}\, e^{-(x+x_0+vt)^2/4Dt} \right] . \tag{61}$$

Notice that the magnitude of the image particle is different from that of the initial particle, while the velocities of the initial and image particles are the same.

We again integrate this expression over all time and obtain, for the average time that the particle spends

in $[x, x+dx]$ before it dies:

$$T(x) = \begin{cases} \dfrac{dx}{v}\left[1 - e^{-vx/D}\right] & x < x_0\,, \\ \dfrac{dx}{v}\, e^{-vx/D}\left[e^{vx_0/D} - 1\right] & x > x_0\,. \end{cases} \tag{62a}$$

For $v \to 0$, Eq. (60) is recovered, while in the opposite limit of $v \to \infty$, (62a) reduces to

$$T(x) \to \begin{cases} \dfrac{dx}{v} & x < x_0\,, \\ \dfrac{dx}{v}\, e^{-v(x-x_0)/D} & x > x_0\,. \end{cases} \tag{62b}$$

As one might expect, the time spent in $[x, x+dx]$ with $x < x_0$ is just that of a ballistic particle, while it is exponentially unlikely for the particle to reach the classically forbidden region $x > x_0$ for large Péclet number, vx/D.

Diffusion in a finite interval

Suppose that an isotropically diffusing particle is constrained to remain within the interval $[0, L]$ and is eventually absorbed at $x = 0$. We again want the time $T(x)$ that the particle spends in $[x, x+dx]$ before it dies. As in the previous two subsections, we need the spatial probability distribution for a diffusing particle with absorbing boundary conditions at 0 and at L. A straightforward computation of this distribution is unwieldy, as it involves either an infinite Fourier series or an infinite number of images.

However, we can avoid this complication by noticing that we only want the integral of the probability distribution over all time, which corresponds to its Laplace transform at Laplace variable $s = 0$. The Laplace transform satisfies $s\widetilde{P} - \delta(x - x_0) = D\widetilde{P}_{xx}$, where $\widetilde{P}$ denotes the Laplace transform and the subscript denotes partial differentiation. For $s = 0$, this reduces to the Laplace equation

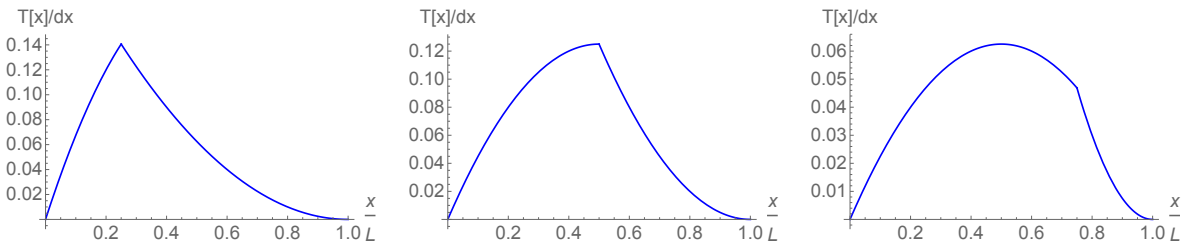
$$D\widetilde{P}_{xx} = -\delta(x - x_0)\,.$$

We solve this equation separately in the subdomains $x < x_0$ and $x > x_0$, impose the boundary conditions, continuity of the solution at $x = x_0$, and the joining condition $D\big(\widetilde{P}_x|_> - \widetilde{P}_x|_<\big) = -1$ to give, after standard steps,

$$\widetilde{P} = \frac{x_<}{D}\Big(1 - \frac{x_>}{L}\Big)\,, \tag{63}$$

where $\widetilde{P}_x|_>$ is the derivative just to the right of x_0 (and similarly for $\widetilde{P}_x|_<$), and $x_< = \min(x, x_0)$, $x_> = \max(x, x_0)$.

Finally, to obtain $T(x)$, we need to multiply the above distribution by the probability that the particle



(a) (b) (c)

FIG. 17 The average residence time density $T(x)/dx$ for a diffusing particle that is constrained to remain within the interval $[0, L]$ until it exits at $x = 0$. Shown are the cases: (a) $x_0 = 0.25L$, (b) $x_0 = 0.5L$ and (c) $x_0 = 0.75L$.

eventually exits the strip at $x = 0$, which is simply $1 - \frac{x}{L}$. Thus we have

$$T(x) = dx \int_0^\infty dt\, P(x,t) \left(1 - \frac{x}{L}\right) = \widetilde{P}(x, s=0) \left(1 - \frac{x}{L}\right) dx\,,$$

$$= \begin{cases} \dfrac{x}{D}\left(1 - \dfrac{x_0}{L}\right)\left(1 - \dfrac{x}{L}\right) dx & x < x_0\,, \\[2ex] \dfrac{x_0}{D}\left(1 - \dfrac{x}{L}\right)^2 dx & x > x_0\,. \end{cases} \tag{64}$$

The maximum residence time occurs at $x = x_0$ for $x_0 < L/2$ and then "sticks" at $x = L/2$ for $x_0 \geq L/2$, with a cusp always occurring at $x = x_0$ (Fig. 17). In the limit $L \to \infty$, we recover the result (60) for diffusion on the semi-infinite line.

Diffusion exterior to a sphere in dimension $d > 2$

We now determine the residence time for a diffusing particle that wanders in the region exterior to an absorbing sphere of radius a, a geometry that is the analog of the semi-infinite system in one dimension. Without loss of generality, we take the initial condition to be a spherical shell of unit total probability at radius r_0. We first treat the case of spatial dimensions $d > 2$ and then the special case of $d = 2$.

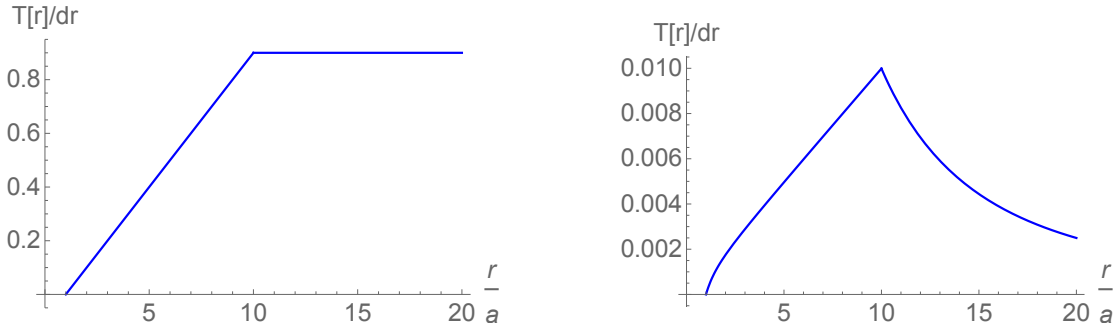
For general d, we need to solve the Laplace equation

$$D\nabla^2 \widetilde{P} = -\frac{1}{\Omega_d\, r_0^{d-1}}\, \delta(r - r_0)\,, \tag{65}$$

where Ω_d is the surface area of a d-dimensional unit sphere and r_0 is the radial coordinate of the starting point. Because of the spherically symmetric source term, angular coordinates are irrelevant. We therefore separately solve $\widetilde{P}'' + \frac{d-1}{r}\widetilde{P}' = 0$ in the subdomains $a < r < r_0$ and $r_0 < r$, and then impose the absorbing boundary condition at $r = a$ and the joining condition by integrating (65) over an infinitesimal interval than includes x_0. The result of these standard manipulations is

$$\widetilde{P}(r) = \frac{1}{D(2-d)\Omega_d}\left[\left(\frac{r_<}{a}\right)^{2-d} - 1\right] r_>^{2-d}\,. \tag{66}$$

To obtain $T(r)$, the average residence time in a shell of radius r and thickness dr, we again need to multiply the above expression by the probability that the particle eventually hits the sphere, which, for



(a) (b)

FIG. 18 The residence time density $T(r)/dr$ for a diffusing particle that starts at $r_0 = 10$ exterior to a sphere of radius 1 in: (a) $d = 3$ and (b) $d = 5$.

$d > 2$, is simply $(a/r)^{d-2}$ (197). Thus we have

$$\begin{aligned} T(r) &= \Omega_d\, r^{d-1} \widetilde{P}(r) \left(\frac{a}{r}\right)^{d-2} dr\,, \\ &= \begin{cases} \dfrac{dr}{D(d-2)} \left(\dfrac{a}{r}\right)^{d-2} \left[1 - \left(\dfrac{a}{r}\right)^{d-2}\right] \dfrac{r^{d-1}}{r_0^{d-2}} & r < r_0\,, \\[2ex] \dfrac{dr}{D(d-2)} \left(\dfrac{a}{r}\right)^{d-2} \left[1 - \left(\dfrac{a}{r_0}\right)^{d-2}\right] r & r > r_0\,. \end{cases} \end{aligned} \tag{67}$$

Two representative results are shown in Fig. 18. For large spatial dimension, a particle that eventually hits the sphere of radius a must do so quickly. Thus the residence time in the domain $r > r_0$ must necessarily be small, as shown in Fig. 18 (b) for $d = 5$.

In spatial dimension $d = 2$, the result analogous to Eq. (66) is

$$\widetilde{P}(r) = \frac{1}{2\pi D}\, \ln \frac{r_<}{a}\,. \tag{68}$$

Since a diffusing particle always reaches the absorbing sphere in $d = 2$, we immediately have

$$T(r) = 2\pi r \widetilde{P}(r)\, dr. \tag{69}$$

2. Visitation by a Discrete Random Walk

We now investigate the corresponding residence time for a symmetric random walk in the semi-infinite one-dimensional domain $[0, \infty]$. The walk starts at lattice site x_0 and is absorbed when it first reaches $x = 0$. The analog of the residence time is $N(x)$, the number of times that the random walk visits site x (excluding the initial visit if $x = x_0$) before the walk dies. We use the generating function approach to derive this quantity for the case of $x_0 = 1$.

Average number of revisits to $x = 1$

For a random walk that starts at $x_0 = 1$ and is absorbed at $x = 0$, the number of steps in the walk is necessarily odd. For convenience, we write this number as $2n+1$, with n an arbitrary non-negative integer. We define $A(n,k)$ as the number of random-walk paths that start at $x_0 = 1$, take the first step to the right (thus upward in the space-time representation of Fig. 19), and make k revisits to $x = 1$, before dying at the $(2n+1)$st step. The number of such paths was found in (35) and is given by

$$A(n,k) = \frac{k\,(2n-k-1)!}{(n-k)!\,n!}\,, \tag{70}$$

which happens to be directly related to the triangular Catalan numbers (12; 149). To compute the average number of revisits to $x = 1$, we will need $\mathbf{P}(k\,|\,n)$, the conditional probability for a path to make exactly k revisits to $x = 1$ before dying at step $2n+1$. This probability is

$$\mathbf{P}(k\,|\,n) = A(n,k)/C_n\,, \tag{71}$$

where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the nth Catalan number (216), which counts the total number of random walks of $2n$ steps that start at $x = 0$, remain in the region $x \geq 0$, and return to $x = 0$ at step $2n$. For what follows, we will also need

$$P(n) = C_n/2^{2n}\,, \tag{72}$$

the probability that a random walk starting at $x_0 = 1$ first hits 0 at step $2n+1$.

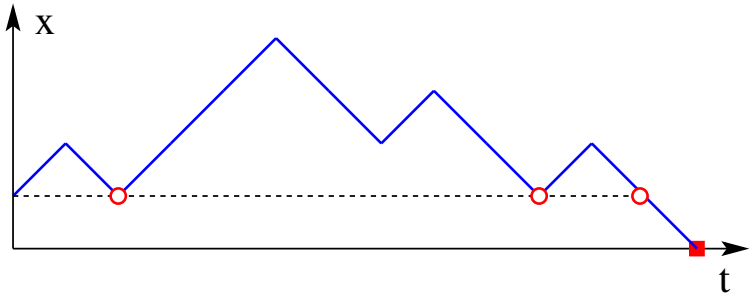


FIG. 19 Space-time trajectory of a one-dimensional random walk of $2n+1 = 13$ steps that starts at $x = 1$ and makes 3 revisits to $x = 1$ (red circles) before being absorbed at $x = 0$ (square). This path contributes to $A(6,3)$.

From Eqs. (70) and (71), we have

$$\mathbf{P}(k\,|\,n) = \frac{k\,(2n-k-1)!\,(n+1)!}{(n-k)!\,(2n)!} = k\binom{n+1}{k+1}\Big/\binom{2n}{k+1}\,. \tag{73}$$

Thus the number of revisits to $x = 1$, averaged over all walks of $2n+1$ steps is given by

$$\langle k\rangle_n = \sum_{k=1}^{\infty} k\,\mathbf{P}(k\,|\,n)\,. \tag{74}$$

Using expression (73) for $\mathbf{P}(k\,|\,n)$ in the above average, we obtain the remarkably simple result

$$\langle k\rangle_n = \sum_{k=1}^{\infty} k^2\binom{n+1}{k+1}\Big/\binom{2n}{k+1} = \frac{3n}{n+2}\,. \tag{75}$$

For long paths of $2n+1$ steps, there are, on average, 3 revisits to $x=1$, after which the walk immediately dies.

We now determine the number of revisits to $x=1$ upon also averaging over all n. This double average is

$$\langle\langle k\rangle\rangle = \sum_{n,k\geq 1} k\,\mathcal{P}(n,k)\,. \tag{76}$$

Here $\mathcal{P}(n,k)$ is the joint probability that the walk first reaches $x=0$ at step $2n+1$, *and* the walk makes k revisits to $x=1$ within $2n+1$ steps. This joint probability is

$$\mathcal{P}(n,k) = \mathbf{P}(k\,|\,n)\,P(n) = \frac{A(n,k)}{C_n}\,\frac{C_n}{2^{2n}} = \frac{A(n,k)}{2^{2n}}\,. \tag{77}$$

The average in (76) may now be expressed in terms of the generating function for $\mathcal{P}(n,k)$:

$$\begin{aligned} g(x,y) = \sum_{n,k\geq 1} \mathcal{P}(n,k)\,x^n\,y^k &= \sum_{n,k\geq 1} \frac{1}{2^{2n}}\,A(n,k)\,x^n\,y^k\,, \\ &= \sum_{n,k\geq 1} \frac{1}{2^{2n}}\,\frac{(2n-k-1)!\,k}{(n-k)!\,n!}\,x^n\,y^k\,, \\ &= \frac{xy}{2-xy+2\sqrt{1-x}}\,, \end{aligned} \tag{78}$$

Which was derived in (35) (see also (217)). In terms of the generating function, we immediately obtain the remarkably simple result

$$\langle\langle k\rangle\rangle = \sum_{n,k\geq 1} k\,\mathcal{P}(n,k)\,x^n\,y^k\,\Big|_{(1,1)} = y\,\frac{\partial g}{\partial y}\,\Big|_{(1,1)} = 2\,. \tag{79}$$

There are, on average, 2 revisits to $x=1$ in the ensemble of all random walks that start at $x=1$, take their first step to the right, and are eventually absorbed at $x=0$.

We can extend Eq. (75) to higher integer moments of the average number of revisits to $x=1$ for walks of $2n+1$ steps. The first few of these fixed-n moments are:

$$\begin{aligned} \langle k^2\rangle_n &= \frac{n(13n-1)}{(2+n)(3+n)}\,, \\ \langle k^3\rangle_n &= \frac{15n^2(5n-1)}{(2+n)(3+n)(4+n)}\,, \\ \langle k^4\rangle_n &= \frac{541n^2-196n^2+11n^2+4n}{(2+n)(3+n)(4+n)(5+n)}\,, \end{aligned} \tag{80a}$$

etc. We can similarly compute the higher integer moments of the number of revisits to $x=1$, averaged over all walk lengths, and the first few are:

$$\langle\langle k^2\rangle\rangle = 6 \qquad \langle\langle k^3\rangle\rangle = 26 \qquad \langle\langle k^4\rangle\rangle = 150 \qquad \langle\langle k^5\rangle\rangle = 1082\,, \tag{80b}$$

etc. Parenthetically, these numbers are also the first few ones in sequence A000629 in the On-Line Encyclopedia of Integer Sequences (209)

In the next paragraph, we will also need the generating function when the first step of the walk can equiprobably be to the right or to the left. This leads to the possibility that the total number of steps $2n+1=1$, i.e., $n=0$, for which the number of revisits to 1 equals zero. The generating function for the joint probability $\mathcal{P}(n,k)$ for this ensemble of random walks therefore is

$$G(x,y)=\sum_{n,k\geq 0}\mathcal{P}(n,k)\,x^n\,y^k=\tfrac{1}{2}\big[1+g(x,y)\big]\,,$$
$$=\frac{1}{2}\left(1+\frac{xy}{2-xy+2\sqrt{1-x}}\right)\,, \tag{81}$$

where the term $\frac{1}{2}$ comes from the walk that initially steps to the left and is immediately absorbed. Notice that $y\frac{\partial G}{\partial y}\big|_{(1,1)}=1$, which is consistent with Eq. 78: half of all paths die immediately upon the first step, and thus never return to 1, while the other half return twice, on average, as derived in Eq. 79.

Average number of visits to $x=2$

We now extend the above approach to determine the number of revisits to $x=2$ for a walk that starts at $x=1$ and is constrained to take its first step to the right. For this purpose, we define three random variables that characterize this set of walks:

- $2n+1$, the total number of steps in the walk when it dies;
- k, the number of visits to $x=2$ (*including* the first visit);
- ℓ, the number of excursions that lie above the level $x=1$.

Since an excursion is a path that lies between two successive returns to $x=1$ (and thus always remains above $x=1$), the minimal length excursion is the path $1\to 2\to 1$.

We want the ensemble average of the number of revisits to $x=2$. To facilitate this calculation, it is useful to define the three-variable generating function

$$\mathcal{G}(x,y,z)=\sum_{n\geq 1}\sum_{1\leq\ell\leq n}\sum_{\ell\leq k\leq n}\mathcal{P}\,(n,k,\ell)\;x^n\,y^k\,z^\ell\,, \tag{82}$$

which encodes all paths according to (n,k,ℓ). We also label each successive excursion of the path above $x=1$ by the index $1\leq i\leq\ell$, and we introduce the variables $2m_i$ and j_i, respectively, for the number of steps in the ith such excursion, and the number of returns to $x=2$ in this excursion (Fig. 20). As shown in this figure, $2m_i$ counts the number of steps that lie above $x=2$. Thus for an excursion that goes from $x=1$ to $x=2$ and immediately returns to $x=1$, $m_i=0$. In addition, j_i counts the number of *revisits* to 2, so that the total number of visits to $x=2$ in the ith excursion above $x=1$ is j_i+1. The variables j_i, m_i, and ℓ must satisfy the geometric constraints (see Fig. 20):

$$\begin{aligned} &j_1+j_2+\cdots+j_\ell+\ell=k\,,\\ &m_1+m_2+\cdots+m_l+\ell=n\,,\\ &j_i\leq m_i\,. \end{aligned} \tag{83}$$

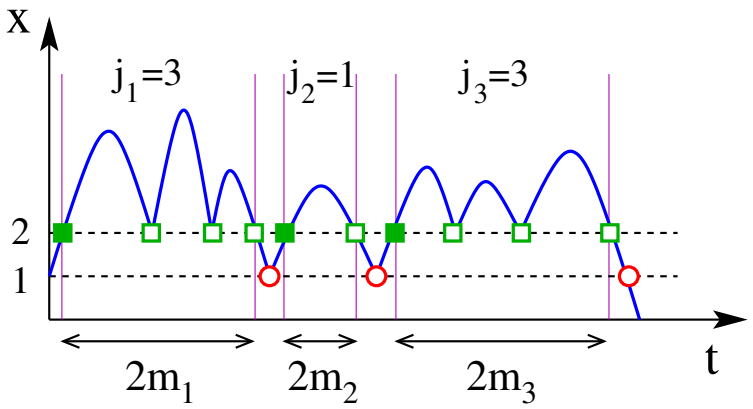


FIG. 20 Schematic space-time trajectory of a random walk that starts at $x = 1$ and has 3 revisits to $x = 1$ (red circles) and $k = 10$ revisits to $x = 2$ (green squares). There are $\ell = 3$ excursions above $x = 1$. Immediately after a revisit to $x = 1$, the next revisit to $x = 2$ is shown as a solid green square.

Using these definitions, the three-variable generating function $\mathcal{G}(x, y, z)$ can be functionally expressed in terms of $G(x, y)$ defined in Eq. 81 as (see the Appendix for details of this derivation)

$$\mathcal{G}(x, y, z) = \sum_{\ell \geq 1} P(\ell)\ [x\, y\, z\, G(x, y)]^{\ell}\ , \tag{84}$$

where $P(\ell)$ is the probability that there are ℓ excursions above $x = 1$ averaged over walks of any length, which is also the distribution of the number of returns to $x = 1$. Thus one may compute $P(\ell)$ as the marginal of the joint distribution of the number of steps and the number of excursions:

$$P(\ell) = \sum_{n \geq 1} \mathcal{P}(n, \ell) = \sum_{n \geq 1} \frac{1}{2^{2n}}\, A(n, \ell)\,. \tag{85}$$

The above sum starts at $n = 1$ because we are imposing the condition that the first step of the walk is to the right. Consequently the three-variable generating function in Eq. 84 will ultimately be expressed in terms of the restricted generating function g. Comparing the above formula with the first line of Eq. 78, we obtain

$$\begin{aligned}
\mathcal{G}(x, y, z) &= \sum_{\ell \geq 1} P(\ell)\, \big[x\, y\, z\, G(x, y)\big]^{\ell} \\
&= \sum_{\ell \geq 1} \sum_{n \geq 1} \frac{1}{2^{2n}}\, A(n, \ell)\, \big[x\, y\, z\, G(x, y)\big]^{\ell} \\
&= g\big(1, x\, y\, z\, G(x, y)\big)\,.
\end{aligned} \tag{86}$$

It is now straightforward to calculate of $\langle k \rangle$. From the definition of the generating function Eq. 82, we have

$$\begin{aligned}
\langle k \rangle &= y\, \frac{\partial \mathcal{G}}{\partial y}\Big|_{x=y=z=1}\ , \\
&= y\, \frac{\partial g}{\partial y}\Big|_{x=y=1} \left[G(1, 1) + y \frac{\partial G}{\partial y}\Big|_{x=y=1}\right]\ , \\
&= 2 \times (1 + 1)\ , \\
&= 4\,.
\end{aligned} \tag{87}$$

A random walk thus visits $x = 2$ twice as often as $x = 1$, as already predicted by the continuum solution (Eq. 60).

Average number of visits to $x > 2$

The ensemble average of the number of visits to a given level $x > 2$ may be readily computed by induction. We start by calculating the average number of visits to $x = 3$, and it will become apparent that this approach applies for any $x > 2$. Each time a random walk reaches $x = 2$, there are two possibilities at the next step: the walk may step forward to $x = 3$ or step back to $x = 1$, each with probability $\frac{1}{2}$. Let us first assume that the walk goes to $x = 3$, which occurs with probability $\frac{1}{2}$. Each time this event occurs, we now ask: what is the average number of visits to $x = 3$ (including this first visit) before the walk returns to $x = 2$?

With probability $\frac{1}{2}$, the walk may immediately return to $x = 2$, in which case, there is one visit to $x = 3$. On the other hand, if the walk steps to $x = 4$, we have the same situation as that discussed before for the number of revisits to 1. Namely, if we view $x = 3$ as the starting point, we know that there are 2 revisits to $x = 3$ and thus 3 visits to $x = 3$, on average, before the walk steps back to $x = 2$. Thus each time $x = 3$ is reached, there are

$$\left(\frac{1}{2} \times 1\right) + \left(\frac{1}{2} \times 3\right) = 2$$

two visits, on average, to $x = 3$.

For a walk that reaches $x = 2$, the average number of visits to $x = 3$ for this visit to $x = 2$ therefore is

$$\left(\frac{1}{2} \times 0\right) + \left(\frac{1}{2} \times 2\right) = 1\,.$$

The first term corresponds to the contribution from a walk that steps from $x = 2$ to $x = 1$ without hitting $x = 3$, and the second term is the contribution when the walk steps from $x = 2$ to $x = 3$.

To summarize, each time the walk visits $x = 2$, there is, on average, one visit to $x = 3$, before the walk is at $x = 2$ again. Clearly, this reasoning that determines the number of visits to $x + 1$ for each visit to x applies inductively for any level $x \geq 2$. Thus we conclude that the average number of times that a random walk visits a given level $x \geq 2$, equals 4, in agreement with the simulation results in Fig. 16(a). Clearly, our argument also applies for any starting point of the walk x_0, as long as we restrict to coordinates with $x > x_0 + 1$.

Time of the First Revisit

In addition to the *number* of revisits to $x = 1$ by a random walk excursion that starts at $x = 1$ and is eventually absorbed, we are interested in the *time* at which the first revisit occurs. This time characterizes the shape of the space-time trajectory of a random walk. Since the walk starts at $x = 1$ and ends at $x = 0$, its space-time shape is essentially that of a Brownian excursion — a Brownian trajectory that starts at $x = 0$, remains above $x = 0$ for all $0 < t < T$, and returns to $x = 0$ for the first time at $t = T$. The average shape of a Brownian excursion has been shown to be semi-circular (14; 69). From this shape, we might

anticipate that the first revisit to $x = 1$ is unlikely to occur for t near $T/2$ because such a revisit involves a large fluctuation from the average trajectory. Instead, it seems more likely that the first revisit to $x = 1$ will occur either near the beginning or the end of the excursion, a feature that evokes the famous arcsine laws (97; 152; 173). We now show that this expectation is correct.

Time of first return to 1 *for fixed walk length* n

Consider a random walk that starts at $x = 1$, takes its first step to the right, and is absorbed at $x = 0$ after $T = 2n + 1$ steps. What is the probability $P(2m \,|\, T)$ that such a walk revisits $x = 1$ *for the first time* at step $\tau_1 = 2m$? Since the walk necessarily revisits $x = 1$ at step $2n$ by definition, and the walk could revisit $x = 1$ immediately after 2 steps, m satisfies the constraint $1 \le m \le n$. The number of walks that revisit $x = 1$ after $2m$ steps may be obtained by decomposing the full path into two constituents (Fig. 21):

- Excursions of $(2m - 2)$ steps that wander in the domain $x \ge 2$ — the number of such paths is C_{m-1};
- Excursions of $(2n - 2m)$ steps that wander in the domain $x \ge 1$ — the number of such paths is C_{n-m}.

The first part accounts for the first return to $x = 1$ at step $2m$ and the second part accounts for the remaining path of $2n - 2m$ steps.

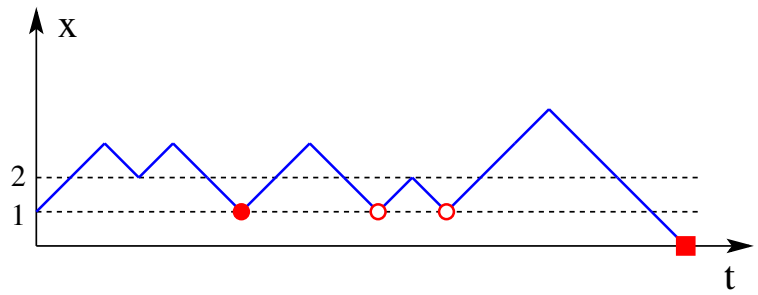


FIG. 21 Space-time trajectory of a one-dimensional random walk that starts at $x = 1$ and first revisits $x = 1$ at step 6 (solid circle). Subsequent revisits to $x = 1$ are indicated by open circles and the walk is absorbed when it first reaches $x = 0$ (square).

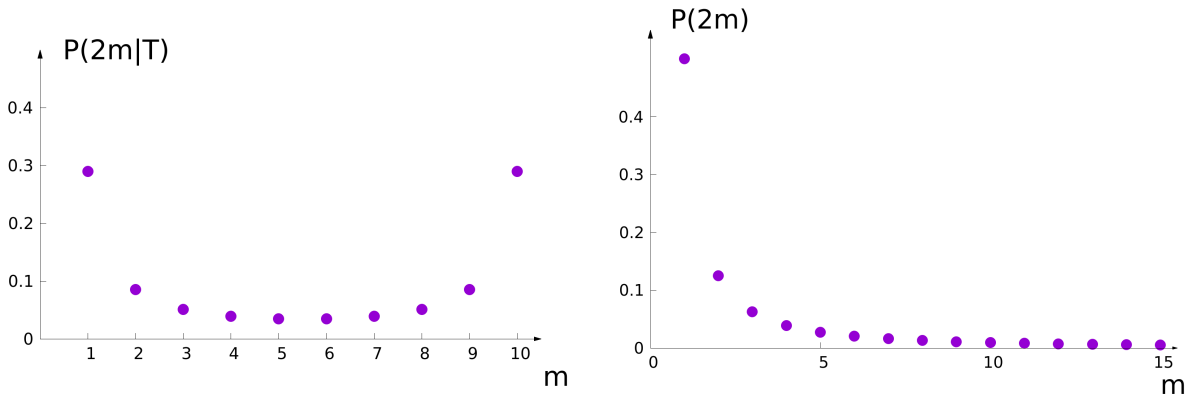
The required probability is then simply the product of these two numbers divided by the total number of walks that start at $x = 1$ and are absorbed after $2n + 1$ steps, which is C_n. Therefore

$$\mathbf{P}\,(2m \,|\, T) = \frac{C_{m-1}\, C_{n-m}}{C_n} = \frac{n+1}{m\,(n-m+1)} \frac{\binom{2m-2}{m-1}\binom{2n-2m}{n-m}}{\binom{2n}{n}}\,. \tag{88}$$

Conditioned on $T = 2n + 1$, the average value of τ_1, the time of the first revisit, can be immediately seen to be

$$\langle \tau_1 \rangle_n = n + 1, \tag{89}$$

because $\mathbf{P}\,(\tau_1 = 2m \,|\, T = 2n + 1)$ is symmetric under $m \to n + 1 - m$. This result may also be obtained by direct calculation of course. Because of the bimodal nature of the underlying probability distribution, the average value is very different from the typical value. The average corresponds to the minimum of the probability distribution (Fig. 22(a)), just as in the arcsine law for the time of the last zero of a Brownian motion.



(a) (b)

FIG. 22 (a) Conditional distribution $P(\tau_1 = 2m \,|\, T = 2n+1)$ for $n = 10$. Note that typical (i.e. most likely) values of τ_1 are 2 and $2n$, while the average value is $\langle \tau_1 \rangle = n + 1$. (b) Distribution $P(\tau_1 = 2m)$ of the first revisit time to 1.

Time of first return to 1 *for any* n

From the conditional probability $\mathbf{P}(2m \,|\, T)$, we may now compute the joint probability $\mathcal{P}(2m, T)$:

$$\begin{aligned} \mathcal{P}(2m, T) &= \mathbf{P}(2m \,|\, T)\; P(T) \\ &= \frac{C_{m-1}\, C_{n-m}}{C_n} \times \frac{C_n}{2^{2n}} = \frac{C_{m-1}\, C_{n-m}}{2^{2n}}. \end{aligned} \tag{90}$$

With this result, we can readily obtain the distribution $P(2m)$, the probability for a path of any length to perform an excursion of $2m$ steps that lies above $x = 2$ between steps 1 and $2m - 1$ (with the first step constrained to go from $x = 1$ to $x = 2$):

$$\begin{aligned} P(2m) &= \sum_{n \geq m} \mathcal{P}(2m, T) \\ &= C_{m-1} \sum_{n \geq m} \frac{C_{n-m}}{2^{2n}} = \frac{C_{m-1}}{2^{2m-1}}\,. \end{aligned} \tag{91}$$

This distribution is normalized, because $\sum_{m \geq 1} C_{m-1}/2^{2m-1} = 1$. The Markovian nature of the random walk means that there is no memory between what happens after step $2m$ and the probability that the walk first revisits $x = 1$ at step $2m$. Hence $P(2m)$ is simply the probability that a symmetric random walk first returns to its starting point at step $2m$, which asymptotically scales as $m^{-3/2}$ (97). Because of this scaling, the average time for the first return, $\langle\langle \tau_1 \rangle\rangle = \sum_m 2m\, P(2m)$, is infinite, even though $P(2m)$ is peaked at $m = 1$ (Fig. 22(b)).

3. Comments

The main qualitative feature of the average residence time at a given point is that it vanishes (often linearly) as the distance between this point and the absorber. That is, a diffusing particle simply does not

linger when it is close to an absorbing point. Another interesting feature is the fact that, in low dimensions ($d \leq 3$), the average residence time at any point beyond the starting point is constant: it is simply equal to the average residence time at the starting point. This is no longer the case for $d \geq 4$. It would be of interest to understand why this transition differs from the well known transition between recurrence and transience, which happens between $d = 2$ and $d = 3$.

For the discrete random walk, we found that the first revisit to $x = 1$ occurs near the start or the end of the path. This means that it is very unlikely that there will be a large deviation toward the boundary and away from the average position of the path near the middle of an excursion. This suggests that an individual Brownian excursion always remains close to its average shape. We hope to investigate this behavior in future work, with a view to obtaining a full characterization of the fluctuations around an excursion's average semi-circular shape.

We illustrated in this work on residence times how simple tools may be put to fruitful use in the derivation of explicit results on random walks or Brownian motion. This approach complements that presented in previous sections, on random convex hulls, persistence exponents and lead changes.

Let me now turn to a seemingly totally different part of my research work in the past few years, on the analysis and modelling of urban segregation phenomena – where we will still occasionally encounter random walks and other stochastic artefacts.

II. SEGREGATION IN THE CITY

I developed an interest in urban segregation problems through both a taste for social and geographical questions and a taste for multi-agent models, of which Thomas Schelling's segregation model is a prominent example (203; 204; 205). For a theoretical physicist, the analogy between the Schelling model and interacting particle systems (such as the Ising model) is very appealing indeed.

Somewhat unoriginally therefore, I first started working on segregation phenomena through the Schelling model. The idea was to introduce further complexity by allowing agents to change groups over the course of the dynamics. We studied this new model with Aurélien Hazan, and we realized that introducing switching agents was akin to exerting a form of random control on the system. Antoine Lucquiaud, whose PhD thesis I co-supervise, is now working on the theoretical aspects of this problem: is it possible to "tame" a complex system via coordinated random interventions?

While working on this model, the question (and the desire!) of confronting it with real data quickly arose. Most papers on the Schelling model remain purely theoretical (with, at best, *in silico* simulations) and do not tackle the issue of working with real data. There is at least one good reason for this: the Schelling model is not *per se* a representation of real processes at work in a real city. It is of course very simple and abstract compared to a real city. But its great achievement is to answer the following question: is a high level of intolerance a *necessary* condition for segregation to emerge? And the answer is "no" – because we can conceive of at least one context (the Schelling model), albeit an abstract one, in which agents are reasonably tolerant and, still, an initially well-mixed system is driven into segregation.
Despite this great achievement of Schelling's model, one may wish to confront oneself with real data and bridge the gap (or at least a tiny part of it) between abstract models of segregation and real-world segregation phenomena (much in the spirit of (153)). The primary resource is then data: how to obtain it, and most importantly, how to process and analyze it in order to reveal and capture complex real-world segregation phenomena. Two projects that I have been co-developing in the past couple of years represent efforts to answer these questions. In particular, we – that is, together with my statistician colleague Madalina Olteanu, with PhD candidate Antoine Lucquiaud, and with UCLA Geography Professor William Clark – have been devising multiscalar trajectories built from urban data. These encode spatial dissimilarities at all scales and allow for both statistical analysis and stochastic modelling. We are currently implementing this new method on European data, for the European Commission (we were granted access to this new EU data through a call for projects). We are also implementing it on Californian data, thanks to the collaboration with UCLA.

Another branch of this data-based research on segregation has led us to use neural networks and specifically the Kohonen algorithm to obtain finer descriptions of neighbourhood effects. Part of this work is now being developed in collaboration with a *start-up* company, and will lead to a joint PhD supervision (within the *CIFRE* framework).

A. A Modified Schelling Model

Since Schelling's and Sakoda's papers (203; 204) introducing stochastic modeling on a checkerboard to gain insights about segregation phenomena in urban environments, much effort has been devoted to develop and understand multi-agent systems subject to Schelling dynamics. The existence of similarities with certain types of spin-system and liquid-solid dynamics (110; 111; 169; 175; 218; 221; 229) has been instrumental both in drawing theoretical physicists' attention to segregation phenomena and in providing the basis for advanced analogies with existing results and methods from the statistical physics toolbox. Many variants of Schelling's initial dynamics have been found to exhibit similar properties, and the main characteristic, namely the emergence of segregation, was shown to be a universal property common to a vast class of systems (116; 201).

We investigate here the impact of introducing, within a Schelling-class dynamics with two types of agents, a given fraction f of spatially-fixed agents able to switch, that is: to change spontaneously from one of the two types of agents to the other, rather than simply bearing the same type throughout the simulation. This amounts to imposing a noise or a perturbation in the form of a spatially-fixed random background within the Schelling dynamics and we depart here from the standard context where agent types usually correspond to ethnic groups that cannot change. Let us emphasize that the dynamics studied here is *a priori* different from Glauber and Kawasaki dynamics as discussed in (218), since we follow Schelling and impose a non-zero vacancy density and we do not allow for direct site-exchange between agents. Our model can thus be viewed as an interpolation between a (trivial) dilute system of non interacting particles in zero external field (at $f = 1$, all agents are fixed) and a magnetic system of moving and interacting particles with zero total magnetization (at $f = 0$, agents follow Schelling's preferential-choice dynamics and there is an equal number of agents of both types).

Our motivation is twofold: on the theoretical side, the complexification introduced here differs from changes in the agents' interaction energies (or utility functions in socio-economic contexts) that have been explored in the literature, and we would like to see how it modifies the system's phase diagram. On a more heuristic side, the question is whether the presence of switching agents tends to facilitate mixing and de-segregation, and the aim is to explore a new variant of residential dynamics, with a long-term view to identifying factors that have the logical, mathematical capacity to lead to desegregation or lesser segregation, just as "social distance and preference dynamics do have the logical capacity to combine with socioeconomic inequality between groups to create relatively high levels of ethnic and class segregation" (65).

The new specification we suggest may be interpreted in at least two ways. First, the agents able to switch, being spatially fixed, can be viewed as housing sites that are never empty and for which the landlords (eg government authorities or housing associations) enforce an allocation policy that is blind to agents' types. In such a case, the system should be considered as "open" in the sense that when an agent switches types, it is as if an agent of the former type had moved outside of the system and an agent of the new type had moved in (the total number of agents in the system is fixed, but not that of agents of a given type). Secondly, if

we favour a "closed"-system view, switching can be interpreted as "social mobility". The easiest may be to picture agent types as social groups defined *eg* by income levels. In this case, there can be various reasons why agents switching from one social group to the other may remain fixed spatially. Indeed, an agent that switches from the higher-status type to the lower one may wish to cling to her former status, as embodied by the majority type in her neighbourhood. If the switching occurs the other way around, so that, say, a newly well-off agent persists in living in a poorer neighbourhood, it may be that she is willing to be an active proponent of social mixity and/or she is careful about what will become of her current income status in the future. Along with this social-mobility interpretation of type-switching, let us remark that ethnic types may also be considered here, as these can change, be it through the agents' own feelings or the normative action of some externally imposed new definition.

For the sake of simplicity, we will henceforth refer to the switching dynamics as "type mobility", as opposed to the "spatial mobility" with preferential choice encoded in the standard Schelling dynamics. Let us however stress that the first interpretation (random allocation in an open system) should be equally kept in mind, in particular as it is compatible with open systems and points to a form of intervention that could be no less an actual mean of action in real urban contexts as the promotion of type mobility (especially if type mobility is to be social mobility in both directions!). Therefore, exploring random allocation strategies for housing sites can be viewed as an extension to the urban segregation problem of strategies for efficiency improvement through random noise that have gained attention lately, be it in relationship with democratic representation procedures (184), hierarchical organizations (41; 185) or central bank interventions on financial markets (40).

We shall start from a most simple instance of a Schelling-class system, which meets the general criteria set out in (201). A noticeable feature of the Schelling dynamics in our system is that the moves are "blind" ones, in the sense that when given the opportunity to move, an agent, if she does indeed move, will go to a site picked uniformly at random among vacant sites. Agents are therefore not satisfiers, let alone maximizers — but this is not a pre-requisite for a system to belong to the general class defined in (201), as we understand it.

1. System and notations

Ours is a square-lattice system, featuring a finite number $width \times height$ of sites. At any instant, each site is in one of four states:

1. vacant,
2. occupied by an agent of (pure) type A,
3. occupied by an agent of (pure) type B,
4. occupied by an agent of type C,

TABLE I Simulation parameters

Parameter	Value
ρ	$\in [0.9, 1]$
τ	0.3
f	$\in [0, 1]$
T_M	500
N_E	50
$width, height$	30
p_u	0.2
p_h	10^{-4}
p_s	0.05

where A and B are the main two types, and C corresponds to agents that have the ability to display either type A or type B. There is a fixed number of agents of each three types (these numbers are determined by the occupation density ρ and the fraction of C-agents f, given that there is an equal number of agents of pure type A and pure type B).

We use periodic boundary conditions and discrete-time, ordered, asynchronous updating (73). Agents are therefore examined in turn ; if they are of types A and B, they follow a Schelling dynamics, *viz.* they move to a vacant site (chosen uniformly at random) with probability p_h if they are satisfied with their current position and with probability p_u if they are not. The satisfaction of an agent of type A or B is either 0 or 1, depending on the proportion of their direct neighbours being of the opposite type: if this proportion is smaller than a given parameter τ (usually interpreted as the degree of "tolerance" of the agents (204)) the agent is satisfied, otherwise she is not. Agents of type C display at each instant one of the two types A or B, but do not move with the Schelling dynamics: when examined, they switch their displayed type with probability p_s, regardless of their neighbours' types. The number of C-type agents is expressed as a fraction f of the total number of agents present in the system. Simulation code and results are available online.[8]

Table I summarizes the simulation parameters, with the typical values which they shall be given hereafter, unless otherwise stated. ρ, τ, f, p_h, p_u and p_s were defined above. T_M is the number of time steps in a single realization (one time step corresponds to a complete round over all agents), and N_E is the number of ensemble realizations. *width* and *height* are the dimensions in number of sites of the simulation grid. The values of ρ and τ are chosen so that the pure Schelling dynamics ($f = 0$) leads to a segregated state. (Note that in the basic Schelling-type dynamics we are using here, only direct neighbours are considered, which implies that values of τ tend to be distinguished only relative to the discrete values: $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$; however,

[8] `https://sourceforge.net/p/phase-py/`

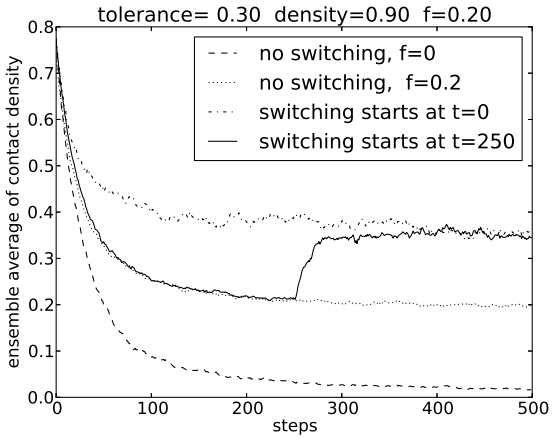


FIG. 23 Ensemble average $\langle x(t)\rangle$ of the contact density in four simulation modes.

since vacant sites are not counted as neighbours, other values of τ can occasionally be distinguished.) The probability of moving for an agent unhappy with her current neighbourhood is set to a value significantly lower than 1 to reflect difficulties (*eg* economic factors) that may prevent unsatisfied agents from moving. The rate p_s of switching for C-agents is then chosen to be four times smaller than p_u to implement distinct timescales for the two types of dynamics (of course a reverse order of these timescales should also be studied later). Finally, the fact that agents who are statisfied with their current location are able to move ($p_h \neq 0$), even at a very small rate, suppresses certain simulation artifacts (see *eg* (218)) and can also be interpreted as reflecting the possibility that agents make a "wrong" move due to misinformation or, again, economic factors.

As recalled in (111; 201; 218), Schelling-class systems are expected to reach some kind of equilibrium, or at least a steady state as far as segregation is concerned. This translates into the reaching of an almost constant limit value for a variable quantifying segregation within the system. Various choices of variables are possible. We shall follow (201) and work with a contact density $x(t)$, defined as the average over all occupied sites at time t, of the ratio between the number of neighbours of the opposite type and the total number of neighbours (vacant neighbouring sites are not counted). The contact density is normalized by one half, so that when the two populations A and B are well mixed in the system, $x(t)$ is close to 1, whereas it is close to 0 when the system exhibits segregation. We shall write $\langle \dots \rangle$ for averages over all ensemble realizations.

2. General results

Let us first examine the influence of a non-zero fraction of switching agents on the limit value $x_\infty = \lim \langle x(t)\rangle$ of the contact density (time-averaged on the time-steps after a steady state has been reached), with four simulation modes:

1. no switching agents are present ($f = 0$);

2. switching agents are present ($f = 0.2$) but do not switch ($p_s = 0$);

3. a given fraction $f = 0.2$ of switching agents participate, being "activated" only after a fixed delay (250 time steps);

4. a given fraction $f = 0.2$ of switching agents participate, and are "activated" at the beginnning of the simulation.

The first case is clear: this is a standard instance of a Schelling system.

For the second and third cases, note that C-type agents may be "de-activated", i.e. p_s may be set to 0 so that C-agents keep the same displayed type throughout the simulation. They thus act as fixed, non-moving agents of one or the other of the two types A and B.[9] Therefore, even though they are not themselves moving, they do participate in the dynamics by influencing the satisfaction of their neighbours, and by "persisting" in their residential choice whatever the current composition of their neighbourhood.

The fourth case simply corresponds to an activation of C-type agents from the beginning of the simulation.

The occupation density is set to 0.9, the tolerance τ to 0.3. Starting from a well-mixed state (drawn from a uniform distribution), one observes that:

1. in the absence of switching agents, the system quickly relaxes to a segregated state with a value of x_∞ very close to 0, as expected for the chosen values of ρ and τ (knowing the phase diagram of a typical Schelling system as in (111));

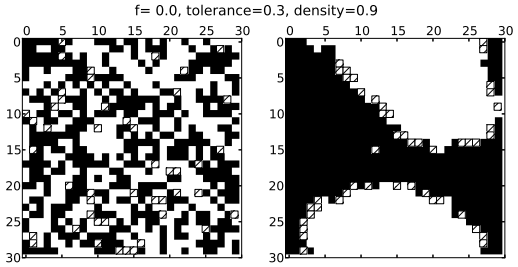
2. in the presence of switching agents that persist as fixed A or B agents, the relaxation is slower, and leads to a value of x_∞ very close to 0.2;

3. in the presence of switching agents that start as fixed A or B agents for 250 time steps, relaxation starts similarly to the previous case; however, once activation takes place at $t = 250$, the ensemble average $\langle x(t) \rangle$ makes a sudden jump to reach a higher limit value close to 0.35;

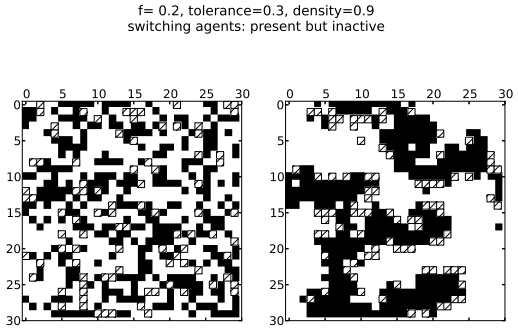
4. in the full dynamics, i.e. with C-type agents activated from $t = 0$, the system relaxes toward a moderate value of x_∞ close to 0.35.

The introduction of switching agents leads to a higher contact density, as could be intuitively expected. In the second variant, when C-agents are de-activated, the limit contact density x_∞ is already significantly larger than in the pure Schelling variant. Actually it coincides with the value of f in this particular case. This hints at the fact that it may simply be a background level of contact density due to local, point

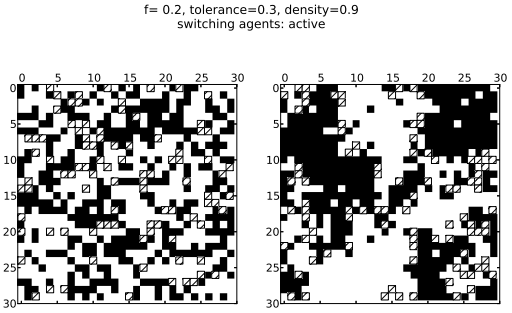
[9] Indeed in the second variant, C-agents stay at the same site and with the same displayed type for ever. Links with so called extremists in opinion propagation models (81) could be explored.



(a) Sample initial and final configurations in the absence of switching agents ($f = 0$)



(b) Sample initial and final configurations in the presence of *inactive* switching agents ($f = 0.2$)



(c) Sample initial and final configurations in the presence of *active* switching agents ($f = 0.2$)

FIG. 24 Sample initial and final configurations in the absence (a) or presence of active (c) or inactive (b) switching agents. Initial configurations are on the left-hand side. (Black = A agents, white = B agents, hatched = empty site, C agents appear according to their current displayed type.) Final configurations show strict phase separation in the absence of switching agents, a mixture of patches with fuzzy interfaces in the presence of active switching agents and an intermediate pattern in the presence of inactive switching agents.

mixity around fixed, persisting C-agents that accounts for a higher value of x_∞ – rather than *bona fide* mixity across the whole system. Visual observation of patterns in the final, steady state of our numerical simulations (Fig. 24) shows that this type of effect is not the only thing happening in the presence of unactivated switching agents. Compared to the complete phase separation observed in the absence of switching agents, the final configurations observed in the presence of persisting agents is an intermediate step towards the more complex pattern (mixture of patches and fuzzy interfaces) formed in the presence of activated switching agents.[10]

The activation of switching, at time $t = 250$ in the third variant and time $t = 0$ in the fourth one, entices indeed a significant de-segregation phenomenon, leading quickly, in both cases, to a limit value $x_\infty \in [0.3, 0.4]$. Such a value for the contact density can correspond to a mixture of phases, as described in (119): eg two "pure" clusters separated by a well-mixed area, or a mixture of homogeneous patches (see Fig. 24).

Notice that such patchy mixtures are a way of "optimizing" the *average* neighbourhood (but not necessarily the *typical* neighbourhood) in a residential system at fixed $\tau < \frac{1}{2}$: indeed, a notable feature of Schelling systems is that, when they lie in the segregated phase, their average equilibrium contact density is significantly less than the agents' tolerance. This is precisely one of the features that made Schelling's model famous, since one could naively expect that when agents "tolerate" up to one half of neighbours to be of the type opposite to theirs, the system would in the long run reach some kind of steady state in which each agent had a neighbourhood comprising of one half of agents from the other type. But that was firstly assuming that *average* neighbourhood and *typical* neighbourhood would coincide, and secondly overlooking the possibility of a phase transition occuring at some value of the tolerance — which is the case (130; 204; 205).

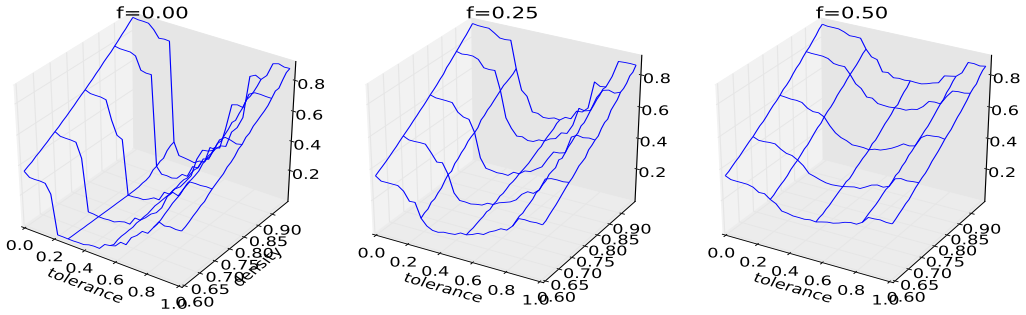
3. Phase diagram and transitions

We wish now to look at the transitions occuring in the system both when $f = 0$ and ρ, τ vary (where we should retrieve at least part of the results described in (111)), and when f increases. Of course, our main interest lies in the latter case, and the relative bareness of our Schelling dynamics (with only direct neighbours taken into account and moves being blind) is not suited for an in-depth study of the system's phase diagram.

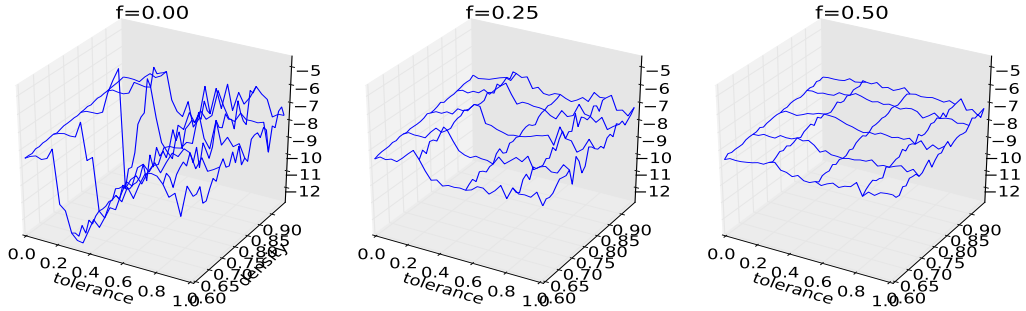
To facilitate further analysis of the system's behaviour and identify the transitions, we shall follow (111) and introduce, on top of the ensemble-averaged contact density $\langle x(t) \rangle$ and its limit x_∞, a quantity analogous to the susceptibility of thermodynamical systems, and another analogous to the specific heat. The former is defined as

$$\chi(t) = \frac{\langle [x(t)]^2 \rangle - \langle x(t) \rangle^2}{\tau},$$

[10] Note that in the presence of persisting agents, any form of phase separation is constrained by the fixed (and random) spatial pattern of persisting agents.



(a) Contact density



(b) Analogue of susceptibility (log of)

FIG. 25 Contact density x_∞ and analogue of susceptibility χ_∞, in terms of ρ and τ as f increases.

and the latter as

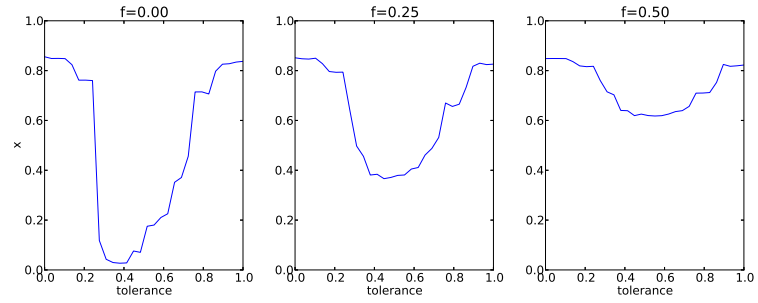
$$C(t) = \langle [E(t)]^2 \rangle - \langle E(t) \rangle^2,$$

where E is defined in analogy with the energy in the Blume-Emery-Griffiths spin model (42):

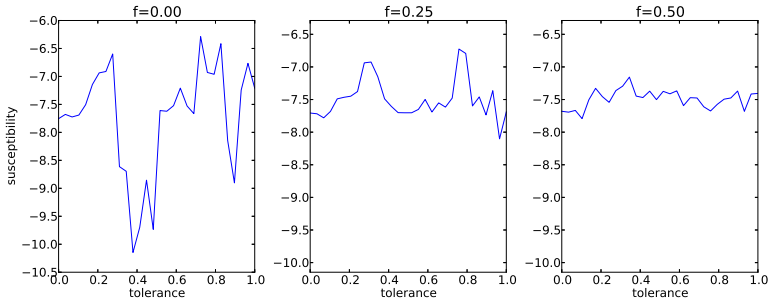
$$E = -\sum_{i,j} c_i c_j - (2\tau - 1) \sum_{i,j} c_i^2 c_j^2,$$

the sums being over pairs of neighbours, and $c_i = 1$ for a site occupied by an agent of (displayed) type A, -1 for an agent of (displayed) type B and 0 for a vacant site. We shall write χ_∞ and C_∞ for the time-averaged steady-state values of $\chi(t)$ and $C(t)$.

Gauvin *et al.* have argued and explored the validity of such an analogy with thermodynamical spin systems (110; 111). We shall follow their lead in trying to identify transitions by looking for rapid changes in the value of x_∞ accompanied by peaks in χ_∞ and/or C_∞.



(a) Contact density



(b) Analogue of susceptibility (log of)

FIG. 26 Profile of the contact density x_∞ (a) and the analogue of the susceptibility χ_∞ (b) as τ varies, for a fixed occupation density $\rho = 0.95$.

4. Transitions at $f = 0$

We look here briefly at the $(f = 0)$ -section of the phase diagram, which corresponds to the first picture on the left in the top part of Fig. 25 for the (f, ρ, τ)-diagram. Of course, our system is a much simplified version of the Schelling system studied in (111), and in particular we expect changes mainly to occur at a couple of discrete values of τ: $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$. Fig. 25 shows that, at high occupation densities, our simulation results are indeed compatible with phase transitions occuring at $\tau = 0.25$ and $\tau = 0.75$. This is also supported by the behaviour of the susceptibility, as seen in the first figure on the left in the bottom part of Fig. 25.

The change in the value of the contact density at $\tau = 0.25$ seems more abrupt than the one at $\tau = 0.75$, as indicated also by the profile of $x_\infty(\tau)$ at fixed $\rho = 0.95$ (leftmost plot in Fig. 26 (a)). Both transitions are marked by peaks of the susceptibility (leftmost plot in Fig. 26 (b)).

In (111), at low vacancy densities, the authors identified a first-order phase transition around $\tau_c = 0.75$ and a "frozen" state below $\tau_f = 0.4$. In between these values, their system would eventually reach a state of low contact density, that is, a segregated state. The transition from a state of frozen dynamics to the segregated phase was further investigated in (200) and found to correspond to a jamming transition that could be reproduced *via* a patch model and described by deterministic equations. In our system, the

low-vacancy-density, low-tolerance state is separated from the segregated state by a more abrupt (possibly first-order) transition. This state can be understood as follows: at low τ (that is essentially $\tau = 0$ as soon as $\tau < 0.25$), all agent are unsatisfied, unless they have only neighbours of the same type as theirs. Therefore, given the opportunity to move, they will, because moving does not depend on their finding a more "welcoming" site (contrary to what is the case in (111), leading to the dynamics freezing). Eventually, blind moves lead to a reasonably well-mixed state, for reasons that differ from those behind the frozen, well-mixed state in (111) where the system simply remains in its initial state, which happens to be drawn from a well-mixed distribution.

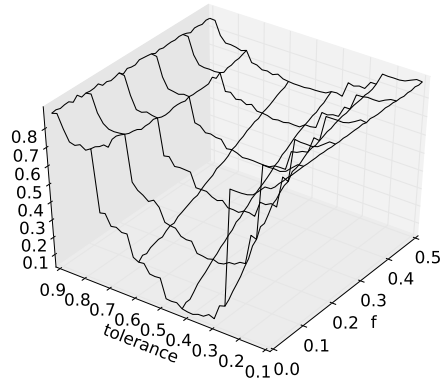
Thus, the general shape of the phase diagram at $f = 0$ and high occupation density is very much comparable to that of a classical Schelling system as in (111), and we proceed to the case $f \neq 0$ in the next paragraph.

5. Transitions along f

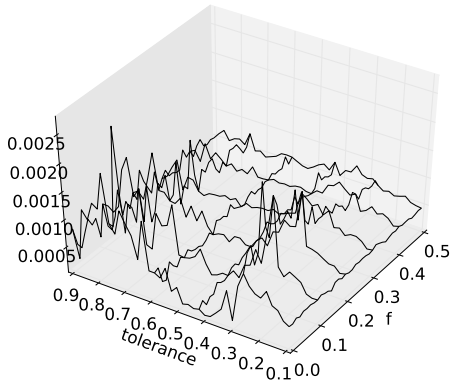
We examine now cross sections of the phase diagram at constant ρ, letting f, the fraction of switching agents present in the system, increase. Intuitively, if f is sufficiently large, one expects a well-mixed system, as the situation then becomes that of fixed agents switching with equal probabilities to the A or B types. This is what one can observe in the upper part of Fig. 27, where indeed the mixed phase (high contact density) gains more and more ground as f increases. As we have seen in section II.A.3, when $f = 0$ two transitions are noticed, one near $\tau = 0.25$ and one near $\tau = 0.75$. Given that at large values of f, only a high contact-density phase survives, the corresponding transition lines in the f-τ plane should vanish or bend toward each other, jutting into the well-mixed phase and enclosing a segregated phase. This is precisely what one observes in the simulations, as pictured in the upper part of Fig. 27.

The nature of the transition occuring across these lines, at different points, is of particular interest, all the more as it varies with f: at small values of f, one has the transitions described in the previous subsection. They are compatible with thermodynamical transitions, marked by peaks in the analogues of susceptibility and specific heat (lower parts of Fig. 27). Near higher values of f, the transition lines are no longer parallel to the f axis, and thus can be crossed *via* an increase in f. For instance, at fixed occupation density $\rho = 0.95$ and tolerance $\tau = 0.3$, Fig. 28 shows the profiles of the contact density, the analogue of susceptibility, and the analogue of specific heat as the fraction of switching agents f is increased from 0 to 1. The change in the contact density is not abrupt but peaks can be observed in χ_∞ and C_∞ around $f \approx 0.25$, with a finite-size effect (15; 38) shown in Fig. 29.

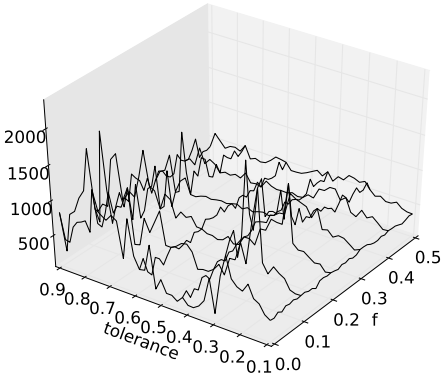
Further up in the phase diagram shown in the top part of Fig. 27, near $f \approx 0.5$, a more gradual change from a low to a high contact density is observed at intermediate values of the tolerance τ when f is increased. No peaks are to be noted here in the susceptibility or specific heat analogues. This can indicate that the nature of the transition has changed along the line, from discontinuous to a smooth, gradual change.



(a) Contact density



(b) Analogue of susceptibility



(c) Analogue of specific heat

FIG. 27 Phase diagram in terms of f and τ, at constant $\rho = 0.95$. Contact density x_∞ (a), analogue of susceptibility χ_∞ (b), and analogue of specific heat C_∞ (c).

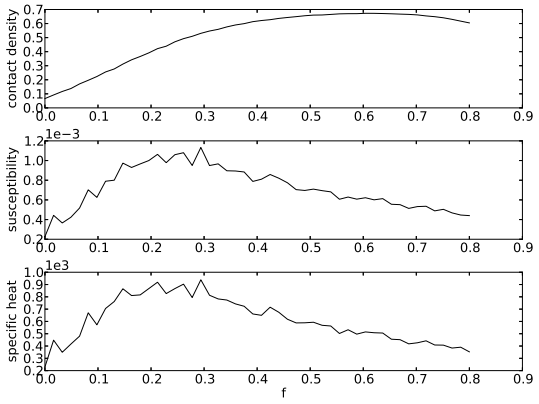


FIG. 28 Transition along f, at constant $\rho = 0.95$ and $\tau = 0.3$. Contact density x_∞ and the analogues of susceptibility χ_∞ and specific heat C_∞ are plotted as the fraction f of switching agents increases.

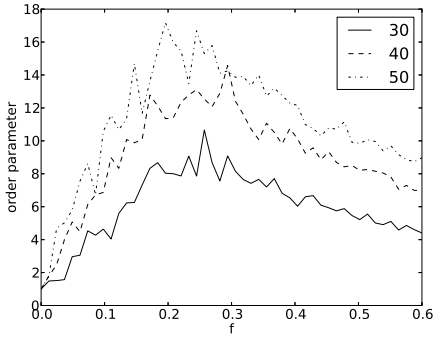


FIG. 29 Finite-size effect. The analogue of specific heat C_∞, normalized by its initial value, is plotted for various grid sizes, at $\rho = 0.95$ and $\tau = 0.3$ (with an ensemble of realizations other than the one used for Fig. 28). The larger the system, the more marked is the peak, as one would expect in thermodynamical phase transitions.

6. Comments

We have examined the numerical behaviour of a system driven by a mixture of two dynamics: (i) a standard Schelling preference dynamics where agents of two types move stochastically across a grid according to their preferred composition of neighbourhood — which can be interpreted as residential choice, and (ii) a dynamics where some agents stay fixed but have the ability to switch from one type to the other, which can be interpreted either as "type" mobility or as random allocation for given sites (without the possibility that these remain vacant).

Ongoing theoretical work (esp. Antoine Lucquiaud's PhD thesis) aims at offering a model for this mixture of dynamics, with the ability to predict – quantitatively – the observed behaviour.

Numerous variants are also possible, of course, but one will find useful leads in the confrontation with real data (24; 49; 105; 119; 140) and in the development of interdisciplinary work with sociologists, geographers and urban-system scientists. We turn to more data-based approaches in the following sections.

B. Multiscalar Trajectory Convergence Analysis

Segregation is often simply perceived as spatial separation of two or more groups, and therefore measured in terms of the relative proportions of each group in the different neighbourhoods of a city, as we tipycally did in the previous section. However, segregation is essentially a spatial and multiscalar phenomenon, as pointed out for instance by (148; 153; 179). An individual perceives segregation all the more acutely as she has to go *a longer way* from her home to discover what the city in its entirety might look like. Imagine the extreme case of a city where two groups A and B live in total separation, thus forming two ghettos. An individual living at the heart of one of the ghettos would have to explore the whole city to find out that it actually comprises equal proportions of both groups. On the contrary, starting from some parts of the city that are *better mixed*, for instance on the boundary between the two ghettos, there would be no need to cover so large an area to come to the same realization. This basic observation has led us to imagine a mathematical framework that allows to capture and measure spatial dissimilarities as a multiscalar phenomenon across the city.

Consider ever larger neighbourhoods around a starting point or areal unit. As detailed in the first subsection below, enlarging the area all the way up to the whole city produces for each point in the city a trajectory. That is, the sequence of values taken by the variable under consideration from its value at the most local level around the starting point to its average value at the metropolitan level. For a given starting point, how "long" it takes for the sequence to reach the city's average thus indicates how "distant" the point is from the city's global distribution.

Further, the same mapping of points into trajectories allows to easily characterize a null model. In a perfectly mixed city (the "uniform" city), the aggregation process used to build the trajectories would be equivalent to adding units drawn at random (as when drawing balls in an urn without replacement). We introduce in the second subsection below a way of quantifying deviations from such random sequences, based on the Ballot theorem already encountered in the first part of this thesis. We illustrate our ideas and methods with public data available for the city of Paris.

1. From bespoke neighbourhoods to trajectories

Geographical and demographical data is often available at a given spatial level – for example census blocks in the USA or IRIS (*Îlots regroupés pour l'information statistique*) in France, see (226) and (126). Aggregating such blocks poses a number of statistical questions, the most famous of which is maybe the so-called "modifiable areal unit problem", as stated in (177; 178). Also, in most instances, elementary spatial units will exhibit dissimilarities. Indeed, if the data gives the number of medical practitioners, or the quantiles of the income distribution within the unit, then units will generally differ from one another (114). Such dissimilarities may or may not present spatial patterns. If they do, the geographical system is said to exhibit spatial segregation, especially if the variables considered correspond to the relative proportions of

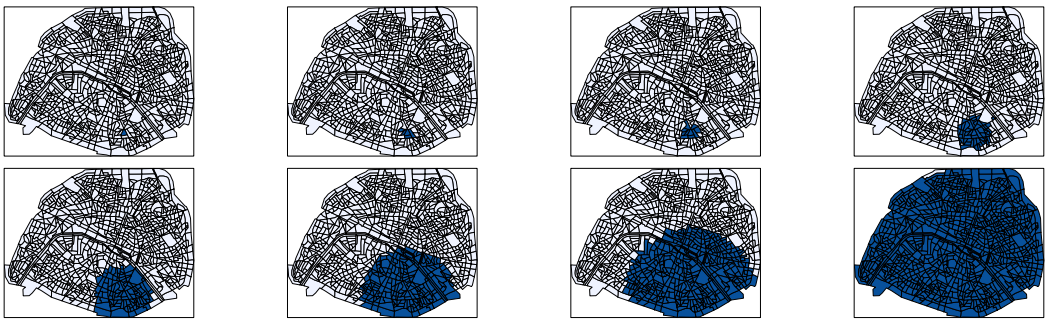


FIG. 30 Some of the steps in a sequence of bespoke neighbourhoods centered around one of Paris's statistical units (so-called IRIS). The first point in the sequence corresponds to the unit alone (top left). A second unit is then aggregated to the first one, according to some prescribed rule – here for the sake of simplicity, we choose the nearest neighbouring unit in terms of distance from centroid to centroid. The procedure is iterated until eventually one has aggregated all the city's units around the first one. This yields, for each starting point, a sequence of bespoke neighbourhoods, representing all scales, from the most local one available in the data to the metropolitan one.

different population groups: (48; 51; 74; 75; 76; 86; 94; 95; 96; 106; 127; 165; 195; 206; 234).

Modifiable areal units may actually prove to be a valuable tool to quantify segregation and spatial patterns of dissimilarities, see (9; 67; 118; 131; 147; 148; 164). Instead of only comparing individual units, one may gain greater insight into relative spatial differences by taking into account the broader picture: that is, by considering ever larger neighbourhoods around a unit, as if through a zoom lens – see Figure 30. Evaluating a given statistical variable at each step of the aggregation procedure produces a sequence of values that reflects not only the initial point but also the singularity of its position within the city. The sequence of bespoke neighbourhoods around a starting point acts as a sequence of filters: unveiling step by step the city's "face". Indeed the final value of all trajectories, whatever the starting point, will be the city's average for the variable considered. But details of a given trajectory, such as the number of aggregating steps needed to converge to the city's average characterize how different from the whole city the starting point is – how singular it is in the whole city.

Although relatively easy to describe, the actual computing of trajectories presents a number of difficulties from a statistical point of view. We recall some of them in the following paragraph.

2. Defining and computing trajectories

Consider a grid comprising N fundamental spatial units (typically, individuals or the smallest statistical units available). For a given unit, define $G_n(i)$, the cluster formed by itself and its $n-1$ nearest neighbouring units. Thus, $G_1(i)$ is just the i-th unit on its own, while $G_N(i)$ is the whole grid (whichever i from which one starts).

Now suppose one has a statistical variable ξ defined on each element of

$$\mathcal{G} = \{G_n(i),\, 1 \leq i, n \leq N\}$$

Note that $\mathcal{G}$ is simply the set of all bespoke neighbourhoods, with all possible starting points in the city. Then, define for each starting point i a function f_i such that $f_i(n)$ gives the value of ξ computed on $G_n(i)$.

Interpreting n as an index, one has for each starting unit i a sequence representing the trajectory that takes ξ from its value on unit i to its value on the whole grid. Formally, $(f_i(n))_{1\leq n\leq N}$ is what we call the trajectory of i for the variable ξ.
Once trajectories have been built from the data available in a given city, a statistical analysis may be carried out in order to:

- identify units that exhibit similar trajectories – this may be done for instance using clustering algorithms;

- detect the crossing over between two regimes: the local one and the global one – in the former, ξ may take values significantly distinct from the one on the whole grid; in the latter, it takes values very similar to the city's average.

Let us emphasize links between the first point above and multiscalar approaches used in recent papers, *eg* (5; 67; 179). In particular, (5) use vertical slices or cross-sections of the trajectories defined here, i.e. values at certain points only. But looking at full trajectories allows us to examine the second point listed above: the number of aggregated neighbouring blocks needed around a starting block to get close enough to the city's average value for the variable under consideration. We shall call this the "*radius of convergence*". Geographically this plays the role of a local urban radius: it is a proxy for the area one needs to explore locally to obtain a reasonably good perception of the city as a whole. We give examples of such analysis further in this section.

3. Constructing trajectories from actual data

One of the statistical subtleties arises upon building datasets so that they are defined on $\mathcal{G}$, the full set of all possible bespoke neighbourhoods in the city.

Consider indeed the following standard situation: data is available in the form of a size factor s_i (*eg* a number of inhabitants) for each spatial unit, as well as the value of a possibly multi-dimensional variable ξ. This could be the number of offices or services of such and such type available in the unit, the social housing rate, an average income, a local density of public transportation, quantiles of a distribution, a full distribution...

When grouping n units, one needs to compute the value of the variable on the new aggregate. This is done easily in the case when ξ is simply a number:

$$\xi(G_n(i)) = \sum_{j\in G_n(i)} \xi(j). \tag{92}$$

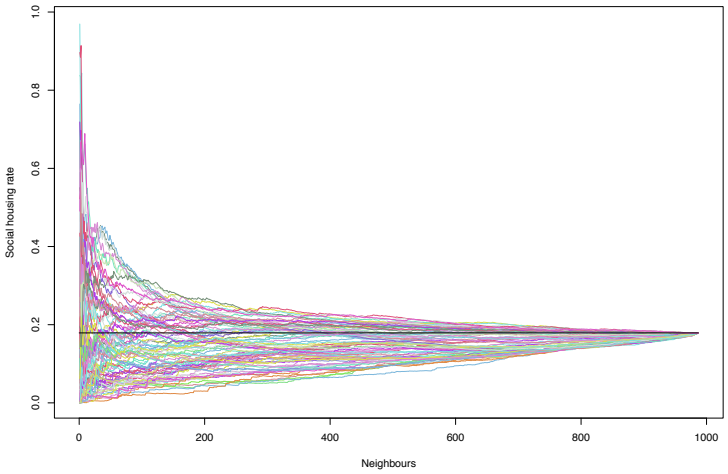


FIG. 31 Trajectories for the social housing rate, starting from some (10%) of the 936 statistical units (so-called IRIS) in Paris. The solid line corresponds to the city's average (17.9%).

Similarly, $\xi(G_n(i))$ is readily computed when $\xi(i)$ is a rate or an average:

$$\xi(G_n(i)) = \frac{\sum_{j\in G_n(i)} s_j \xi(j)}{\sum_{j\in G_n(i)} s_j} \,, \tag{93}$$

with s_i the population of unit i (or another relevant size factor).

When ξ is a distribution that is known entirely (*eg* one knows the income of every single household in any of the units), one simply aggregates the individual datasets to obtain the dataset for a group of spatial units, and the corresponding empirical distribution is obtained readily.

A more difficult case, alas very frequent, is that when ξ retains only certain percentiles of a distribution. For instance:

$$\xi(i) = \{2354; 4684; 6546; 8138; 10542; 13622; 17058; 22202; 30862\} \,, \tag{94}$$

are the deciles of the income distribution (in euros per year) for a given IRIS (census block) in the northern part of Paris. In this case, one is led to either (or both) rely on an Ansatz for the shape of the underlying distribution (*eg* assume that it is log-normal, or exponential or other) or simulate the full dataset (with assumptions on the intradecile distributions), see (50; 107; 112; 113; 129; 219). An example is given in the next subsection, as we are faced with this problem when we consider income data for the city of Paris, available only in the form of quantiles for each census block.

Lastly, one does not have to be working in the discrete framework of statistical blocks or population count data. For certain variables, such as availability of public transport networks, one may work directly in terms of a local density. For instance, let M be the total number of metro stations in Paris. Let $m_i(l)$ be the number of stations in a disk of radius l centered on IRIS i, $p_i(l)$ the population living in the same disk, and P the total population in Paris. Then define

$$r_i(l) = \frac{m_i(l)/M}{p_i(l)/P} = \frac{m_i(l)/p_i(l)}{M/P}. \tag{95}$$

This is similar to the "representation" defined for social classes in (153). It quantifies whether the local density of metro stations per inhabitant is smaller ($r_i(l) < 1$), larger ($r_i(l) > 1$) or equal ($r_i(l) = 1$) to the city's average density. And it allows to build trajectories, for each IRIS, letting l vary from some small value ε to the full radius of the city.

4. Example: spatial dissimilarities in Paris

As an illustration, let us work with two types of data available for the city of Paris from France's census bureau, called INSEE ("*Institut national de la statistique et des études économiques*").

Social housing rate.

For each IRIS, the number of housing units and the number, among these, of social housing ones are available. One is then in a situation where computing the social housing rate for any group of IRIS is easy, and the trajectories can be computed for all IRIS in Paris. Some of these trajectories are shown in Figure 31.

One observes groups of trajectories that tend to start higher or lower than the metropolitan average, and converge to it more or less quickly, obviously with a strong spatial dependency as far as the initial units are concerned.

Let us define the radius of convergence, that is, the point where each path enters (and remains) into a given interval (here ± 0.05) around the city's average social housing rate. Looking at these radii reveals different scales of convergence to the global mean, from one district to another and, inside each district, from one IRIS to another, as can be seen in Figure 32: IRIS blocks from Paris's 11th district converge much "faster" than those in the 8th district, for instance.

One may then classify IRIS blocks according to their radii of convergence to the city's mean, and represent them accordingly on a geographical map of Paris (Figure 33). First, note that boundary effects are clearly not predominant as peripheral western and north-eastern parts of the city, for instance, do not exhibit the same radii of convergence to the city's average. In fact, boundaries, by forcing the aggregation of IRIS blocks closer to the city centre and beyond (rather than neighbouring IRIS blocks just outside the city) tend to smooth patterns rather than exacerbate them. A particularly interesting feature revealed by our method is that the western (W) part of the city is "further away" from the whole city than the north-eastern (NE) part: trajectories for IRIS in the W part need 2 to 4 times as many aggregation steps to converge to the city's average than IRIS in the NE part. Both parts correspond to extreme points for the variable in question, with a concentration of social housing in the NE part and a substantially lower than average rate in the W part (local social housing rates are also visible in Figure 33). Starting from these extremal points above and below the city's average, with symmetrical geographical position in the city, the W and NE parts could have had similar radii of convergence to the city's mean. However, this is clearly not the case, thus revealing a higher level of singularity, as far as social housing is concerned, in the W part of the city than in the NE part.

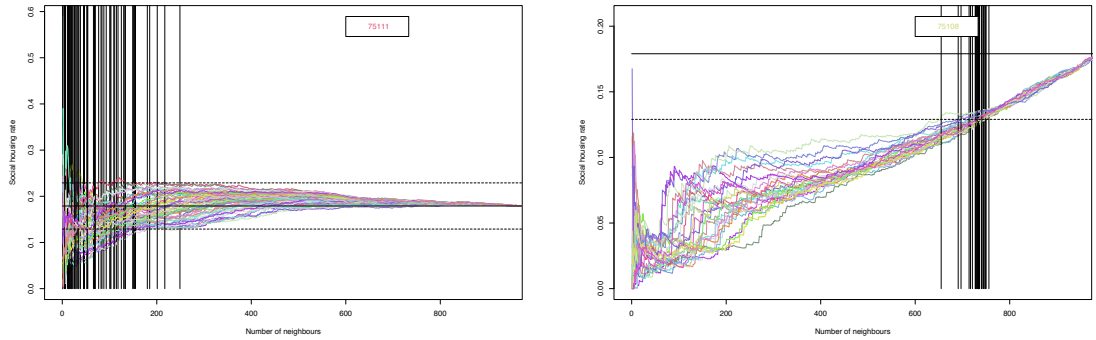


FIG. 32 Trajectories for the social housing rate, starting from each IRIS (census block) in Paris's 11th (left) and 8th (right) districts. Colours correspond to each different IRIS taken as starting point. The solid flat line gives the city's average social housing rate (17.9%), the dashed lines correspond to ±0.05 around this average. Solid vertical lines correspond to radii of convergence: the radius of convergence of a trajectory is defined as the point where the path last enters the ±0.05-interval (and therefore remains inside the interval afterwards). One observes that IRIS in the 11th district converge much faster to the city's average than IRIS in the 8th district.

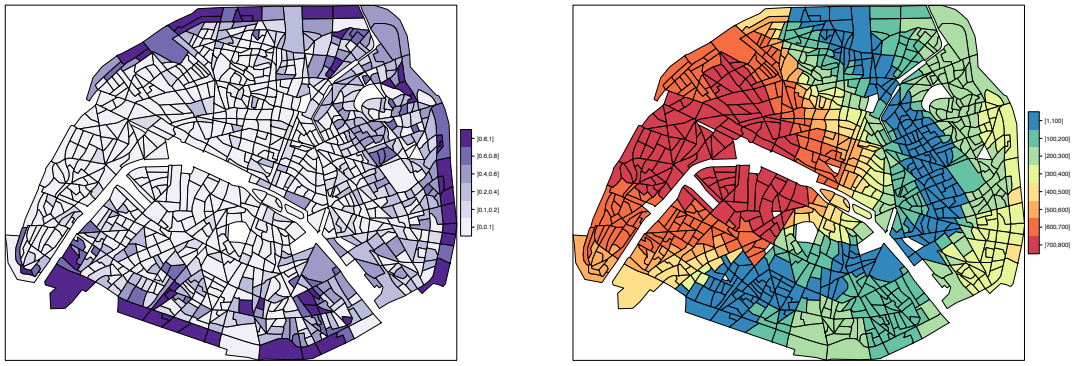


FIG. 33 IRIS blocks coloured according to their radius of convergence to the global average social housing rate (right). Radii are expressed in terms of the number of aggregation steps. For comparison, the map on the left-hand side shows IRIS areas coloured according to the local rate of social housing. (White areas correspond to parks, riverbanks and other IRIS where data is not available.) Observe that IRIS in the peripheral western part (W) and the peripheral eastern and north-eastern part (NE) of Paris present similar geographical positions and correspond to extreme values for the social housing rate (very high in the NE part, very low in the W part). Nonetheless, IRIS in the W part converge much slower to the city's mean, revealing a higher level of seclusion.

Conversely, one observes a quasi ring of IRIS blocks with relatively short scales of convergence to the city's mean. These are the historically socially mixed areas along the *boulevards*, the former *faubourgs* that used to be just outside the city's walls before these were transformed into *boulevards*. Our method thus reveals a lasting imprint visible in terms of distance to the city's average for a variable, the social housing rate, that may be taken as a proxy to social diversity.

Income distribution.

We now consider a second example, and build trajectories for the income distribution in Paris.

As explained at the beginning of this section, when working with distributions available only through

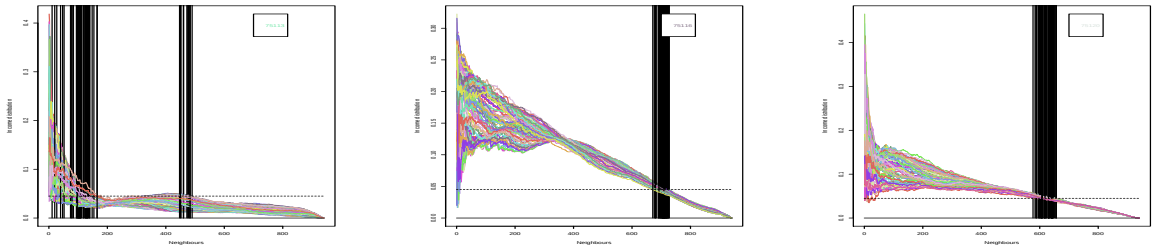


FIG. 34 Trajectories for the income distribution, starting from each IRIS (census block) in Paris's 13th (left), 16th (middle) and 20th (right) districts. Ordinates correspond to the Kolmogorov-Smirnov distance (KS) between a group of blocks' income distribution and the whole city's income distribution. Solid vertical lines correspond to radii of convergence: the radius of convergence of a trajectory is defined as the point where the path last enters within 0.05 of the city's average in terms of KS distance.

their quantiles, computing points for each group of blocks is a slightly more involved task. In this case, we chose to estimate the parameters of the best-fitting distribution (which happened to be log-normal) given a block's quantiles. Then, from this distribution, we simulated data corresponding to the number of households in the block. Thus we obtained a full ensemble of simulated households for every possible group of neighbouring blocks, from one block only to the whole city.

Let us look at trajectories for IRIS blocks in three Parisian districts of comparable sizes and population numbers: the 13th, 16th and 20th districts – see Figure 34. The first one corresponds to the south-western peripheral part of the city, the second one to the south-eastern peripheral part of the city, and the third one to the north-eastern peripheral part.

The 16th tends to be further away from the full city's picture, as it takes generally longer for its blocks to converge to the city's distribution than for blocks in the 13th or in the 20th districts. The 13th district exhibits an interesting behaviour, with some trajectories that first come close to the city's mean but then bounce up again further from it. This can be understood as follows: some IRIS blocks in the 13th district belong to relatively well-mixed neighbourhoods, so initially their trajectory approaches the city's average distribution. But the aggregation process around them leads to incorporate much less wealthy blocks on the border of the city, at a stage where wealthier blocks towards the center of the city and beyond have not yet been aggregated in large enough numbers to counterbalance the former. This sends trajectories away from the city's distribution again. A similar effect may (and does) take place for neighbourhoods relatively close to blocks that are much wealthier than the city as a whole (*e.g.* in the north-western 17th district, not shown here).

Let us now look at a map of Paris with IRIS blocks coloured according to their radius of convergence for the income distribution (Figure 35). The picture is here much more symmetric than for the social housing rate. Indeed, wealthier (south-western) and poorer (north-eastern) parts of the city both exhibit similar, slower convergence to the city's global distribution, compared to well-mixed areas such as the *faubourgs*. However the wealthier, south-western part is again the last one to converge and constitutes a larger, more secluded area, including all of the 7th, 16th and 17th districts, and most of the 6th and (more

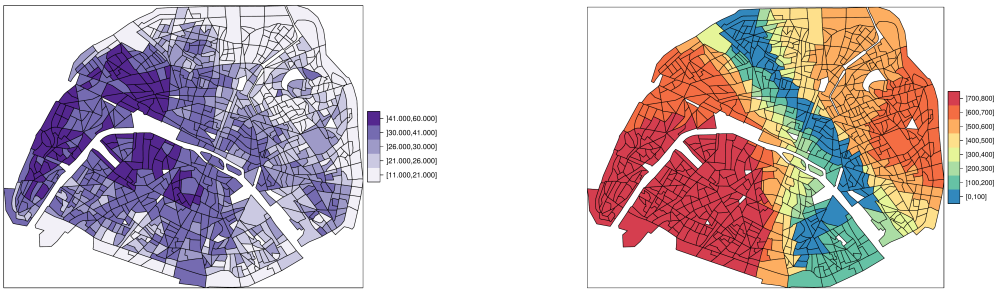


FIG. 35 IRIS blocks coloured according to their radius of convergence to the city's income distribution (right). For comparison, the map on the left-hand side shows IRIS areas coloured according to the median of the (local) income distribution. Distance between distributions is measured with the Kolmogorov-Smirnov distance. (White areas correspond to parks, riverbanks and other IRIS where data is not available.)

surprisingly) the 14th district. Remarkably, the 8th district (around the *Champs Élysées*) does not stand out as a secluded area for the income distribution: this is because, geographically, the 8th is much closer to well-mixed neighbourhoods than, for instance, the 16th district.

Metro and tramway station density.

We come to a third type of data, for which one may define a local density. Recall the definition of

$$r_i(l) = \frac{m_i(l)/M}{p_i(l)/P}, \tag{96}$$

where M is the total number of metro and tramway stations in Paris, $m_i(l)$ the number of stations in a disk of radius l centered on IRIS i, $p_i(l)$ the population living in the same disk, and P the total population in Paris.

This quantifies whether the local density of metro stations per inhabitant is smaller ($r_i(l) < 1$), larger ($r_i(l) > 1$) or equal ($r_i(l) = 1$) to the city's average density. In other terms, as far as the number of metro and tramway stations is concerned, is the population living in the area under consideration deprived, well-served or neither?

Letting l vary, we may then compute surface areas (equivalently, radii) for which, around a given point, one obtains a density per inhabitant similar to the average one in the whole city.

As can be seen in Figure 36, our method provides a different picture from the one obtained by simply looking at local densities. The fact that the Parisian metropolitan network is mostly structured around a North-South axis is clearly visible in terms of convergence radii, whereas local densities almost exclusively show the higher concentration of metro stations in the city center. Note also that local densities are oversensitive to the local presence of stations (*eg* along tramway lines in peripheral IRIS blocks).

The quantity r_i defined above (Eq. 96) is similar to the "representation" defined in (153) in order to examine patterns of residential segregation. Let us also follow in their footsteps by emphasizing, in the next section, the benefits of thinking in terms of what segregation is *not*. Indeed, Louf and Barthelemy used their representation index to define a null model; we show here how our method of trajectories may be used to define similarly a reference model in terms of a uniform city.

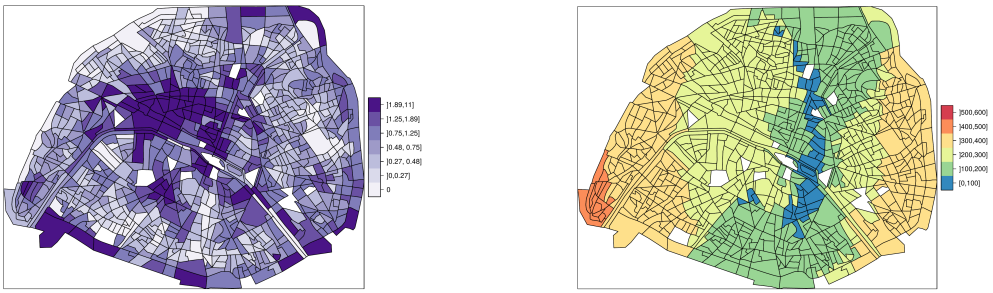


FIG. 36 IRIS areas coloured according to their radius of convergence to the city's average per-inhabitant density of metro and tramway stations (right). For comparison, the map on the left-hand side shows IRIS areas coloured according to the local density of metro and tramway stations (within 500 meters of an IRIS's centroid). (White areas correspond to parks, riverbanks and other IRIS where data is not available.)

5. The uniform city as a null model: random walks to the mean

The importance of null models has perhaps become more palpable in the last few decades, as new mathematical and statistical tools have come into play in a wider variety of fields, from urban geography to history and archeology ((101; 170)).

The null model against which segregation should be defined is the "unsegregated" city, as pointed out by (153). In the unsegregated city, distributions are uniform across space. So that, for instance, if the rate of social housing in the whole city is ρ, then any housing unit in the city, irrespective of its spatial position, has probability ρ of being a social housing unit: there is not any spatial structure in the distribution of social housing units. This means that when considering a sequence of units defined solely according to some spatial rule (*eg* aggregating sequentially nearest neighbours), no sign of this spatial rule will appear in the sequence: everything will be as if one were drawing units at random from a well-mixed urn.

Any deviation from such random behaviour will be a sign that an underlying structure leads to a biased shuffling of the units. Note that the underlying structure may be a spatial one, as in the viewpoint we are taking here, but it may also be of another nature depending on the type of data one is examining.

A null model

Let us examine here the somewhat ideal case when one is dealing with count data, i.e. data at the individual level. Suppose each individual belongs to one of two groups A and B, *eg* people belonging to one of two social groups, or housing units being of one of two types. Start from a given point in the city, and increase the population, one by one, from the individual located there to the whole population in the city. Write $X_i = 1$ if the i-th individual is of type A, $X_i = 0$ otherwise. Then setting $S_0 = 0$ and

$$S_n = \sum_{i=1}^{n} X_i \tag{97}$$

produces a trajectory, that is: a sequence of points (n, S_n), with n varying from 0 to the total number N of individual units in the city. S is simply the count data for group A.

In the case when the city is perfectly mixed, the trajectory thus produced will be a random walk S that,

given its current value S_{n-1}, either moves up with probability

$$\rho_n = \frac{\rho N - S_{n-1}}{N - n + 1}, \tag{98}$$

or stays put with probability $1 - \rho_n$, where ρ represents the fraction of group A in the total population (N). For the more mathematically inclined: such a random walk may also be viewed as a standard problem of drawing without replacement from an urn containing N balls, ρN of which being of one type and the rest being of another type. S_n would thus follow a well-known hypergeometric distribution – see (97).

Other sequences of interest that may be defined similarly are

- the sequence of averages:

$$M_n = S_n/n;$$

- the sequence of differences D_n, when one counts $+1$ for an individual of group A and -1 otherwise:

$$D_n = 2S_n - n.$$

Both sequences, M and D, are random walks that move up with probability ρ_n or down with probability $1 - \rho_n$, ρ_n being as defined above in equation (98).

All three types of trajectories converge to well-identified final values: S converges to $N\rho$, M converges to ρ and D converges to $N(2\rho - 1)$. From a probabilistic point of view, any of these three sequences is fully characterized by ρ and N. Here they represent three facets of the same null model: a perfectly mixed city.

Now, if one has count data for a given variable in a city, one may build N sequences (one for each starting point). One thus produces a sample of trajectories that would, if the city were well-mixed, exhibit the statistical properties of a set of N random sequences obtained by drawing balls without replacement from a well-mixed urn. Any statistically significant deviation from such a set of random sequences signals the presence of spatial dissimilarities at multiple scales. We look at such deviations in the next subsection.

Deviation from random sequences

Count data sequences are similar to sequences that one encounters when counting votes in an election. Imagine an election with two candidates, A and B. Counting the votes cast for candidate A produces the S-sequence defined in the previous subsection, with $N\rho$ the final number of votes for A. A famous result, or rather a famous set of results, linked to count data in elections, is the so-called Ballot theorem: (2; 18; 34; 97; 223; 236). In its simplest form, it gives the probability that A be always in the lead, all through the counting process, given that A wins the election with a final share ρ of the votes. This is $2\rho - 1$ (note that if A wins the election, necessarily $\rho > 1/2$).

A very natural interpretation of the Ballot theorem in terms of urban count data is the following. If I start from some point in the city and look at the first two housing units around me, and then the first three, and so on, what is the probability that I will always see a minority of social housing units, given that the

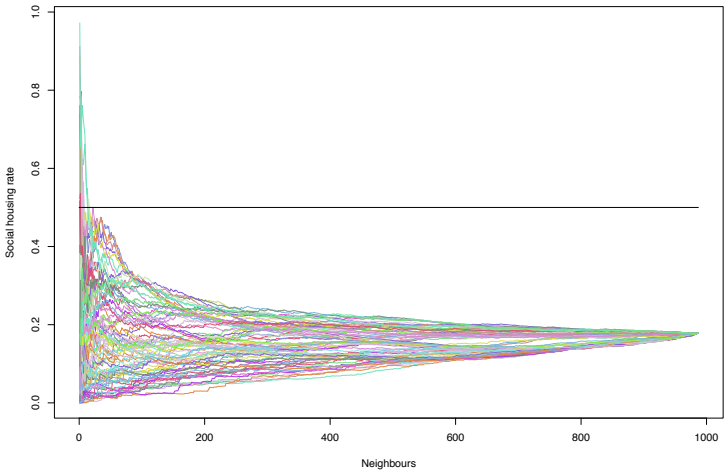


FIG. 37 Trajectories for the social housing rate, starting from some (10%) of the 936 statistical units (so-called IRIS) in Paris. The solid line corresponds to a rate of 0.5. According to the Ballot Theorem, were the city "well-mixed" in terms of social housing, just above 64% of the trajectories would always remain below the 0.5 mark. In Paris, almost 85% of them do so.

fraction of social housing in the whole city is, say, 20%? The Ballot theorem tells us this probability is 60%. So that, even if social housing is in a strong minority and if the city is perfectly mixed, when zooming out from a point chosen at random there is only a 60% probability that one will never come across a bespoke neighbourhood where social housing is in the majority.

On count data, a first test for the presence of an underlying spatial structure therefore simply consists in examining significant deviation from the Ballot theorem. In Paris for instance, the total fraction of social housing in the city is about 17.9%. Therefore 64.2% of the trajectories should always stay below the horizontal line $y = 0.5$. In actuality, almost 85% of them always stay below 0.5 (Figure 37).

Obviously, one should account for the fact that we only have a finite sample of sequences – this is typically done through statistical hypothesis testing. In this specific context, we need a test to decide whether the observed deviation from the Ballot theorem is significant or could be a simple random sampling effect. To this end, define the following sequence of Bernoulli random variables:

$$Y_i = \begin{cases} 1 & \text{if trajectory for the } i\text{-th IRIS is always below } 0.5 \\ 0 & \text{otherwise.} \end{cases}$$

According to the Ballot Theorem,

$$\text{Prob}\,(Y_i = 1) = 0.642, \quad \forall 1 \leq i \leq N.$$

If all N trajectories were "drawn" independently from a well-mixed city, the sum $\sum_{i=1}^{N} Y_i$ would follow a binomial distribution with mean $0.642N$ and variance $0.23N$. Here a simple goodness-of-fit test shows that the Parisian trajectories for the social housing rate do not obey the Ballot Theorem, with a p-value of

order[11] 10^{-16}.

Further, deviations from the null model of an unsegregated city may be characterized by a more complete statistical test. Indeed, with N starting points in a city, one builds N sequences, which should be compared to N full drawings of a well-mixed urn. By full drawing, we mean drawing all balls, one by one, without replacement, until the urn is therefore left empty. In partial drawings, where one draws n balls from an urn containing N balls, standard statistical tests include the hypergeometric one, similar to Fisher's exact test (64; 102; 103). In fact, one could use such a test at every single point of any of the trajectories. However, this would not be using the available information to the full. In future work, we will make use of a multiple test for the full set of trajectories' points $S_n^{(j)}$, with $1 \leq n \leq N$ and $1 \leq j \leq N$ – see for instance (207). We will also present ways of characterizing deviations from the null model for data types other than count data. This will lead to new ways of measuring segregation levels. A thorough, systematic investigation of these new indices, including rigorous comparisons with commonly used indices (see eg (106; 165)), will be carried out.

6. Comments

Aggregation of spatial data units has been so far generally considered as a difficulty to circumvent rather than an opportunity to analyze spatial dissimilarities. The new method introduced here uses aggregation as a means to extract information on the relative singularity of each spatial unit within the city as a whole. Aggregation procedures may indeed reveal features that are not so easily seized from other perspectives. Another example is the recent use by (8) of aggregation to explore the behaviour (and more specifically the scaling laws) of various statistical variables across a phase space corresponding to almost all possible city boundaries' definitions in England and Wales. Aggregation is akin to an exploration of space from a given starting point, and it is this observation that forms the basis of the method we have presented here.

A question we have not enlarged upon here is that of the definition of convergence. For instance, we chose a ± 0.05-interval for convergence to the city's average social housing rate. In fact, one may tune this threshold so as to study in greater detail areas that converge more slowly or more quickly. Our current work focuses on this question: looking at a transform of the trajetories defined here, we explore the full range of possible convergence thresholds and define a more accurate measure than using arbitrary thresholds (this will soon be published (66)). It will also be interesting to explore models able to reproduce stylized aspects of observed trajectories: hidden Markov models, multi-agent models, intermittent diffusion models.

We are also currently working on a difficulty that arises in practice: one rarely has access to individual unit data. The data is often already aggregated at some basic statistical unit level. This means trajectories are observed only at certain points determined by the sequence of sizes of the statistical units instead of at every individual point. So that comparisons with Ballot Theorem results need to be made through what is

[11] The minuteness of this p-value is due to the fact that the perfectly mixed city is an extreme, unrealistic case. It provides a scale that is too extreme, if one wishes to use it directly as a measure of segregation. However, it is an absolute scale, much in the same sense as the absolute temperature in physics. As such, it may be used reliably for comparisons.

called *subordinated* random walks – which are precisely sequences observed at a number of different points rather than at every single step. Results on this will also appear in the coming months (154).

C. A Neural Network Measure of Segregation

Let me now present another data-based approach, that provides a multidimensional analysis of segregation phenomena. We also introduce segregation indices that can be sensibly and robustly defined from such a multidimensional picture.

Specifically, let us see how to define a data-based, multidimensional segregation index *via* Self-Organizing Maps (141; 142). SOM's intrinsic multidimensionality presents many benefits for the study of such a complex system as a city. This has only recently been noted (10; 232) and our work shall hopefully contribute to SOM becoming a standard tool in urban sociology and geography. All the more so as, most importantly and specifically, the topology obtained by SOM allows for useful comparisons with the actual geographical topology. Precisely, our idea is to measure segregation from correlations between SOM distances and geographical distances.

1. Data and variables

We use databases from INSEE (*Institut National de la Statistique et des Études Économiques*, France's national agency for economical data), IGN (*Institut Géographique National*, France's national agency for geographical data) and RATP (*Régie Autonome des Transports Parisiens*, Paris public transport agency).

In INSEE data, census blocks are called IRIS (*Ilots Regroupés pour l'Information Statistique*). Unfortunately, neither do they correspond to a fixed surface area nor to a fixed number of inhabitants, although they correspond on average to blocks of around $2,000$ inhabitants. Thus, the city of Paris, with a total population of over 2 million people comprises just under $1,000$ IRIS. Some IRIS blocks are also purely geographical, with no or very few inhabitants. They appear nonetheless in INSEE data because of services and facilities they may offer. They also appear in the contour data provided by IGN to draw geographical maps and spatial representations.

At the census block level, INSEE provides data such as the number and types of shops, the number and types of public service offices, the number and type of health facilities. They also provide quantiles of the income distribution within each census block – except those where the number of households is too small and combined with a high level of income. For this and other similar reasons, the number of IRIS for which data is available differs from one variable to the other.

The metropolitan authority for transportation provides geographical coordinates for all access points to underground, tramway and bus stations in the Paris area. Using this data (a version of it processed by the OpenStreetMap project (186)), we have computed for each census block the number of underground and tramway lines available within an 800 meter radius (from the centroid of the block). Note that a station with two lines counts twice as a station with just one line.

Variables

For this exploratory case study, we have retained three sets of variables for each census block:

Set	Label	Variable
1	Decile.1	1st decile of income dist.
1	Mediane	Median of income dist.
1	Decile.9	9th decile of income dist.
1	Part_patrim	Share of financial and patrimonial income
1	Part_min_soc	Share of minimum social benefits
2	age_moy	Average age
2	std_age	Standard deviation of age dist.
2	moins_18_moy	Average number of inhabitants under 18 y.o.
2	Diplom_moy	Average level of education (1: pre-secondary; 5: postgrad.)
2	std_diplom	Standard deviation of education level
3	taux_hlm	Fraction of social housing
3	1_D_EP_Sum	Number of EP primary schools
3	1_D_Public_Sum	Number of state primary schools
3	Col_EP_Sum	Number of EP secondary schools
3	Col_Public_Sum	Number of state secondary schools
3	1_D_Priv_Sum	Number of non-state primary schools
3	Col_Priv_Sum	Number of non-state secondary schools
3	Commerce_Sum	Number of shops
3	Services_Sum	Post offices, local administration offices
3	Sports_Sum	Number of sports facilities
3	Action_sociale	Social services offices
3	Medecins_Sum	Number of medical doctors (GPs and specialists)
3	Sante_Sum	Number of hospitals, pharmacies
3	Transport_Sum	Number of train stations and travel agents
3	Somme_lignes	Number of metro and tram stops (within 800m)

TABLE II Variables used for this exploratory study. EP stands for "Éducation Prioritaire", a state-sponsored scheme for schools in deprived neighbourhoods. (When counting shops, schools, services etc, the 10 nearest neighbouring blocks of a given IRIS are taken into account.)

Set 1 revenue and income: first and ninth deciles as well as median of the income distribution, fraction of revenue coming from assets and other patrimonial sources, fraction of revenue coming from minimal social benefits[12]. From the available data, these variables could be computed for 853 IRIS blocks in Paris.

Set 2 population: age (average and standard deviation), number of people under 18 years old, education level (coded in 5 groups, average and standard deviation). From the available data, these variables could be computed for 943 IRIS blocks in Paris.

Set 3 urban facilities and services: rate of social housing, access to public transport, number of shops, access to medical and health services, number of sports facilities, number of primary and secondary schools (including primary schools in special urban and education development projects – called *éducation prioritaire* (EP) in France). From the available data, these variables could be computed for 980 IRIS blocks in Paris.

A full list of the variables used in this study is given in Table II.

2. Self-Organizing Map Approach

We use a multidimensional classification algorithm known as Self-Organizing Map and first introduced by T. Kohonen (141; 142). All results were obtained using the R-package SOMbrero (44), which performs SOM combined with a hierarchical agglomerative clustering (HAC). Given that we are dealing with just under one thousand census blocks, we choose to train the algorithm to produce an 8x8 map for each of the three sets of variables, and then to classify individual IRIS blocks into four groups for sets 1 and 2 and six groups for the third set. The number of clusters is chosen from the dendrograms so as to provide a meaningful classification, with a balance between too little separation among groups and too few details in their description: four to six types of neighbourhoods compose a reasonably coarse-grained rendering of geographical and sociological details.

The online SOM method implemented in the SOMbrero package being a stochastic algorithm, we allowed for 100 runs on each set of variables. We then extracted results from the runs exhibiting the best explained variance ratios. These are shown on Figures 38, and we proceed to analyze them.

[12] These are social benefits paid to prevent people from falling into extreme poverty. They vary from 300 euros to about 800 euros per month.

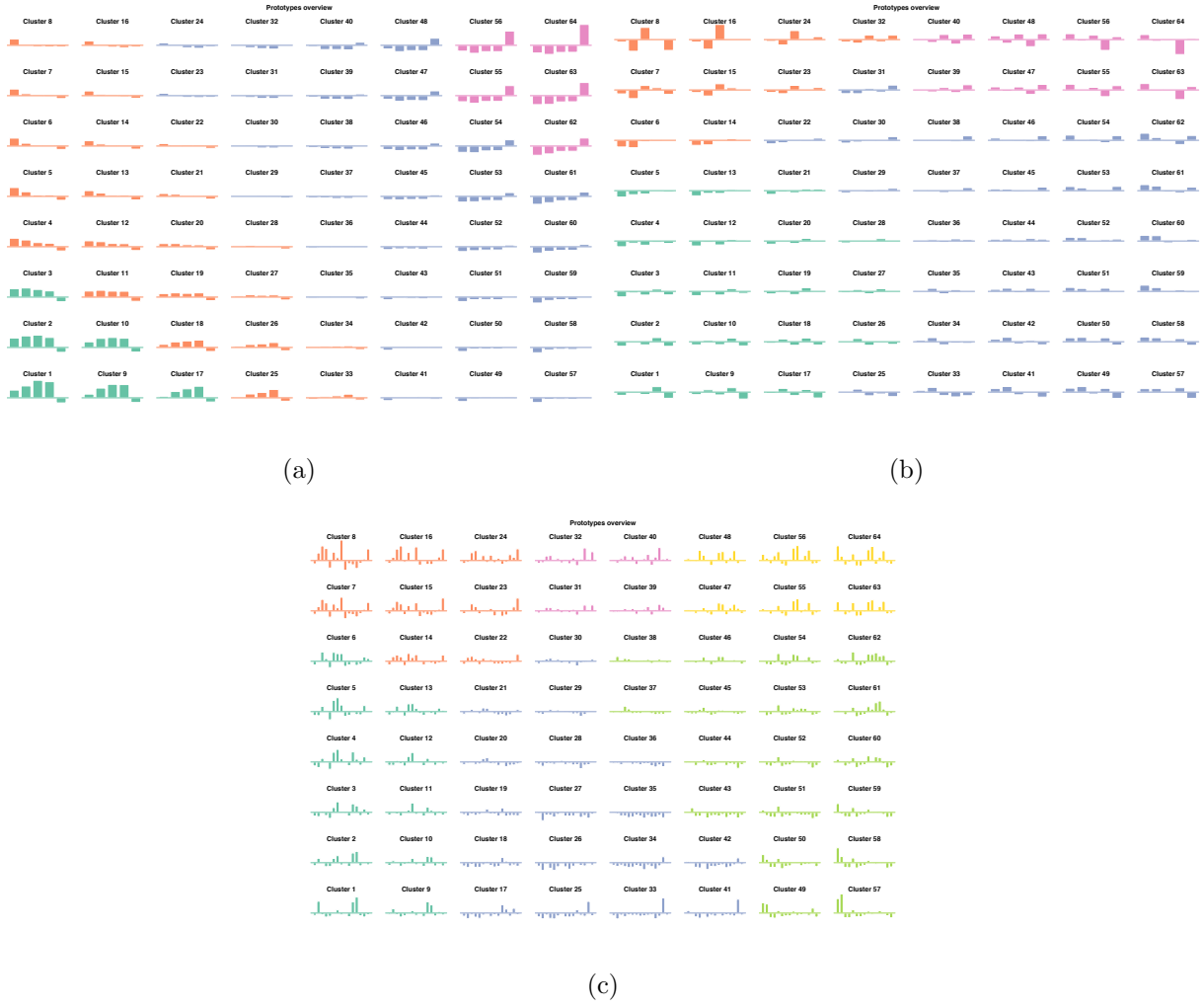


FIG. 38 Kohonen maps for the variables of set 1 (subfigure a), 2 (b) and 3 (c), with values of the variables for the prototypes in each class. Colours indicate groups of clusters obtained by hierarchical agglomerative clustering (HAC): 4 groups for Sets 1 and 2, 6 for Set 3.

SOM on Set 1

The first set of variables correspond to income and revenue variables. These are the most commonly used in socio-economic segregation studies (along with ethnic group variables – let us recall here that ethnic statistics are not available in France). As can be seen on Figure 38, SOM followed by HAC yields here well ordered groups, parallel to the diagonal, corresponding to four easily identifiable types of census blocks (see Table III). There are blocks (group 4) where the population is clearly poorer than the Parisian average, with a first decile of just above 7,000 euros per year. At the other end of the spectrum, blocks where the population is not only richer than average (specifically the top 10 percent are much wealthier, with a ninth decile above 129,000 euros per year) but also with a very substantial part of revenues coming from financial and other patrimonial assets: 40 percent on average (group 1). In between these two groups, the other two

Group	1st decile	Median income	9th decile	Revenue from assets	Revenue from social benefits
1	13,405	44,367	129,538	40	0.2
2	12,504	33,281	73,929	22	0.5
3	9,471	23,963	50,338	13	1.2
4	7,390	15,081	30,840	7	3.6
All	10,672	28,411	65,309	19	

TABLE III Per-group averages of some of the variables in Set 1 (in euros for quantiles of income distribution, in percent for the share of revenue drawn from financial and other assets).

types of blocks correspond to upper (2) and lower (3) middle classes, with again a difference in the level of patrimonial income (22 percent vs 13 percent). Patrimonial income thus seems to work as an order variable (in the sense that it characterizes the cluster to which a block belongs). It is significantly correlated with spatial segregation: Figure 39 shows indeed a high level of spatial homogeneity for the groups derived from the first set of variables.

SOM on Set 2

SOM followed by HAC run on the second set of variables yields a less structured clustering than in the case of the first set of variables (see Figure 38). However, one can identify four groups and the following underlying trends: the bottom of the Kohonen map (groups 1 and 3) corresponds to blocks with fewer

FIG. 39 Spatial distribution of groups 1 to 4 for Set 1. (Colours correspond to groups as defined on Figure 38. Areas in white correspond to parks, train stations and blocks for which data is not available.)

Group	Age	Age std deviation	Children per household	Education
1	39.2	17.7	0.3	2.8
2	39.3	14.4	0.6	2.7
3	44.5	18.6	0.4	2.7
4	46.4	16.7	0.5	2.1
All	42.6		0.4	2.6

TABLE IV Per-group averages of some of the variables in set 2 (education level is from pre-secondary (1) to postgraduate (5) level).

children and a higher education level than the top part (groups 2 and 4). Also group 1 has younger heads of households on average than group 3. Similarly in group 2 the population is comparatively younger than in group 4, and with a higher education level (see Table IV).

FIG. 40 Spatial distribution of groups 1 to 4 for Set 2. (Colours correspond to groups as defined on Figure 38.)

Spatially, groups have a wider distributions (Fig. 40), but one still notes that certain areas are particularly representative of a given cluster, eg the northern-north-eastern part of the city for group 4.

SOM on Set 3

Processing the third set of variables with SOM and HAC allows one to distinguish six well-identified types of census blocks, that pave the Kohonen map (Fig. 38 and Table V):

1. areas with more medical services, more private schools, fewer shops and less access to public transports;

Group	Social housing	Medical doctors	EP schools	Shops	Public transport
1	5	298	0.1	362	7.2
2	5	181	1.4	626	16.6
3	9	172	0.5	236	7.2
4	10	129	5.4	406	12.4
5	34	129	2.8	177	5.9
6	42	100	10.5	182	6.3
All	18	173	2.8	288	8

TABLE V Per-group averages of some of the variables in Set 3 (social housing rate is a percentage per block, other variables are raw numbers for each block and its ten nearest neighbours; access to public transport is the number of lines within 800 meters of a block's centroïd).

2. areas with many shops, facilities and the highest access to public transports;

3. areas with a slight concentration of social housing (9 percent) and below average for all other variables;

4. areas with a high level of access to public transports, many shops and facilities and also a certain number of EP primary shools;

5. areas with a significant proportion of social housing (34 percent), very few EP primary schools, and the lowest access to public transport;

6. areas with the highest proportion of social housing (42 percent), the largest number of EP schools, and low access to public transport and other facilities.

On this set of variables, a multidimensional approach such as ours sheds light on residential patterns, allowing to refine as one sees fit the level of description. For instance, in this case if one opts for only five groups, groups number 5 and 6 will be merged, as may be seen from their proximity on the Kohonen map.

If one looks at the spatial distribution of the groups (Fig. 41), there are again some district areas (called *Arrondissements* in Paris) that emerge as particularly representative of a given cluster: parts of the 5th, 6th and 16th Arrondissements for group 1, parts of the 1st, 2nd, 8th and 9th for group 2, parts of the 7th and 16th for group 3, the area around *Place de la République* for group 4, 13th, 19th and 20th for group 5, and north-eastern parts of the 18th for group 6.

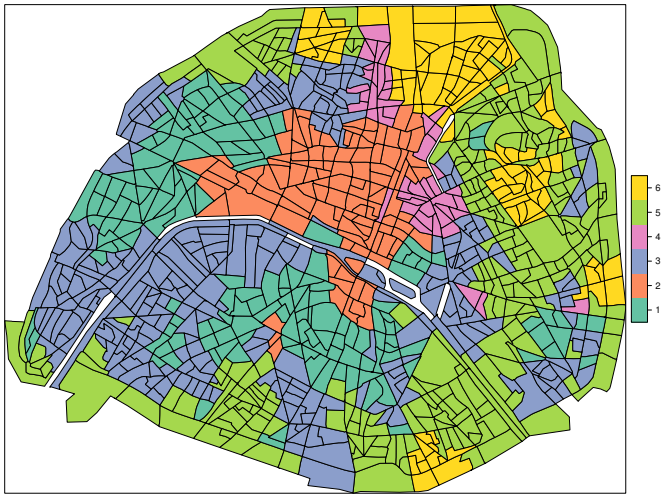


FIG. 41 Spatial distribution of clusters 1 to 6 for Set 3. (Colours correspond to groups as defined on Figure 38.)

3. Segregation indices

Thanks to its multidimensional nature, the approach followed in the previous sections yields typologies of neighbourhoods according to multiple variables taken simultaneously into account. Now, for a given set of variables, are all types of neighbourhoods well mixed across the city, or are there any spatial patterns? In other terms, is there any form of spatial segregation along some of the variables considered here? Is there a set of variables for which segregation patterns are stronger than for the other two?
Looking at the maps on Figures 39, 40 and 41, one sees spatial patterns – but, how significant are they? We introduce in this section a method that allows to quantify them in a new manner, thanks to specific aspects of the SOM algorithm.

Indeed, a well-known difficulty arising in the study of segregation phenomena is that of defining indices to measure the actual level of segregation (19; 95; 196). Since most INSEE and other public data in Paris is only available at the IRIS level, let us think at this level. Then measuring segregation amounts to quantifying two things:

1. how different IRIS blocks are from one another (that is, individual census blocks are not all alike, eg some have a younger population than others, some have a richer population than others);

2. how much spatial concentration occurs for IRIS blocks of a given type (in terms of the SOM groups obtained in the previous section, this means how far one stands from a uniform distribution over the whole city for blocks belonging to each group).

The first point may be addressed using the stochastic nature of the SOM algorithm. Indeed, if all blocks were (almost) identical for a given set of variables then the clustering obtained would not be very robust to a re-run of the algorithm with a different seed (43). (See paragraph II.C.3 below.)

The second point may be treated from two perspectives. One can measure the geographical dispersion of blocks in a given cluster. Or, one can make use of one of SOM's specific properties. Indeed, proximity on the Kohonen map means proximity in the state space of the chosen set of variables. Thus comparing Kohonen distances and geographical distances gives a measure of segregation. Based on this observation, we introduce ideas for a new segregation index. Beforehand, we compute some standard segregation indices.

Some entropy and information theory indices

A class of segregation indices is formed by the so-called entropy indices (124; 225). Based on information theoretic entropy, they measure the difference between the entropy of the global distribution at the city scale and local distributions (we shall consider here each IRIS within its administrative neighbourhood – with 91 such neighbourhoods in Paris). A standard information theory index is the H index defined in (195; 196), implemented in the R package *seg* (123). Other standard segregation indices include the R-index (relative diversity) and the D-index (dissimilarity).

A summary of values obtained for these three standard segregation indices is given in Table VI. All three indices indicates stronger segregation on the third set of variables, and weaker on the second one. The first set appears close to the third one in terms of segregation.

Set of variables	Index 1	Index 2	Index 3
1	0.64	0.43	0.50
2	0.44	0.22	0.28
3	0.72	0.49	0.60

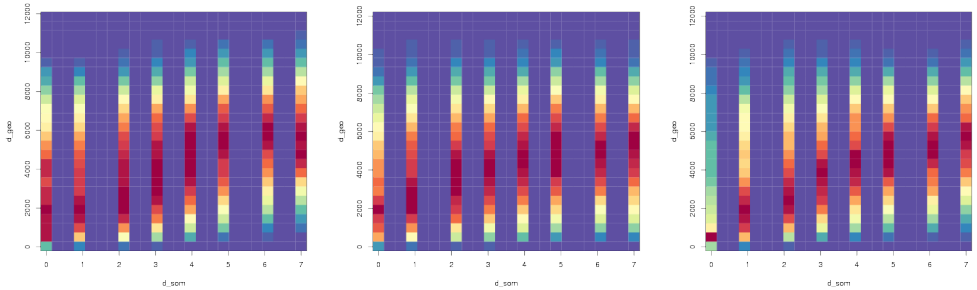
TABLE VI Values of various segregation indices for the clusters produced by SOM on our three sets of variables. Index 1 is the D-index. Index 2 is the R-index. Index 3 is the H-index. See (195; 196) and the R package *seg* (123)

SOM-based segregation index

As mentioned in the introduction of this section, a new integrated segregation index can be obtained by comparing Kohonen distances and geographical distances for all pairs of IRIS blocks.

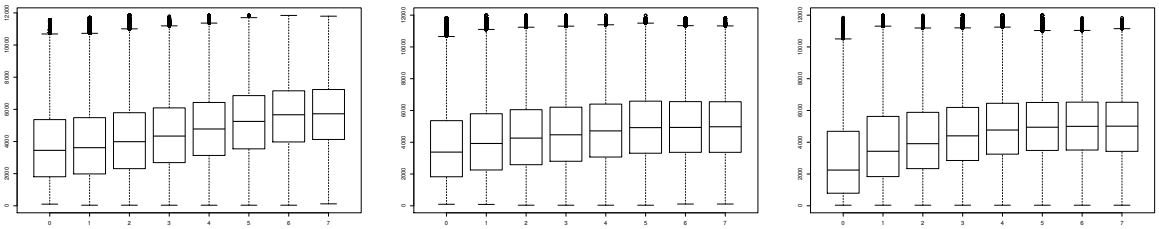
Indeed, proximity on the Kohonen map corresponds to proximity in the multidimensional space of the variables. The closer two blocks are on the Kohonen map, the more similar they are for the variables under consideration. Now, if the city were well mixed, one would not observe any particular spatial pattern: geographical distances would be independent of Kohonen distances. Thus any correlation between geographical and Kohonen distances signals spatial patterns, i.e. the presence of segregation, the level of which is well quantified by the actual value of the correlation.

One convenient way of visualizing this correlation is to represent the density of pairwise geographical distances for each value of the Kohonen distance, as on Figures 42 and 43. Were the city well mixed,



(a) (b) (c)

FIG. 42 Visual representation of a SOM-based segregation index for IRIS blocks considered with the variables of Sets 1 (a), 2 (b) and 3 (c). For each value of distance on the 8x8 Kohonen map, the density of pairwise geographical distances is shown (red means higher density). At a point with coordinates $(0, 2000)$ one reads the probability density to find a pair of IRIS blocks that are in the same Kohonen class while their centroids are 2 km apart in the actual city. If IRIS blocks from each Kohonen class were uniformly scattered across the city, geographical distance distributions would be the same for every Kohonen distance.



(a) (b) (c)

FIG. 43 Box plot representations of the index presented on Figure 42.

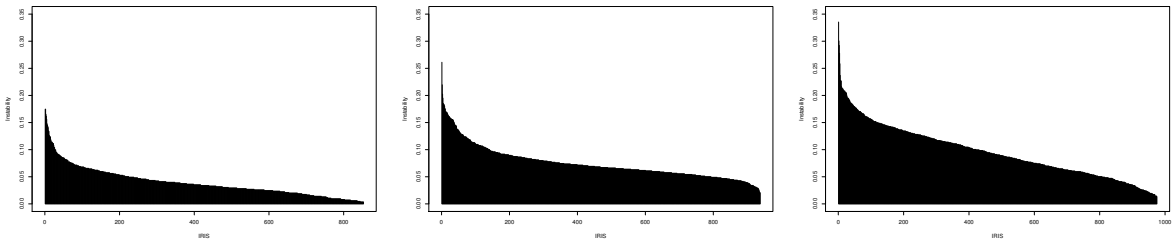
geographical distance distributions would be the same for every Kohonen distance.[13] This is not the case here, revealing in particular greater levels of segregation for the variables of Set 1, contrary to what appeared with standard segregation indices (see Table VI). Indeed, looking at numerical values (see Table VII), one observes that the correlation between Kohonen and geographical distances is twice as large for Set 1 as for Set 2, and about one-and-a-half as large for Set 1 as for Set 3.

Numerical values are to be compared with similar correlation coefficients computed in imaginary, extreme cases. For instance, if one considers four groups fully separated (with four "pure" neighbourhoods on the four quadrants of a square city), the correlation may be computed exactly and is about 0.61. Given this value for the most extreme (unrealistic) case, one may say that the city of Paris exhibits significant levels of spatial segregation on all three sets of variables.

[13] Note that such a simple, direct measure of the correlation between geographical distance and Kohonen distance is well suited to intricate patterns of segregation, as observed in real cities. However, if one considers artificial patterns with much regularity, this correlation measure works well on checkerboard patterns (provided the mesh is not too small), but obviously not as well on concentric patterns. More work is needed to circumvent this difficulty.

Set of variables	SOM-based segregation index
1	0.26
2	0.13
3	0.18

TABLE VII Values of the SOM-based segregation index on the three sets of variables. This is defined as the correlation between Kohonen distances and geographical distances.



(a) (b) (c)

FIG. 44 Volatility part of the SOM-based segregation index on each of the three sets of variables defined in Table II. Plots show level of fickleness for each IRIS block, in decreasing order of magnitude, for Sets 1 (left), 2 (centre) and 3 (right).

Quantifying the heterogeneity of IRIS blocks

Any comparison between geographical distances and the distances obtained through a clustering algorithm obviously relies on the robustness of the classification. One should therefore control for the volatility level of the clustering: for a given pair of IRIS blocks, how sure can one be that they lie at a given Kohonen distance? The stochastic nature of the SOM algorithm allows one to address this question. Indeed, volatility can be measured by looking at the fraction of pairs of areal units that are stable – i.e. that always belong together to the same cluster (or the same vicinity) through a large number of SOM runs launched with different, random seeds (43; 45; 46). One may then count, for each unit, the number of unstable pairs to which it belongs, thus defining its level of "fickleness".

As an illustration, we show on Figure 44 levels of fickleness for the three sets of variables considered in the previous sections of this paper. SOM classification appears robust across all three sets of variables, with a maximum of 10% of fickle pairs observed on Set 3. Also, the classification is more robust on Set 1 than on Sets 2 or 3, indicating a greater level of differentiation among blocks along the variables of Set 1.

The larger the number of fickle units, the less heterogeneity there is among them and therefore the less definite will any classification be. In the extreme case in which all areal units present the same characteristics, any choice of classes will be arbitrary and this will be reflected in the fact that every pair of units will be unstable and every unit will be fickle. Note that this is not only a measure of the robustness of the classification: it actually measures the level of heterogeneity between areal units, independently of their

Set of variables	Percentage of fickle pairs
1	4.0
2	7.5
3	9.6

TABLE VIII Percentage of unstable pairs computed on 100 runs of the SOM algorithm for each set of variables. An unstable pair is defined as a pair of IRIS blocks that do not tend to be always in the same vicinity or always *not* in the same vicinity on the Kohonen map through the 100 runs.

spatial distribution. In this respect, levels of fickleness form a second part of a SOM-based segregation index (the volatility part). They complement the correlation defined in the previous paragraph, which measures the spatial aspect of segregation.

4. Comments

The use of self-organizing maps to explore socio-economical and geographical data presents many promising features. The method is intrinsically multidimensional, and the clustering obtained is very much suited to interdisciplinary work: for instance, typologies of urban areas infered from sociological surveys and other forms of studies can be compared to clusters obtained from SOM, and their robustness can be tested by performing a large number of runs with different seeds.

Further, this new method allows for the definition of a general, robust segregation measure, both on the social and spatial levels – the former is measured by the percentage of fickle pairs, and the latter by the correlation between Kohonen and geographical distances. We are working onward to develop a full methodological framework, with a thorough comparative study of the new segregation index introduced here.

CONCLUSION

This thesis surveys the two main tracks in the research I have been conducting since my PhD. New, other projects have stemmed recently – too recently to provide results for presentation, but let me give some details about them here. Reversing the order of the main text, I shall start with interdisciplinary projects.

A year ago I started a collaboration with historian Karine Le Bail (CNRS/EHESS) on the analysis and modelling of data linked to the phenomenon that is known as *Épuration*, in which, at the end and in the aftermath of World War II, people were tried for having collaborated in various ways with the Nazis. Karine Le Bail focuses on the performing arts, and investigates the complex and very diverse situations and trajectories of musicians, singers and other artists. Our research project, currently funded under a CNRS interdisciplinary grant,[14] involves collecting data from national, regional and local archive centres and implementing a bespoke processing scheme so as to be able to both visualize and analyze this data with contemporary statistical techniques. In particular, through a network representation and via the use of stochastic and latent block models, we are seeking to establish whether or not certain groups (female artists, classical musicians,...) have been discriminated against in the *Épuration* process. We should be able to pre-publish some first results in the coming semester.

Another current project, one that has just started, is a PHC[15] project with the University of La Havana, Cuba. Together with Paris-1 colleague Madalina Olteanu and Cuban mathematician Sira Allende, we are coordinating a 14-strong collaboration to work on socio-spatial dynamics in French and Cuban cities. This is a three-year project, that includes the co-supervision of a PhD thesis, the organizing of three workshops and a lecture series.

This complements other ongoing research projects with William Clark and Mark Handcock at UCLA, with whom we are working (Madalina Olteanu and I, as well as two of our students, Antoine Lucquiaud and Alex Mourer) on new extensions of the trajectory convergence method presented in Section II.B. We are working on refining the convergence criterion by considering the full range of possible convergence thresholds. This leads to a more rigorous mathematical definition of the singularity of each neighbourhood in a metropolitan environment. Simultaneously, we are applying this new method to a growing number of cities: Paris and Los Angeles were the first two, and we are currently processing data on large and medium-size European cities, as commissioned by the European Research Centre. Our results and those of other European-based research teams (we are the only France-based one) will be presented at a workshop in Brussels on November 27th. We are also working on the network version of the trajectory method. We should have a pre-publication early in 2019 on this.

Also, the multidimensional, neural network analysis presented in Section II.C has been implemented on French cities other than Paris by a Master student, Sarah Soleiman, working as an intern in the *start-up* company *MeilleursAgents.com*. The company's main objective is to provide the best possible representation

---

[14] This is in the INFINITI framework: *INterFaces Interdisciplinaires, NumérIque et ThéorIque*.

[15] *Partenariat Hubert Curien*, funded by the French Foreign Office.

of the French real estate market. They have been testing our multidimensional analysis into the algorithm they are using to estimate house prices. Ms Soleiman has shown that it leads indeed to significantly better estimations. We will shortly be applying for a PhD funding[16] to develop this work further.

Still on an interdisciplinary note, I started a few months ago a collaboration with philologist Jean-Baptiste Camps (ENC/PSL). We are studying manuscript transmission through both numerical simulations and Markov chain models (in particular, *birth-and-death* processes). We explore the variety of manuscript tree patterns that emerge when key parameters in the transmission process are varied across their ranges, viz. *fecondity* and *decimation* rates. This should allow us to address a long-standing issue in philology: the predominance of root bifurcation (known as bifidity) in text genealogical trees (known as *stemmata*). It has been argued that the over-representation of bifid stemmata may be an artefact of the *common error* methods used by philologists to reconstruct text genealogies, and we are investigating whether the observed ubiquitous bifidity could simply be an endogeneous phenomenon resulting from the manuscript transmission process.

Coming now back to theoretical aspects of stochastic processes, let me mention some of the current projects on which I am working.

On convex hulls of Brownian and Lévy processes, a collaboration was initiated with Raphaël Lachièze-Rey (Univ. Paris-Descartes) after the 2018 Stochastic Geometry Days. We are working on second-order results regarding the number of facets on the boundary of Brownian convex hulls.

Work on residence times, with Sidney Redner, is also on-going, as well as work on a predator/prey foraging model, in further collaboration with Olivier Bénichou (CNRS/UPMC).

Most of my time though, in this part of my work, is devoted to another type of foraging model. This has been in collaboration with Paul Krapivsky (Univ. Boston). The model may be viewed as representing a foraging animal that consumes food at unit rate and lives in an environment with one unit of food per unit space. To keep things simple, consider a Brownian or Lévy particle on a line and say it survives as long as the *range* it has explored remains, at every instant, larger than the elapsed time. Writing $R(t)$ for the range at time t, we are computing the probability that $R(s) > s$ for all $s \leq t$. A variant is the one-sided problem where one considers only the supremum $M(t)$ instead of the range $R(t)$. This variant may be solved starting from existing results in the literature, but we also found a completely new method – using path decompositions and transformations. The full, double-sided problem is not amenable to the same technique, precisely because one needs to control *both* sides when dealing with the range, and not just the upper one. We are hoping to work on the double-sided problem with Pierre Vallois (Univ. Nancy), after this HDR thesis has been defended.

Belonging also to the more theoretical side of my research is the work we are doing with Antoine Lucquiaud on extensions of the Ballot theorem to so-called *subordinated* random walks. The question stemmed out of the trajectory convergence method that Madalina Olteanu and I developed. When it came to using

[16] Within the CIFRE framework (*Conventions Industrielles de Formation par la REcherche*).

random walk results in the analysis of the trajectories, we had to take into account the fact that these are formed by aggregating areal units of various sizes. This means that the trajectories do not progress one step at a time but rather they behave as subordinated random walks: there is an underlying "full trajectory" where one step corresponds *eg* to one individual or one housing unit, but the observed trajectory is a subsample, at certain steps only (steps corresponding to the sizes of the areal units).
With Antoine Lucquiaud, Madalina Olteanu and several other colleagues (Chiara Cammarota (KCL Londres), Aurélien Hazan (UPEC, Créteil), Alejandro Lage-Castellanos and Roberto Mulet (UH, La Havane)), we are also working on a theoretical model of the Schelling-type dynamics presented in Section II.A. We are trying to combine statistical physics tools (from spin glass models) with probability theoretic approaches. A Markov chain model of the standard Schelling dynamics was studied by one of my Master students (Matthieu Vert) a couple of years ago, and we are hoping to be able to adapt it to the new variant. Incidentally, the original Markov chain model is itself of interest, as on-going work with Annie Millet has shown: we are investigating the details of the thermodynamical limit in the context of this specific Markov chain (long time limit *vs* large system limit as opposed to taking both limits simultaneously).

Let the last lines of this conclusion be to warmly thank all my colleagues, students, and collaborators for their patience and their enduring goodwill in sharing their knowledge, their energy and their expertise. Research is a collective endeavour – and that plays no small part in making it such a beautiful craft.

REFERENCES

[1] Abramson, Josh, Jim Pitman, Nathan Ross, and Geronimo Uribe Bravo (2011), "Convex minorants of random walks and Lévy processes," Electron. Commun. Probab. **16**, no. 38, 423–434.

[2] Addario-Berry, L, and BA Reed (2008), "Ballot theorems, old and new," Horizons of combinatorics **17**, 9–35.

[3] Agmon, Noam (1984), "Residence times in diffusion processes," The Journal of Chemical Physics **81** (8), 3644–3647.

[4] Alberts, B, A. Johnson, and J. Lewis (2002), "Molecular biology of the cell," Chap. 3, 4th ed. (Garland Science, New York).

[5] Andersson, Eva K, and Bo Malmberg (2015), "Contextual effects on educational attainment in individualised, scalable neighbourhoods: Differences across gender and social class," Urban Studies **52** (12), 2117–2133.

[6] André, Désiré (1887), "Solution directe du probleme résolu par M. Bertrand," CR Acad. Sci. Paris **105** (436), 7.

[7] Applebaum, David (2009), *Lévy processes and stochastic calculus* (Cambridge University Press).

[8] Arcaute, Elsa, Erez Hatna, Peter Ferguson, Hyejin Youn, Anders Johansson, and Michael Batty (2015), "Constructing cities, deconstructing scaling laws," Journal of The Royal Society Interface **12** (102), 20140745.

[9] Ardestani, Babak Mahdavi, David O'Sullivan, and Peter Davis (2018), "A multi-scaled agent-based model of residential segregation applied to a real metropolitan area," Computers, Environment and Urban Systems **69**, 1 – 16.

[10] Arribas-Bel, Daniel, Peter Nijkamp, and Henk Scholten (2011), "Multidimensional urban sprawl in Europe: A self-organizing map approach," Computers, Environment and Urban Systems **35** (4), 263–275.

[11] Badel-Chagnon, A, J. Nessi, L. Buffat, and S. Hazout (1994), "Iso-depth contour map of a molecular surface," Journal of Molecular Graphics **12** (3), 162 – 168.

[12] Bailey, DF (1996), "Counting arrangements of 1's and -1's," Mathematics Magazine **69** (2), 128–131.

[13] Baldassarri, A, J. P. Bouchaud, I. Dornic, and C. Godrèche (1999), "Statistics of persistent events: An exactly soluble model," Physical Review E **59**, R20–R23.

[14] Baldassarri, Andrea, Francesca Colaiori, and Claudio Castellano (2003), "Average shape of a fluctuation: Universality in excursions of stochastic processes," Physical Review Letters **90**, 060601.

[15] Barber, M N (1983), in *Phase transitions and critical phenomena*, Vol. 8, edited by C. Domb and J. L. Lebowitz (New York: Academic Press) p. 145.

[16] Barbier, Émile (1887), "Généralisation du problème résolu par M. Bertrand," CR Acad. Sci. Paris **105** (1887), 407.

[17] Barndorff-Nielsen, O, and G. Baxter (1963), "Combinatorial lemmas in higher dimensions," Transactions of the American Mathematical Society **108** (2), 313–325.

[18] Barton, David E, and Colin L Mallows (1965), "Some aspects of the random sequence," The Annals of Mathematical Statistics , 236–260.

[19] Batty, Michael (1976), "Entropy in spatial aggregation," Geographical Analysis **8** (1), 1–21.

[20] Baxter, G (1961), "A combinatorial lemma for complex numbers," Ann. Math. Stat. **32** (3), 901.

[21] Ben-Naim, E, and P L Krapivsky (2014), "Persistence of random walk records," Journal of Physics A: Mathematical and Theoretical **47** (25), 255002.

[22] Ben-Naim, E, and P. L. Krapivsky (2014), "Slow kinetics of Brownian maxima," Physical Review Letters **113**, 030604.

[23] Ben-Naim, Eli, Paul L Krapivsky, and Julien Randon-Furling (2016), "Maxima of two random walks: universal statistics of lead changes," Journal of Physics A: Mathematical and Theoretical **49** (20), 205003.

[24] Benenson, I, I. Omer, and E. Hatna (2002), "Entity-based modeling of urban residential dynamics: the case of Yaffo, Tel Aviv," Environment and Planning B **29**, 491–512.

[25] Bénichou, O, C Chevalier, J Klafter, B Meyer, and R Voituriez (2010), "Geometry-controlled kinetics," Nature Chemistry **2** (6), 472–477.

[26] Bénichou, O, M. Coppey, M. Moreau, P.-H. Suet, and R. Voituriez (2005), "Optimal search strategies for hidden targets," Physical Review Letters **94**, 198101.

[27] Bénichou, O, and R. Voituriez (2014), "From first-passage times of random walks in confinement to geometry-controlled kinetics," Physics Reports **539** (4), 225 – 284.

[28] Bénichou, Olivier, and Jean Desbois (2008), "Exit and occupation times for Brownian motion on graphs with general drift and diffusion constant," Journal of Physics A: Mathematical and Theoretical **42** (1), 015004.

[29] Bertoin, J (1991), "Décomposition du mouvement brownien avec dérive en un minimum local par juxtaposition de ses excursions positives et negatives," Séminaire de Probabilités de Strasbourg **25**, 330–344.

[30] Bertoin, J (1993), "Splitting at the infimum and excursions in half-lines for random walks and Lévy processes," Stochastic Processes and their Applications **47** (1), 17–35.

[31] Bertoin, J, L. Chaumont, and J. Pitman (2003), "Path transformations of first passage bridges," Electronic Communications in Probability **8**, 155–166.

[32] Bertoin, J, and J. Pitman (1994), "Path transformations connecting Brownian bridge, excursion and meander," Bulletin des sciences mathématiques **118** (2), 147–166.

[33] Bertoin, Jean (1998), *Lévy processes*, Vol. 121 (Cambridge university press).

[34] Bertrand, Joseph (1887), "Solution d'un problème," CR Acad. Sci. Paris **105**, 369.

[35] Bhat, U, S Redner, and O Bénichou (2017), "Starvation dynamics of a greedy forager," Journal of Statistical Mechanics: Theory and Experiment **2017** (7), 073213.

[36] Biane, P (1986), "Relations entre pont et excursion du mouvement brownien réel," Ann. Inst. Henri Poincaré **22** (1), 1–7.

[37] Biane, P, and G. Letac (2011), "The mean perimeter of some random plane convex sets generated by a Brownian motion," J. Theor. Prob. **24** (2), 330–341.

[38] Binder, K (1987), "Theory of first-order phase transitions," Rep. Prog. Phys. **50**, 783–859.

[39] Bingham, NH, and RA Doney (1988), "On higher-dimensional analogues of the arc-sine law," Journal of Applied Probability **25** (1), 120–131.

[40] Biondo, A E, A. Pluchino, A. Rapisarda, and D. Helbing (2013), "Are random strategies more successful than technical ones?" Preprint: arXiv:q-fin/1303.4351.

[41] Biondo, A E, A. Rapisarda, and C. Garofalo (2011), "Efficient promotion strategies in hierarchical organizations," Physica A: Statistical Mechanics and its Applications **390** (20), 3496 – 3511.

[42] Blume, M, V. J. Emery, and R. B. Griffiths (1971), "Ising model for the λ transition and phase separation in He^3-He^4 mixtures," Physical Review A **4**, 1071–1077.

[43] de Bodt, Eric, Marie Cottrell, and Michel Verleysen (2002), "Statistical tools to assess the reliability of self-organizing maps," Neural Networks **15, 8-9**, 967–978.

[44] Boelaert, Julien, Laura Bendhaiba, Madalina Olteanu, and Nathalie Villa-Vialaneix (2014), "SOMbrero: An R Package for Numeric and Non-numeric Self-Organizing Maps," in *Advances in Self-Organizing Maps and Learning Vector Quantization: Proceedings of the 10th International Workshop, WSOM 2014*, edited by Thomas Villmann, Frank-Michael Schleif, Marika Kaden, and Mandy Lange (Springer International Publishing) pp. 219–228.

[45] Bourgeois, Nicolas, Marie Cottrell, Benjamin Déruelle, Stéphane Lamassé, and Patrick Letrémy (2015), "How to improve robustness in kohonen maps and display additional information in factorial analysis: application to text mining," Neurocomputing **147**, 120–135.

[46] Bourgeois, Nicolas, Marie Cottrell, Stéphane Lamassé, and Madalina Olteanu (2015), "Search for meaning through the study of co-occurrences in texts," in *International Work-Conference on Artificial Neural Networks* (Springer) pp. 578–591.

[47] Bray, Alan J, Satya N. Majumdar, and Grgory Schehr (2013), "Persistence and first-passage properties in nonequilibrium systems," Advances in Physics **62** (3), 225–361.

[48] Brown, Lawrence A, and Su-Yeul Chung (2006), "Spatial segregation, segregation indices and the geographical perspective," Population, Space and Place **12** (2), 125–143.

[49] Bruch, E E, and R. D. Mare (2006), "Neighborhood choice and neighborhood change," Amer. Jour. Sociol. **112** (3), 667–709.

[50] Busetti, Fabio (2017), "Quantile aggregation of density forecasts," Oxford Bulletin of Economics and Statistics **79** (4), 495–512.

[51] Caridi, I, J.P. Pinasco, N. Saintier, and P. Schiaffino (2017), "Characterizing segregation in the Schelling Voter model," Physica A: Statistical Mechanics and its Applications **487**, 125 – 142.

[52] Chandrasekhar, S (1949), "Brownian motion, dynamical friction, and stellar dynamics," Reviews of Modern Physics **21** (3), 383.

[53] Chassaing, P, and S. Janson (2001), "A Vervaat-like path transformation for the reflected Brownian bridge conditioned on its local time at 0," The Annals of Probability **29** (4), 1755–1779.

[54] Chaumont, L (1996), "Conditionings and path decompositions for Lévy processes," Stochastic processes and their applications **64** (1), 39–54.

[55] Chaumont, L (2000), "An extension of Vervaat's transformation and its consequences," Journal of Theoretical Probability **13** (1), 259–277.

[56] Chaumont, L, D.G. Holson, and M. Yor (2001), "Some consequences of the cyclic exchangeability property for exponential functionals of Lévy processes," in *Séminaire de Probabilités XXXV*, Lecture Notes in Mathematics, Vol. 1755, edited by J. Azéma, M. Émery, M. Ledoux, and M. Yor (Springer Berlin Heidelberg) pp. 334–347.

[57] Chaumont, Loïc (1994), *Processus de Lévy et conditionnement*, Ph.D. thesis (UPMC).

[58] Chaumont, Loïc, Ronald Doney, *et al.* (2005), "On lévy processes conditioned to stay positive." Electronic Journal of Probability **10**, 948–961.

[59] Chaumont, Loïc, and Gerónimo Uribe Bravo (2015), "Shifting processes with cyclically exchangeable increments at random," in *XI Symposium on Probability and Stochastic Processes* (Springer) pp. 101–117.

[60] Chechkin, Aleksei V, Ralf Metzler, Vsevolod Y Gonchar, Joseph Klafter, and Leonid V Tanatarov (2003), "First passage and arrival time densities for Lévy flights and the failure of the method of images," Journal of Physics A: Mathematical and General **36** (41), L537.

[61] Chung, K L (1976), "Excursions in Brownian motion," Ark. Math. **14**, 155–177.

[62] Chupeau, M, O. Bénichou, and S. N. Majumdar (2015), "Convex hull of a Brownian motion in confinement," Physical Review E **91** (5), 050104.

[63] Chupeau, M, O. Bénichou, and S. N. Majumdar (2015), "Mean perimeter of the convex hull of a random walk in a semi-infinite medium," Physical Review E **92** (2), 022145.

[64] Chvátal, Vasek (1979), "The tail of the hypergeometric distribution," Discrete Mathematics **25** (3), 285–287.

[65] Clark, W A V, and M. Fossett (2008), "Understanding the social context of the Schelling segregation model," PNAS **105** (11), 4109–4114.

[66] Clark, William, Madalina Olteanu, and Julien Randon-Furling (in prep.), .

[67] Clark, William A V, Eva Andersson, John Östh, and Bo Malmberg (2015), "A multiscalar analysis of neighborhood composition in Los Angeles, 2000–2010: A location-based approach to segregation and diversity," Annals of the Association of American Geographers **105** (6), 1260–1284.

[68] Claussen, Gunnar, Alexander K. Hartmann, and Satya N. Majumdar (2015), "Convex hulls of random walks: Large-deviation properties," Physical Review E **91**, 052104.

[69] Colaiori, Francesca, Andrea Baldassarri, and Claudio Castellano (2004), "Average trajectory of returning walks," Physical Review E **69** (4), 041105.

[70] Coleman, Ryan G, and Kim A. Sharp (2006), "Travel depth, a new shape descriptor for macromolecules: Application to ligand binding," Journal of Molecular Biology **362** (3), 441 – 458.

[71] Condamin, S, V Tejedor, and O Bénichou (2007), "Occupation times of random walks in confined geometries: From random trap model to diffusion-limited reactions," Physical Review E **76** (5), 050102.

[72] Cook, PR, and D. Marenduzzo (2009), "Entropic organization of interphase chromosomes," J. Cell. Biol. **188** (6), 825–834.

[73] Cornforth, D, D. G. Green, and D. Newth (2005), "Ordered asynchronous processes in multi-agent systems," Physica D **204**, 70–82.

[74] Cortese, Charles F, R Frank Falk, and Jack K Cohen (1976), "Further considerations on the methodological analysis of segregation indices," American Sociological Review , 630–637.

[75] Cottrell, Marie, Madalina Olteanu, Julien Randon-Furling, and Aurelien Hazan (2017), "Multidimensional urban segregation: an exploratory case study," in *12th International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM+)* (IEEE) pp. 1–7.

[76] Cowgill, Donald O, and Mary S Cowgill (1951), "An index of segregation based on block statistics," American Sociological Review **16** (6), 825–831.

[77] Cranston, M, P. Hsu, and P. March (1989), "Smoothness of the convex hull of planar Brownian motion," Annals Prob. **17** (1), 144.

[78] Daniels, HE (1982), "Sequential tests constructed from images," The Annals of Statistics , 394–400.

[79] Davydov, Yu (2012), "On convex hull of d-dimensional fractional Brownian motion ," Statistics and Probability Letters **82** (1), 37 – 39.

[80] De Gennes, P G (1976), "Dynamics of Entangled Polymer Solutions. I. The Rouse Model," Macromolecules **9** (4), 587–593.

[81] Deffuant, G, F. Amblard, G. Weisbuch, and T. Faure (2002), "How can extremism prevail? a study based on the relative agreement interaction mode," J. Artif. Soc. Social. Simul. **5** (4), 1.

[82] Derrida, Bernard, Vincent Hakim, and Reuven Zeitak (1996), "Persistent spins in the linear diffusion approximation of phase ordering and zeros of stationary gaussian processes," Physical Review Letters **77**, 2871–2874.

[83] Desbois, Jean (2007), "Occupation times for planar and higher dimensional Brownian motion," Journal of Physics A: Mathematical and Theoretical **40** (10), 2251.

[84] Doi, M, and S. F. Edwards (1988), *The theory of polymer dynamics*, International Series of Monographs on Physics (Oxford University Press).

[85] Doob, Joseph L (1949), "Heuristic approach to the Kolmogorov-Smirnov theorems," The Annals of Mathematical Statistics **20** (3), 393–403.

[86] Duncan, Otis Dudley, and Beverly Duncan (1955), "A methodological analysis of segregation indexes," American Sociological Review **20** (2), 210–217.

[87] Durrett, R T, and D. L. Iglehart (1977), "Functionals of Brownian meander and Brownian excursion," The Annals of Probability **5** (1), 130–135.

[88] Dybiec, B, E. Gudowska-Nowak, and P. Hänggi (2006), "Lévy-Brownian motion on finite intervals: Mean first passage time analysis," Physical Review E **73**, 046104.

[89] Edwards, S F (1965), "The statistical mechanics of polymers with excluded volume," Proc. Phys. Soc. **85** (4), 613.

[90] Einstein, A (1905), "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen," Annalen der Physik **322** (8), 549–560.

[91] El Bachir, M (1983), *L'enveloppe convexe du mouvement brownien*, Ph.D. thesis (Université Paul Sabatier, Toulouse, France).

[92] Eldan, Ronen (2014), "Volumetric properties of the convex hull of an n-dimensional Brownian motion," Electron. J. Probab. **19**, no. 45, 1–34.

[93] Evans, S N (1985), "On the Hausdorff dimension of Brownian cone points," Math. Proc. Camb. Philos. Soc. **98**, 343–353.

[94] Farley, Reynolds, and Karl E Taeuber (1968), "Population trends and residential segregation since 1960," Science **159** (3818), 953–956.

[95] Feitosa, Flávia F, Gilberto Camara, Antonio Miguel Vieira Monteiro, Thomas Koschitzki, and Marcelino PS Silva (2007), "Global and local spatial indices of urban segregation," International Journal of Geographical Information Science **21** (3), 299–323.

[96] Feitosa, Flávia F, Quang Bao Le, Paul LG Vlek, Antonio Miguel V Monteiro, and Roberta Rosemback (2012), "Countering urban segregation in brazilian cities: policy-oriented explorations using agent-based simulation," Environment and Planning B: Planning and Design **39** (6), 1131–1150.

[97] Feller, W (1971), *An introduction to probability theory and its applications*, Wiley Series in Probability and Mathematical Statistics (Wiley).

[98] Ferraro, M, and L Zaninetti (2001), "Number of times a site is visited in two-dimensional random walks," Physical

Review E **64** (5), 056107.

[99] Ferraro, M, and L Zaninetti (2004), "Statistics of visits to sites in random walks," Physica A: Statistical Mechanics and its Applications **338** (3-4), 307–318.

[100] Ferraro, M, and L Zaninetti (2006), "Mean number of visits to sites in levy flights," Physical Review E **73** (5), 057102.

[101] Filet, Clara (2017), "An attempt to estimate the impact of the spread of economic flows on Latenian urbanization," Frontiers in Digital Humanities **3**, 10.

[102] Fisher, Ronald Aylmer (1925), *Statistical Methods for Research Workers* (Genesis Publishing Pvt Ltd).

[103] Fisher, Ronald Aylmer (1945), "A new test for 2×2 tables," Nature **156** (3961), 388.

[104] Fitzsimmons, PJ, and R.K. Getoor (1995), "Occupation time distributions for Lévy bridges and excursions," Stochastic Processes and their Applications **58** (1), 73 – 89.

[105] Fossett, M (2006), "Ethnic preferences, social distance dynamics, and residential segregation: Theoretical explorations using simulation analysis," J. Math. Sociol. **30** (185–274).

[106] Fossett, Mark (2017), *New methods for measuring and analyzing segregation*, Vol. 42 (Springer).

[107] Fotheringham, A Stewart, and David WS Wong (1991), "The modifiable areal unit problem in multivariate statistical analysis," Environment and Planning A **23** (7), 1025–1044.

[108] Fougre, F, and J. Desbois (1993), "Moments of inertia and the shapes of Brownian paths," J. Phys. A: Math. Gen. **26**, 7253–7262.

[109] Galambos, J (1987), *The asymptotic theory of extreme order statistics* (R.E. Krieger Pub. Co.).

[110] Gauvin, L (2010), *Modélisation de systèmes socio-économiques l'aide des outils de physique statistique*, Ph.D. thesis (Université Pierre et Marie Curie).

[111] Gauvin, L, J. Vannimenus, and J.-P. Nadal (2009), "Phase diagram of a Schelling segregation model," The European Physical Journal B **70**, 293–304.

[112] Genest, Christian (1992), "Vincentization revisited," The Annals of Statistics , 1137–1142.

[113] Genest, Christian, and James V Zidek (1986), "Combining probability distributions: A critique and an annotated bibliography," Statistical Science , 114–135.

[114] Glaeser, Edward L, Jed Kolko, and Albert Saiz (2001), "Consumer city," Journal of Economic Geography **1** (1), 27–50.

[115] Goel, NS, and N. Richter-Dyn (2004), *Stochastic Models in Biology* (Blackburn Press).

[116] Grauwin, S, F. Goffette-Nagot, and P. Jensen (2012), "Dynamic models of residential segregation: an analytical solution," J. Publ. Econ. **96**, 124–141.

[117] Haber, C, S.A. Ruiz, and D. Wirtz (2000), "Shape anisotropy of a single random-walk polymer," Proceedings of the National Academy of Sciences **97** (20), 10792–10795.

[118] Harris, Richard, and Dewi Owen (2017), "Implementing a multilevel index of dissimilarity in R with a case study of the changing scales of residential ethnic segregation in England and Wales," Environment and Planning B: Urban Analytics and City Science , 2399808317748328.

[119] Hatna, E, and I. Benenson (2012), "The Schelling model of ethnic residential dynamics: beyond the integrated-segregated dichotomy of patterns," J. Artif. Soc. Social. Simul. **15** (1), 6.

[120] Havlin, S, S.V. Buldyrev, H.E. Stanley, and G.H. Weiss (1991), "Probability distribution of the interface width in surface roughening: analogy with a Lévy flight," Journal of Physics A: Mathematical and General **24** (16), L925.

[121] Henkel, M, H. Hinrichsen, M. Pleimling, and S. Lübeck (2011), *Non-Equilibrium Phase Transitions: Volume 2: Ageing and Dynamical Scaling Far from Equilibrium*, Theoretical and Mathematical Physics (Springer).

[122] Holm, L, and C. Sander (1993), "Protein structure comparison by alignment of distance matrices," Journal of Molecular Biology **233** (1), 123 – 138.

[123] Hong, Seong-Yun, David O'Sullivan, and Yukio Sadahiro (2014), "Implementing spatial segregation measures in R," PloS one **9** (11), e113767.

[124] Iceland, John (2004), "The multigroup entropy index," US Census Bureau **31**, 2006.

[125] Iglehart, D L (1975), "Conditioned limit theorems for random walks," Advances in Applied Probability **7** (2), 237237.

[126] INSEE, (2017), "IRIS – Definition," `https://www.insee.fr/en/metadonnees/definition/c1523`, accessed 2017-11-05.

[127] James, David R, and Karl E Taeuber (1985), "Measures of segregation," Sociological Methodology **15**, 1–32.

[128] Janakiraman, Deepika, and K. L. Sebastian (2014), "Unusual eigenvalue spectrum and relaxation in the Lévy-Ornstein-Uhlenbeck process," Physical Review E **90**, 040101.

[129] Jelinski, Dennis E, and Jianguo Wu (1996), "The modifiable areal unit problem and implications for landscape ecology," Landscape Ecology **11** (3), 129–140.

[130] Jones, F L (1985), "Simulation models of group segregation," Aust. New. Zealand J. Sociol. **21** (3), 431–444.

[131] Jones, Kelvyn, Ron Johnston, David Manley, Dewi Owen, and Chris Charlton (2015), "Ethnic residential segregation: A multilevel, multigroup, multiscale approach exemplified by London in 2011," Demography **52** (6), 1995–2019.

[132] Kabluchko, Z, V. Vysotsky, and D. Zaporozhets (2017), "Convex hulls of random walks: Expected number of faces and face probabilities," Advances in Mathematics **320**, 595–629.

[133] Kabluchko, Zakhar, Vladislav Vysotskiy, and Dmitry Zaporozhets (2018), "A multidimensional analogue of the arcsine law for the number of positive terms in a random walk," Bernoulli (to appear).

[134] Kac, M (1954), "Toeplitz matrices, translation kernels and a related problem in probability theory," Duke Mathematical Journal **21** (3), 501–509.

[135] Kallenberg, O (1999), "Ballot theorems and sojourn laws for stationary processes," Annals Prob. **27** (4), 2011–2019.

[136] Kampf, J, G. Last, and I. Molchanov (2012), "On the convex hull of symmetric stable processes," Proc. Amer. Math. Soc. **140**, 2527–2535.

[137] Kinney, JR (1966), "Convex hull of Brownian motion in d-dimensions," Isr. J. Math. **4** (2), 139–143.

[138] Knight, F B (1963), "Random walks and a sojourn density process of brownian motion," Transactions of the American Mathematical Society **109** (1), 56–86.

[139] Knight, F B (1996), "The uniform law for exchangeable and Lévy process bridges," Astérisque **236**, 171–188.

[140] Koehler, G, and J. Skvoretz (2010), "Residential segregation in university housing: The mathematics of preferences," Social Science Research **39**, 14–24.

[141] Kohonen, T (2012), *Self-Organizing Maps*, Springer Series in Information Sciences (Springer Berlin Heidelberg).

[142] Kohonen, Teuvo (1982), "Self-organized formation of topologically correct feature maps," Biological Cybernetics **43** (1), 59–69.

[143] Koren, T, A.V. Chechkin, and J. Klafter (2007), "On the first passage time and leapover properties of lvy motions," Physica A: Statistical Mechanics and its Applications **379** (1), 10 – 22.

[144] Koren, Tal, Michael Lomholt, Aleksei Chechkin, Joseph Klafter, and Ralf Metzler (2007), "Leapover lengths and first passage time statistics for Lévy flights," Physical Review Letters **99**, 160602.

[145] Koscielny-Bunde, Eva, Armin Bunde, Shlomo Havlin, H. Roman, Yair Goldreich, and Hans-Joachim Schellnhuber (1998), "Indication of a universal persistence law governing atmospheric variability," Physical Review Letters **81**, 729–732.

[146] Krapivsky, Paul, Sidney Redner, and Eli Ben-Naim (2010), *A kinetic view of statistical physics* (Cambridge University Press).

[147] Leckie, George, and Harvey Goldstein (2015), "A multilevel modelling approach to measuring changing patterns of ethnic composition and segregation among London secondary schools, 2001–2010," Journal of the Royal Statistical Society: Series A (Statistics in Society) **178** (2), 405–424.

[148] Leckie, George, Rebecca Pillinger, Kelvyn Jones, and Harvey Goldstein (2012), "Multilevel modeling of social segregation," Journal of Educational and Behavioral Statistics **37** (1), 3–30.

[149] Lee, Kyu-Hwan, and Se-jin Oh (2016), "Catalan triangle numbers and binomial coefficients," arXiv preprint arXiv:1601.06685.

[150] Lee, M, P. Lloyd, X. Zhang, J. M. Schallhorn, K. Sugimoto, A. G. Leach, G. Sapiro, and K. N. Houk (2006), "Shapes of antibody binding sites: qualitative and quantitative analyses based on a geomorphic classification scheme," The Journal of Organic Chemistry **71** (14), 5082–5092.

[151] Lévy, P (1939), "Sur certains processus stochastiques homogènes," Compositio Mathematica **7**, 283–339.

[152] Lévy, P (1948), *Processus stochastiques et mouvement brownien* (Gauthiers-Villars, Paris).

[153] Louf, Rémi, and Marc Barthelemy (2016), "Patterns of residential segregation," PLOS One **11** (6), e0157476.

[154] Lucquiaud, Antoine, Julien Randon-Furling, and Madalina Olteanu (in prep.), .

[155] Luković, Mirko, Theo Geisel, and Stephan Eule (2013), "Area and perimeter covered by anomalous diffusion processes," New Jour. Phys. **15** (6), 063034.

[156] Lupu, T, J. Pitman, and W. Tang (2015), "The Vervaat transform of Brownian bridges and Brownian motion," Electronic Journal of Probability **20**.

[157] Majumdar, S N, and A. Comtet (2004), "Exact maximal height distribution of fluctuating interfaces," Physical Review Letters **92**, 225501.

[158] Majumdar, S N, A. Comtet, and J. Randon-Furling (2010), "Random Convex Hulls and Extreme-Value Statistics," J. Stat. Phys. **138** (6), 955–1009.

[159] Majumdar, Satya, Alberto Rosso, and Andrea Zoia (2010), "Hitting probability for anomalous diffusion processes," Physical Review Letters **104**, 020602.

[160] Majumdar, Satya, Clément Sire, Alan Bray, and Stephen Cornell (1996), "Nontrivial exponent for simple diffusion," Physical Review Letters **77**, 2867–2870.

[161] Majumdar, Satya, and Robert Ziff (2008), "Universal record statistics of random walks and lévy flights," Physical Review Letters **101**, 050601.

[162] Majumdar, Satya N (1999), "Persistence in nonequilibrium systems," Current Science **77** (3), 370.

[163] Majumdar, Satya N, and Jean-Philippe Bouchaud (2008), "Optimal time to sell a stock in the BlackScholes model: comment on Thou shalt buy and hold, by A. Shiryaev, Z. Xu and X.Y. Zhou," Quantitative Finance **8** (8), 753–760.

[164] Manley, David, Ron Johnston, Kelvyn Jones, and Dewi Owen (2015), "Macro-, meso-and microscale segregation: Modeling changing ethnic residential patterns in Auckland, New Zealand, 2001–2013," Annals of the Association of American Geographers **105** (5), 951–967.

[165] Massey, Douglas S, and Nancy A Denton (1988), "The dimensions of residential segregation," Social forces **67** (2), 281–315.

[166] Maximon, Leonard C (2003), "The dilogarithm function for complex argument," Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **459** (2039), 2807–2819.

[167] Meier, R, F. Ackermann, G. Herrmann, S. Posch, and G. Sagerer (1995), "Segmentation of molecular surfaces based on their convex hull," in *Proceedings, International Conference on Image Processing*, Vol. 3, pp. 552–555.

[168] Merikoski, J, J. Maunuksela, M. Myllys, J. Timonen, and M. Alava (2003), "Temporal and spatial persistence of combustion fronts in paper," Physical Review Letters **90**, 024501.

[169] Meyer-Ortmanns, H (2003), "Immigration, integration and ghetto formation," Int. J. Mod Phys. C **14**, 311.

[170] Mills, Barbara J (2017), "Social Network Analysis in Archaeology," Annual Review of Anthropology **46** (1).

[171] Molchanov, Ilya, and Florian Wespi (2016), "Convex hulls of Lévy processes," Electron. Commun. Probab. **21**, 69.

[172] Montroll, Elliott W, and George H Weiss (1965), "Random walks on lattices. II," Journal of Mathematical Physics **6** (2), 167–181.

[173] Mörters, Peter, and Yuval Peres (2010), *Brownian motion*, Vol. 30 (Cambridge University Press).

[174] de Mulatier, Clélia, Alberto Rosso, and Gregory Schehr (2013), "Asymmetric Lévy flights in the presence of absorbing boundaries," Journal of Statistical Mechanics: Theory and Experiment **2013** (10), P10006.

[175] Müller, K, C. Schulze, and D. Stauffer (2008), "Inhomogeneous and self-organised temperature in Schelling-Ising model," Int. J. Mod. Phys. C **19**, 385.

[176] Oerding, Klaus, Stephen Cornell, and Alan Bray (1997), "Non-Markovian persistence and nonequilibrium critical dynamics," Physical Review E **56**, R25–R28.

[177] Openshaw, Stan (1984), *The modifiable areal unit problem* (University of East Anglia).

[178] Openshaw, Stan, and Peter J Taylor (1979), "A million or so correlation coefficients: three experiments on the modifiable areal unit problem," Statistical applications in the spatial sciences **21**, 127–144.

[179] Östh, John, William A V Clark, and Bo Malmberg (2015), "Measuring the scale of segregation using k-nearest neighbor aggregates," Geographical Analysis **47** (1), 34–49.

[180] Parashar, Rishi, Daniel O'Malley, and John Cushman (2008), "Mean first-passage time for superdiffusion in a slit pore

with sticky boundaries," Physical Review E **78**, 052101.

[181] Paul, R, and G. Schehr (2005), "Non-Markovian persistence in the diluted Ising model at criticality," EPL (Europhysics Letters) **72** (5), 719.

[182] Penland, Cécile, and Brian D Ewald (2008), "On modelling physical systems with stochastic models: diffusion versus Lévy processes," Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **366** (1875), 2455–2474.

[183] Perlekar, Prasad, Samriddhi Ray, Dhrubaditya Mitra, and Rahul Pandit (2011), "Persistence problem in two-dimensional fluid turbulence," Physical Review Letters **106**, 054501.

[184] Pluchino, A, C. Garofalo, A. Rapisarda, S. Spagano, and M. Caserta (2011), "Accidental politicians: How randomly selected legislators can improve parliament efficiency," Physica A: Statistical Mechanics and its Applications **390** (2122), 3944 – 3954.

[185] Pluchino, A, A. Rapisarda, and C. Garofalo (2010), "The Peter principle revisited: A computational study," Physica A: Statistical Mechanics and its Applications **389** (3), 467 – 472.

[186] Project, OpenStreetMap (2012), "Intégration des stations du réseau ferré RATP," Accessed March 2017.

[187] Randon-Furling, J (2009), *Statistiques d'extrêmes du mouvement brownien et applications*, Ph.D. thesis (Université Paris-Sud / Paris-Saclay).

[188] Randon-Furling, J (2013), "Convex hull of n planar Brownian paths: an exact formula for the average number of edges," J. Phys. A: Math. Theor. **46**, 015004.

[189] Randon-Furling, J (2014), "Universality and time-scale invariance for the shape of planar Lévy processes," Physical Review E **89** (5), 052112.

[190] Randon-Furling, J (2015), "From Markovian to non-Markovian persistence exponents," EPL (Europhysics Letters) **109** (4), 40015.

[191] Randon-Furling, J, S. N. Majumdar, and A. Comtet (2009), "Convex hull of n planar Brownian motions: Exact results and an application to ecology," Physical Review Letters **103**, 140602.

[192] Randon-Furling, J, and S Redner (2018), "Residence time near an absorbing set," Journal of Statistical Mechanics: Theory and Experiment **2018** (10), 103205.

[193] Randon-Furling, J, and F. Wespi (2017), "Facets on the convex hull of d-dimensional Brownian and Lévy motion," Physical Review E **95** (3), 032129.

[194] Ray, Daniel, *et al.* (1963), "Sojourn times of diffusion processes," Illinois Journal of mathematics **7** (4), 615–630.

[195] Reardon, Sean F, and Glenn Firebaugh (2002), "Measures of Multigroup Segregation," Sociological Methodology **32** (1), 33–67.

[196] Reardon, Sean F, and David OSullivan (2004), "Measures of spatial segregation," Sociological methodology **34** (1), 121–162.

[197] Redner, S (2001), *A Guide to First-Passage Processes* (Cambridge University Press).

[198] Reymbaut, Alexis, Satya N Majumdar, and Alberto Rosso (2011), "The convex hull for a random acceleration process in two dimensions," J. Phys. A: Math. Theor. **44** (41), 415001.

[199] Risau-Gusman, S, S. Ibáñez, and S. Bouzat (2013), "Directed transport induced by α-stable Lévy noises in weakly asymmetric periodic potentials," Physical Review E **87**, 022105.

[200] Rogers, T, and A. McKane (2012), "Jamming and pattern formation in models of segregation," Physical Review E **85**, 041136.

[201] Rogers, T, and A. J. McKane (2011), "A unified framework for Schelling's model of segregation," Journal of Statistical Mechanics: Theory and Experiment **7**, 6.

[202] Rudnick, J, and G. Gaspari (1987), "The shapes of random walks," Science **237** (4813), 384–389.

[203] Sakoda, James M (1971), "The checkerboard model of social interaction," Journal of Mathematical Sociology **1** (1), 119–132.

[204] Schelling, T C (1971), "Dynamic model of segregation," J. Math. Sociol. **1**, 143–186.

[205] Schelling, T C (1978), *Micromotives and Macrobehavior* (New York: W. W. Norton).

[206] Schwartz, Joseph, and Christopher Winship (1980), "The welfare approach to measuring inequality," Sociological methodology **11**, 1–36.

[207] Simes, R John (1986), "An improved Bonferroni procedure for multiple tests of significance," Biometrika **73** (3), 751–754.

[208] Sire, Clément (2007), "Probability distribution of the maximum of a smooth temporal signal," Physical Review Letters **98**, 020601.

[209] Sloane, Neil JA (accessed 12 June 2018), "The on-line encyclopedia of integer sequences,".

[210] Sparre Andersen, E (1953), "On sums of symmetrically dependent random variables," Scandinavian Actuarial Journal **1953** (sup1), 123–138.

[211] Sparre Andersen, E (1954), "On the fluctuations of sums of random variables," Mathematica Scandinavica , 263–285.

[212] Sparre Andersen, E (1955), "On the fluctuations of sums of random variables II," Mathematica Scandinavica , 195–223.

[213] Spitzer, F (1956), "A combinatorial lemma and its application to probability theory," Transactions of the American Mathematical Society **82** (2), 323–339.

[214] Spitzer, F, and H. Widom (1961), "The circumference of a convex polygon," Proceedings of the American Mathematical Society **12** (3), 506–509.

[215] Spitzer, Frank (1964), *Principles of random walk.* (D. Van Nostrand, Princeton NJ).

[216] Stanley, Richard P (2015), *Catalan numbers* (Cambridge University Press).

[217] Stanley, Richard P, and Sergey Fomin (1999), *Enumerative Combinatorics*, Cambridge Studies in Advanced Mathematics, Vol. 2 (Cambridge University Press).

[218] Stauffer, D, and S. Solomon (2007), "Ising, Schelling and self-organising segregation," The European Physical Journal B **57**.

[219] Stoker, Thomas M (1984), "Completeness, distribution restrictions, and the form of aggregate functions," Econometrica: Journal of the Econometric Society , 887–907.

[220] Stout, M, J. Bacardit, J. D. Hirst, and N. Krasnogor (2008), "Prediction of recursive convex hull class assignments for protein residues," Bioinformatics **24** (7), 916.

[221] Sumour, M A, A. H. El-Astal, M. A. Radwan, and M. M. Shabat (2008), "Urban segregation with cheap and expensive residences," Int. J. Mod Phys. C **19** (4), 637–645.

[222] Takács, L (1980), "Expected perimeter length," Am. Math. Monthly **87**, 142.

[223] Takács, Lajos (1962), "Ballot problems," Probability Theory and Related Fields **1** (2), 154–158.

[224] Takács, Lajos (1997), "On the ballot theorems," in *Advances in combinatorial methods and applications to probability and statistics* (Springer) pp. 97–114.

[225] Theil, Henri, and Anthony J Finizza (1971), *A note on the measurement of racial integration of schools by means of informational concepts* (Taylor & Francis).

[226] US Census Bureau, (1994), "Geographic areas reference manual,".

[227] Van Kampen, Nicolaas Godfried (1992), *Stochastic processes in physics and chemistry* (Elsevier).

[228] Vervaat, W (1979), "A relation between Brownian bridge and Brownian excursion," The Annals of Probability , 143–149.

[229] Vinković, D, and A. Kirman (2006), "A physical analogue of the Schelling model," PNAS **103** (51), 19261–19265.

[230] Viswanathan, GM, M.G.E. da Luz, E.P. Raposo, and H.E. Stanley (2011), *The Physics of Foraging: An Introduction to Random Searches and Biological Encounters* (Cambridge University Press).

[231] Wang, Yong, Ling-Yun Wu, Xiang-Sun Zhang, and Luonan Chen (2006), "Automatic classification of protein structures based on convex hull representation by integrated neural network," in *Theory and Applications of Models of Computation: Third International Conference, TAMC 2006, Beijing, China, May 15-20, 2006. Proceedings*, edited by Jin-Yi Cai, S. Barry Cooper, and Angsheng Li (Springer Berlin Heidelberg, Berlin, Heidelberg) pp. 505–514.

[232] Wei, Chunzhu, Pablo Cabrera-Barona, and Thomas Blaschke (2016), "Local geographic variation of public services inequality: does the neighborhood scale matter?" International journal of environmental research and public health **13** (10), 981.

[233] Westwater, MJ (1980), "On Edwards' model for long polymer chains," Comm. Math. Phys. **72** (2), 103–205.

[234] White, Michael J (1983), "The measurement of spatial segregation," American Journal of Sociology **88** (5), 1008–1018.

[235] Whitworth, William Allen (1878), "Arrangements of m things of one sort and n things of another sort, under certain conditions of priority," Messenger of Math **8**, 105–114.

[236] Whitworth, William Allen (1886), *Choice and Chance*, 4th ed. (Deighton Bell).

[237] Wilson, J Anthony, Andreas Bender, Taner Kaya, and Paul A. Clemons (2009), "Alpha shapes applied to molecular shape characterization exhibit novel properties compared to established shape descriptors," Journal of Chemical Information and Modeling **49** (10), 2231–2241.

[238] Worton, BJ (1987), "A review of models of home range for animal movement," Ecological Modelling **38** (34), 277 – 298.

[239] Zagier, Don (2007), "The dilogarithm function," in *Frontiers in number theory, physics, and geometry II* (Springer) pp. 3–65.

[240] Zeng, Lingzao, and Bohou Xu (2010), "Effects of asymmetric Lvy noise in parameter-induced aperiodic stochastic resonance," Physica A: Statistical Mechanics and its Applications **389** (22), 5128 – 5136.

[241] Zumofen, G, and J. Klafter (1995), "Absorbing boundary in one-dimensional anomalous transport," Physical Review E **51**, 2805–2814.

APPENDIX

Relation Between Generating Functions in Section I.D.2

Using the geometric constraints in Eq. 82, the generating function $\mathcal{G}(x,y,z)$ can be re-expressed as

$$\begin{aligned}\mathcal{G}(x,y,z) &= \sum_{n=1}^{\infty}\sum_{\ell=1}^{n}\sum_{k=\ell}^{n}\mathcal{P}\,(n,k,\ell)\;x^n\,y^k\,z^\ell \\ &= \sum_{n=1}^{\infty}\sum_{\ell=1}^{n}\sum_{k=\ell}^{n}P(\ell)\,\mathbf{P}\,(n,k|\ell)\;x^n\,y^k\,z^\ell\,,\end{aligned} \tag{99}$$

where we use $\mathbf{P}(\dots|\dots)$ to denote the conditional joint probability.

Next we write $\mathbf{P}\,(n,k|\ell)$ in terms of the variables m_i and j_i (see Fig. 20):

$$\begin{aligned}\mathcal{G}(x,y,z) &= \sum_{n=1}^{\infty}\sum_{\ell=1}^{n}\sum_{k=\ell}^{n}\sum_{\substack{j_1+\cdots+j_\ell=k-\ell\\ m_1+\cdots+m_\ell=n-\ell\\ j_i\le m_i}}P(\ell)\,\mathbf{P}\,(m_1,\dots,m_\ell,j_1,\dots,j_\ell|\ell)\;x^{m_1+\dots m_\ell+\ell}\,y^{j_1+\dots j_\ell+\ell}\,z^\ell \\ &= \sum_{n=1}^{\infty}\sum_{\ell=1}^{n}\sum_{\substack{m_1+\cdots+m_\ell=n-\ell\\ 0\le j_i\le m_i}}P(\ell)\,\mathbf{P}\,(m_1,\dots,m_\ell,j_1,\dots,j_\ell|\ell)\;x^{m_1+\dots m_\ell+\ell}\,y^{j_1+\dots j_\ell+\ell}\,z^\ell \\ &= \sum_{\ell\ge1}P(\ell)\sum_{n\ge\ell}\sum_{\substack{m_1+\cdots+m_\ell=n-\ell\\ 0\le j_i\le m_i}}\mathbf{P}\,(m_1,\dots,m_\ell,j_1,\dots,j_\ell|\ell)\;x^{m_1+\dots m_\ell+\ell}\,y^{j_1+\dots j_\ell+\ell}\,z^\ell \\ &= \sum_{\ell\ge1}P(\ell)\sum_{m_1,\dots,m_\ell\ge0}\sum_{j_1,\dots,j_\ell=0}^{m_1,\dots,m_\ell}\left[\prod_{i=1}^{\ell}\mathbf{P}\,(m_i,j_i|\ell)\;x^{m_i}\,y^{k_i}\right](xyz)^\ell \\ &= \sum_{\ell\ge1}P(\ell)\,(xyz)^\ell\sum_{m_1,\dots,m_\ell\ge0}\sum_{j_1,\dots,j_\ell=0}^{m_1,\dots,m_\ell}\left[\prod_{i=1}^{\ell}\mathbf{P}\,(m_i,j_i|\ell)\;x^{m_i}\,y^{j_i}\right] \\ &= \sum_{\ell\ge1}P(\ell)\,(xyz)^\ell\prod_{i=1}^{\ell}\left[\sum_{m_i\ge0}\sum_{0\le j_i\le m_i}\mathbf{P}\,(m_i,j_i|\ell)\;x^{m_i}\,y^{j_i}\right].\end{aligned} \tag{100}$$

We now use the fact that the random walk is a Markov process, which implies that $\mathbf{P}\,(m_i,j_i|\ell) = \mathcal{P}\,(m_i,j_i) = A(m_i,j_i)/2^{2m_i}$. Therefore

$$\sum_{m_i\ge0}\sum_{0\le k_i\le m_i}\mathbf{P}\,(m_i,j_i|\ell)\;x^{m_i}\,y^{j_i} = G(x,y)\,. \tag{101}$$

Note that $G(x,y)$, as defined in Eq. 81, appears here and not $g(x,y)$ because upon starting from $x=2$, when coming from $x=1$, the path is not conditioned to immediately move to $x=3$. In fact, the path is allowed to return immediately to $x=1$, as reflected in the fact that, for the ith excursion, m_i may be 0.

Eq. 100 becomes

$$\mathcal{G}(x,y,z) = \sum_{\ell\ge1}P(\ell)\,(xyz)^\ell\prod_{1\le i\le\ell}[G(x,y)] = \sum_{\ell\ge1}P(\ell)\;[x\,y\,z\,G(x,y)]^\ell\,. \tag{102}$$

This is Eq. 84 in the main text.

SUPPLEMENTARY MATERIAL